

Article

Some Dissimilarity Measures of Branching Processes and Optimal Decision Making in the Presence of Potential Pandemics

Niels B. Kammerer ¹ and Wolfgang Stummer ^{2,*}

¹ Königinstrasse 75, 80539 Munich, Germany; nielskammerer@gmx.de

² Department of Mathematics, University of Erlangen–Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany

* Correspondence: stummer@math.fau.de

Received: 26 June 2020; Accepted: 28 July 2020; Published: 8 August 2020



Abstract: We compute exact values respectively bounds of dissimilarity/distinguishability measures—in the sense of the Kullback–Leibler information distance (relative entropy) and some transforms of more general power divergences and Renyi divergences—between two competing discrete-time *Galton–Watson branching processes with immigration* GWI for which the offspring as well as the immigration (importation) is arbitrarily Poisson-distributed; especially, we allow for arbitrary type of extinction-concerning criticality and thus for non-stationarity. We apply this to optimal decision making in the context of the spread of potentially pandemic infectious diseases (such as e.g., the current COVID-19 pandemic), e.g., covering different levels of dangerousness and different kinds of intervention/mitigation strategies. Asymptotic distinguishability behaviour and diffusion limits are investigated, too.

Keywords: Galton–Watson branching processes with immigration; Hellinger integrals; power divergences; Kullback–Leibler information distance/divergence; relative entropy; Renyi divergences; epidemiology; COVID-19 pandemic; Bayesian decision making; INARCH(1) model; GLM model; Bhattacharyya coefficient/distance

Contents

1	Introduction	3
2	The Framework and Application Setups	5
2.1	Process Setup	5
2.2	Connections to Time Series of Counts	6
2.3	Applicability to Epidemiology	8
2.4	Information Measures	12
2.5	Decision Making under Uncertainty	15
2.6	Asymptotical Distinguishability	19
3	Detailed Recursive Analyses of Hellinger Integrals	21
3.1	A First Basic Result	21
3.2	Some Useful Facts for Deeper Analyses	25
3.3	Detailed Analyses of the Exact Recursive Values, i.e., for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$	27
3.4	Some Preparatory Basic Facts for the Remaining Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$	29
3.5	Lower Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$	31
3.6	Goals for Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$	32

3.7	Upper Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},2} \times]0, 1[$	34
3.8	Upper Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},3a} \times]0, 1[$	35
3.9	Upper Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},3b} \times]0, 1[$	36
3.10	Upper Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},3c} \times]0, 1[$	37
3.11	Upper Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},4a} \times]0, 1[$	37
3.12	Upper Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},4b} \times]0, 1[$	37
3.13	Concluding Remarks on Alternative Upper Bounds for all Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times]0, 1[$	37
3.14	Intermezzo 1: Application to Asymptotical Distinguishability	38
3.15	Intermezzo 2: Application to Decision Making under Uncertainty	39
3.15.1	Bayesian Decision Making	39
3.15.2	Neyman-Pearson Testing	41
3.16	Goals for Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus [0, 1])$	41
3.17	Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},2} \times (\mathbb{R} \setminus [0, 1])$	44
3.18	Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},3a} \times (\mathbb{R} \setminus [0, 1])$	45
3.19	Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},3b} \times (\mathbb{R} \setminus [0, 1])$	46
3.20	Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},3c} \times (\mathbb{R} \setminus [0, 1])$	47
3.21	Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},4a} \times (\mathbb{R} \setminus [0, 1])$	47
3.22	Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},4b} \times (\mathbb{R} \setminus [0, 1])$	48
3.23	Concluding Remarks on Alternative Lower Bounds for all Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus [0, 1])$	48
3.24	Upper Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus [0, 1])$	48
4	Power Divergences of Non-Kullback-Leibler-Information-Divergence Type	49
4.1	A First Basic Result	49
4.2	Detailed Analyses of the Exact Recursive Values of $I_{\lambda}(\cdot \cdot)$, i.e., for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{NI}} \cup \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus \{0, 1\})$	51
4.3	Lower Bounds of $I_{\lambda}(\cdot \cdot)$ for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times]0, 1[$	52
4.4	Upper Bounds of $I_{\lambda}(\cdot \cdot)$ for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times]0, 1[$	53
4.5	Lower Bounds of $I_{\lambda}(\cdot \cdot)$ for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus [0, 1])$	53
4.6	Upper Bounds of $I_{\lambda}(\cdot \cdot)$ for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus [0, 1])$	54
4.7	Applications to Bayesian Decision Making	55
5	Kullback-Leibler Information Divergence (Relative Entropy)	55
5.1	Exact Values Respectively Upper Bounds of $I(\cdot \cdot)$	55
5.2	Lower Bounds of $I(\cdot \cdot)$ for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1})$	56
5.3	Applications to Bayesian Decision Making	58
6	Explicit Closed-Form Bounds of Hellinger Integrals	59
6.1	Principal Approach	59
6.2	Explicit Closed-Form Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{NI}} \cup \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus \{0, 1\})$	63
6.3	Explicit Closed-Form Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times]0, 1[$	64
6.4	Explicit Closed-Form Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus [0, 1])$	67
6.5	Totally Explicit Closed-Form Bounds	69
6.6	Closed-Form Bounds for Power Divergences of Non-Kullback-Leibler-Information-Divergence Type	70
6.7	Applications to Decision Making	71

7 Hellinger Integrals and Power Divergences of Galton-Watson Type Diffusion Approximations	71
7.1 Branching-Type Diffusion Approximations	71
7.2 Bounds of Hellinger Integrals for Diffusion Approximations	74
7.3 Bounds of Power Divergences for Diffusion Approximations	79
7.4 Applications to Decision Making	80
A Proofs and Auxiliary Lemmas	81
A.1 Proofs and Auxiliary Lemmas for Section 3	81
A.2 Proofs and Auxiliary Lemmas for Section 5	88
A.3 Proofs and Auxiliary Lemmas for Section 6	94
A.4 Proofs and Auxiliary Lemmas for Section 7	101
References	115

1. Introduction

(This paper is a thoroughly revised, extended and retitled version of the preprint arXiv:1005.3758v1 of both authors) Over the past twenty years, *density-based divergences* $D(P, Q)$ –also known as (dis)similarity measures, directed distances, disparities, distinguishability measures, proximity measures–between probability distributions P and Q , have turned out to be of substantial importance for decisive statistical tasks such as parameter estimation, testing for goodness-of-fit, Bayesian decision procedures, change-point detection, clustering, as well as for other research fields such as information theory, artificial intelligence, machine learning, signal processing (including image and speech processing), pattern recognition, econometrics, and statistical physics. For some comprehensive overviews on the divergence approach to statistics and probability, the reader is referred to the insightful books of e.g., Liese & Vajda [1], Read & Cressie [2], Vajda [3], Csiszár & Shields [4], Stummer [5], Pardo [6], Liese & Miescke [7], Basu et al. [8], Voinov et al. [9], the survey articles of e.g., Liese & Vajda [10], Vajda & van der Meulen [11], the structure-building papers of Stummer & Vajda [12], Kießlinger & Stummer [13] and Broniatowski & Stummer [14], and the references therein. Divergence-based bounds of minimal mean decision risks (e.g., Bayes risks in finance) can be found e.g., in Stummer & Vajda [15] and Stummer & Lao [16].

Amongst the above-mentioned dissimilarity measures, an important omnipresent subclass are the so-called f –divergences of Csiszar [17], Ali & Silvey [18] and Morimoto [19]; important special cases thereof are the total variation distance and the very frequently used λ –order power divergences $I_\lambda(P, Q)$ (also known as alpha-entropies, Cressie-Read measures, Tsallis cross-entropies) with $\lambda \in \mathbb{R}$. The latter cover e.g., the very prominent Kullback-Leibler information divergence $I_1(P, Q)$ (also called relative entropy), the (squared) Hellinger distance $I_{1/2}(P, Q)$, as well as the Pearson chi-square divergence $I_2(P, Q)$. It is well known that the power divergences can be build with the help of the λ –order Hellinger integrals $H_\lambda(P, Q)$ (where e.g., the case $\lambda = 1/2$ corresponds to the well-known Bhattacharyya coefficient), which are information measures of interest by their own and which are also the crucial ingredients of λ –order Renyi divergences $R_\lambda(P, Q)$ (see e.g., Liese & Vajda [1], van Erven & Harremoës [20]); the case $R_{1/2}(P, Q)$ corresponds to the well-known Bhattacharyya distance.

The above-mentioned information/dissimilarity measures have been also investigated in non-static, time-dynamic frameworks such as for various different contexts of *stochastic processes* like *processes with independent increments* (see e.g., Newman [21], Liese [22], Memin & Shiryaev [23], Jacod & Shiryaev [24], Liese & Vajda [1], Linkov & Shevlyakov [25]), *Poisson point processes* (see e.g., Liese [26], Jacod & Shiryaev [24], Liese & Vajda [1]), *diffusion processes and solutions of stochastic differential equations with continuous paths* (see e.g., Kabanov et al. [27], Liese [28], Jacod & Shiryaev [24], Liese & Vajda [1], Vajda [29], Stummer [30–32], Stummer & Vajda [15]), and *generalized binomial processes* (see e.g., Stummer & Lao [16]); further related literature can be found e.g., in references of the aforementioned papers and books.

Another important class of time-dynamic models is given by *discrete-time integer-valued branching processes*, in particular (*Bienaymé*-)Galton-Watson processes without immigration GW respectively with immigration (resp. importation, invasion) GWI, which have numerous applications in biotechnology, population genetics, internet traffic research, clinical trials, asset price modelling, derivative pricing, and many others. As far as important terminology is concerned, we abbreviatingly subsume both models as GW(I) and, simply as GWI in case that GW appears as a parameter-special-case of GWI; recall that a GW(I) is called *subcritical* respectively *critical* respectively *supercritical* if its offspring mean is less than 1 respectively equal to 1 respectively larger than 1.

For applications of GW(I) in *epidemiology*, see e.g., the works of Bartoszynski [33], Ludwig [34], Becker [35,36], Metz [37], Heyde [38], von Bahr & Martin-Löf [39], Ball [40], Jacob [41], Barbour & Reinert [42], Section 1.2 of Britton & Pardoux [43]); for more details see Section 2.3 below.

For connections of GW(I) to *time series of counts* including GLM models, see e.g., Dion, Gauthier & Latour [44], Grunwald et al. [45], Kedem & Fokianos [46], Held, Höhle & Hofmann [47], and Weiß [48]; a more comprehensive discussion can be found in Section 2.2 below.

As far as the combined study of information measures and GW processes is concerned, let us first mention that (transforms of) power divergences have been used for supercritical Galton-Watson processes without immigration for instance as follows: Feigin & Passy [49] study the problem to find an offspring distribution which is closest (in terms of relative entropy type distance) to the original offspring distribution and under which ultimate extinction is certain. Furthermore, Mordecki [50] gives an equivalent characterization for the stable convergence of the corresponding log-likelihood process to a mixed Gaussian limit, in terms of conditions on Hellinger integrals of the involved offspring laws. Moreover, Sriram & Vidyashankar [51] study the properties of offspring-distribution-parameters which minimize the squared Hellinger distance between the model offspring distribution and the corresponding non-parametric maximum likelihood estimator of Guttorp [52]. For the setup of GWI with Poisson offspring and nonstochastic immigration of constant value 1, Linkov & Lunyova [53] investigate the asymptotics of Hellinger integrals in order to deduce large deviation assertions in hypotheses testing problems.

In contrast to the above-mentioned contexts, this paper pursues the following main goals:

- (MG1) for any time horizon and any criticality scenario (allowing for non-stationarities), to compute lower and upper bounds—and sometimes even exact values—of the Hellinger integrals $H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})$, power divergences $I_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})$ and Renyi divergences $R_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})$ of two alternative Galton-Watson branching processes $P_{\mathcal{A}}$ and $P_{\mathcal{H}}$ (on path/scenario space), where (i) $P_{\mathcal{A}}$ has $\text{Poisson}(\beta_{\mathcal{A}})$ distributed offspring as well as $\text{Poisson}(\alpha_{\mathcal{A}})$ distributed immigration, and (ii) $P_{\mathcal{H}}$ has $\text{Poisson}(\beta_{\mathcal{H}})$ distributed offspring as well as $\text{Poisson}(\alpha_{\mathcal{H}})$ distributed immigration; the non-immigration cases are covered as $\alpha_{\mathcal{A}} = \alpha_{\mathcal{H}} = 0$; as a side effect, we also aim for corresponding asymptotic distinguishability results;
- (MG2) to compute the corresponding limit quantities for the context in which (a proper rescaling of) the two alternative Galton-Watson processes with immigration converge to *Feller*-type branching diffusion processes, as the time-lags between the generation-size observations tend to zero;
- (MG3) as an exemplary field of application, to indicate how to use the results of (MG1) for Bayesian decision making in the epidemiological context of an infectious-disease pandemic (e.g., the current COVID-19), where e.g., potential state-budgetary losses can be controlled by alternative public policies (such as e.g., different degrees of lockdown) for mitigations of the time-evolution of the number of infectious persons (being quantified by a GW(I)). Corresponding Neyman-Pearson testing will be treated, too.

Because of the involved Poisson distributions, these goals can be tackled with a high degree of tractability, which is worked out in detail with the following structure (see also the full table of contents after this paragraph): in Section 2, we first introduce (i) the basic ingredients of Galton-Watson

processes together with their interpretations in the above-mentioned pandemic setup where it is essential to study *all* types of criticality (being connected with levels of reproduction numbers), (ii) the employed fundamental information measures such as Hellinger integrals, power divergences and Renyi divergences, (iii) the underlying decision-making framework, as well as (iv) connections to time series of counts and asymptotical distinguishability. Thereafter, we start our detailed technical analyses by giving *recursive* exact values respectively *recursive* bounds—as well as their applications—of Hellinger integrals $H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})$ (see Section 3), power divergences $I_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})$ and Renyi divergences $R_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})$ (see Sections 4 and 5). *Explicit closed-form* bounds of Hellinger integrals $H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})$ will be worked out in Section 6, whereas Section 7 deals with Hellinger integrals and power divergences of the above-mentioned Galton-Watson type diffusion approximations.

2. The Framework and Application Setups

2.1. Process Setup

We investigate dissimilarity measures and apply them to decisions, in the following context. Let the integer-valued random variable X_n ($n \in \mathbb{N}_0$) denote the size of the n th generation of a population (of persons, organisms, spreading news, other kind of objects, etc.) with specified characteristics, and suppose that for the modelling of the time-evolution $n \mapsto X_n$ we have the choice between the following two (e.g., alternative, competing) models (\mathcal{H}) and (\mathcal{A}):

(\mathcal{H}) a discrete-time homogeneous *Galton-Watson process with immigration* *GWI*, given by the recursive description

$$X_0 \in \mathbb{N}; \quad \mathbb{N}_0 \ni X_n = \sum_{k=1}^{X_{n-1}} Y_{n-1,k} + \tilde{Y}_n, \quad n \in \mathbb{N}, \quad (1)$$

where $Y_{n-1,k}$ is the number of offspring of the k th object (e.g., organism, person) within the $(n-1)$ th generation, and \tilde{Y}_n denotes the number of immigrating objects in the n th generation. Notice that we employ an arbitrary *deterministic* (i.e., degenerate random) initial generation size X_0 . We always assume that under the corresponding dynamics-governing law $P_{\mathcal{H}}$

- (GWI1) the collection $Y := \{Y_{n-1,k}, n \in \mathbb{N}, k \in \mathbb{N}\}$ consists of independent and identically distributed (i.i.d.) random variables which are Poisson distributed with parameter $\beta_{\mathcal{H}} > 0$,
- (GWI2) the collection $\tilde{Y} := \{\tilde{Y}_n, n \in \mathbb{N}\}$ consists of i.i.d. random variables which are Poisson distributed with parameter $\alpha_{\mathcal{H}} \geq 0$ (where $\alpha_{\mathcal{H}} = 0$ stands for the degenerate case of having no immigration),
- (GWI3) Y and \tilde{Y} are independent.

(\mathcal{A}) a discrete-time homogeneous *Galton-Watson process with immigration* *GWI* given by the same recursive description (1), but with different dynamics-governing law $P_{\mathcal{A}}$ under which (GWI1) holds with parameter $\beta_{\mathcal{A}} > 0$ (instead of $\beta_{\mathcal{H}} > 0$), (GWI2) holds with $\alpha_{\mathcal{A}} \geq 0$ (instead of $\alpha_{\mathcal{H}} \geq 0$), and (GWI3) holds. As a side remark, in some contexts the two models (\mathcal{H}) and (\mathcal{A}) may function as a “sandwich” of a more complicated not fully known model.

Basic and advanced facts on general *GWI* (introduced by Heathcote [54]) can be found e.g., in the monographs of Athreya & Ney [55], Jagers [56], Asmussen & Hering [57], Haccou [58]; see also e.g., Heyde & Seneta [59], Basawa & Rao [60], Basawa & Scott [61], Sankaranarayanan [62], Wei & Winnicki [63], Winnicki [64], Guttorp [52] as well as Yanev [65] (and also the references therein all those) for adjacent fundamental statistical issues including the involved technical and conceptual challenges.

For the sake of brevity, wherever we introduce or discuss corresponding quantities *simultaneously* for both models \mathcal{H} and \mathcal{A} , we will use the subscript \bullet as a synonym for either the symbol \mathcal{H} or \mathcal{A} .

For illustration, recall the well-known fact that the corresponding conditional probabilities $P_{\bullet}(X_n = \cdot | X_{n-1} = k)$ are again Poisson-distributed, with parameter $\beta_{\bullet} \cdot k + \alpha_{\bullet}$.

In order to achieve a transparently representable structure of our results, we subsume the involved parameters as follows:

- (PS1) \mathcal{P}_{SP} is the set of all constellations $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}})$ of real-valued parameters $\beta_{\mathcal{A}} > 0, \beta_{\mathcal{H}} > 0, \alpha_{\mathcal{A}} > 0, \alpha_{\mathcal{H}} > 0$, such that $\beta_{\mathcal{A}} \neq \beta_{\mathcal{H}}$ or $\alpha_{\mathcal{A}} \neq \alpha_{\mathcal{H}}$ (or both); in other words, both models are non-identical and have non-vanishing immigration;
- (PS2) \mathcal{P}_{NI} is the set of all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}})$ of real-valued parameters $\beta_{\mathcal{A}} > 0, \beta_{\mathcal{H}} > 0, \alpha_{\mathcal{A}} = \alpha_{\mathcal{H}} = 0$, such that $\beta_{\mathcal{A}} \neq \beta_{\mathcal{H}}$; this corresponds to the important special case that both models have no immigration and are non-identical;
- (PS3) the resulting disjoint union will be denoted by $\mathcal{P} = \mathcal{P}_{SP} \cup \mathcal{P}_{NI}$.

Notice that for (unbridgeable) technical reasons, we *do not allow* for “crossovers” between “immigration and no-immigration” (i.e., $\alpha_{\mathcal{A}} = 0$ and $\alpha_{\mathcal{H}} \neq 0$, respectively, $\alpha_{\mathcal{A}} \neq 0$ and $\alpha_{\mathcal{H}} = 0$). For practice, this is not a strong restriction, since one may take e.g., $\alpha_{\mathcal{A}} = 10^{-12}$ and $\alpha_{\mathcal{H}} = 1$.

For the non-immigration case $\alpha_{\bullet} = 0$ one has the following *extinction properties* (see e.g., Harris [66], Athreya & Ney [55]). As usual, let us define the extinction time $\tau := \min \{i \in \mathbb{N} : X_{\ell} = 0 \text{ for all integers } \ell \geq i\}$ if this minimum exists, and $\tau := \infty$ else. Correspondingly, let $\mathcal{B} := \{\tau < \infty\}$ be the extinction set. If the *offspring mean* β_{\bullet} satisfies $\beta_{\bullet} < 1$ —which is called the *subcritical* case— or $\beta_{\bullet} = 1$ —which is known as the *critical* case—then extinction is certain, i.e., there holds $P(\mathcal{B} | X_0 = 1) = 1$. However, if the offspring mean satisfies $\beta_{\bullet} > 1$ —which is called the *supercritical* case—then there is a probability greater than zero, that the population never dies out, i.e., $P(\mathcal{B} | X_0 = 1) \in]0, 1[$. In the latter case, X_n explodes (a.s.) to infinity as $n \rightarrow \infty$.

In contrast, for the (nondegenerate, nonvanishing) immigration case $\alpha_{\bullet} \neq 0$ there is *no extinction*, viz. $P(\mathcal{B} | X_0 = 1) = 0$, although there may be zero population $X_{\ell_0} = 0$ for some intermediate time $\ell_0 \in \mathbb{N}$; but due to the immigration, with probability one there is always a later time $\ell_1 > \ell_0$, such that $X_{\ell_1} > 0$. Nevertheless, also for the setup $\alpha_{\bullet} \neq 0$ it is important to know whether $\beta_{\bullet} \stackrel{\cong}{\leq} 1$ —which is still called (super-, sub-)criticality—since e.g., in the case $\beta_{\bullet} < 1$ the population size X_n converges (as $n \rightarrow \infty$) to a stationary distribution on \mathbb{N} whereas for $\beta_{\bullet} > 1$ the behaviour is non-stationary (non-ergodic), see e.g., Athreya & Ney [55].

At this point, let us emphasize that in our investigations (both for $\alpha_{\bullet} = 0$ and for $\alpha_{\bullet} \neq 0$) we *do allow* for “crossovers” between “different criticalities”, i.e., we deal with all cases $\beta_{\mathcal{A}} \stackrel{\cong}{\leq} 1$ versus all cases $\beta_{\mathcal{H}} \stackrel{\cong}{\leq} 1$; as will be explained in the following, this unifying flexibility is especially important for corresponding epidemiological-model comparisons (e.g., for the sake of decision making).

One of our main goals is to quantitatively compare (the time-evolution of) two competing GWI models \mathcal{H} and \mathcal{A} with respective parameter sets $(\beta_{\mathcal{H}}, \alpha_{\mathcal{H}})$ and $(\beta_{\mathcal{A}}, \alpha_{\mathcal{A}})$, in terms of the information measures $H_{\lambda}(P_{\mathcal{A}} || P_{\mathcal{H}})$ (Hellinger integrals), $I_{\lambda}(P_{\mathcal{A}} || P_{\mathcal{H}})$ (power divergences), $R_{\lambda}(P_{\mathcal{A}} || P_{\mathcal{H}})$ (Renyi divergences). The latter two express a distance (degree of dissimilarity) between \mathcal{H} and \mathcal{A} . From this, we shall particularly derive applications for decision making under uncertainty (including tests).

2.2. Connections to Time Series of Counts

It is well known that a Galton-Watson process with Poisson offspring (with parameter β_{\bullet}) and Poisson immigration (with parameter α_{\bullet}) is “distributionally” equal to each of the following models (listed in “tree-type” chronological order):

- (M1) a Poissonian *Generalized Integer-valued Autoregressive process* GINAR(1) in the sense of Gauthier & Latour [67] (see also Dion, Gauthier & Latour [44], Latour [68], as well as Grunwald et al. [45]), that is, a first-order autoregressive times series with Poissonian thinning (with parameter β_{\bullet}) and Poissonian innovations (with parameter α_{\bullet});

- (M2) Poissonian *first order Conditional Linear Autoregressive model* (Poissonian CLAR(1)) in the sense of Grunwald et al. [45] (and earlier preprints thereof) (since the conditional expectation is $EP_{\bullet}[X_n|\mathcal{F}_{n-1}] = \alpha_{\bullet} + \beta_{\bullet} \cdot X_{n-1}$); this can be equally seen as Poissonian autoregressive *Generalized Linear Model* GLM with identity link function (cf. [45] as well as Chapter 4 of Kedem & Fokianos [46]), that is, an autoregressive GLM with Poisson distribution as random component and the identity link as systematic component; the same model was used (and generalized)
- (M2i) under the name BIN(1) by Rydberg & Shephard [69] for the description of the number X_n of stock transactions/trades recorded up to time n ;
 - (M2ii) under the name *Poisson autoregressive model* PAR(1) by Brandt & Williams [70] for the description of event counts in political and other social science applications;
 - (M2iii) under the name *Autoregressive Conditional Poisson model* ACP(1,0) by Heinen [71];
 - (M2iv) by Held, Höhle & Hofmann [47] as well as Held et al. [72], as a description of the time-evolution of counts from infectious disease surveillance databases, where β_{\bullet} (respectively, α_{\bullet}) is interpreted as driving parameter of epidemic (respectively, endemic) component; in principle, this type of modelling can be also implicitly recovered as a special case of the epidemics-treating work of Finkenstädt, Bjornstad & Grenfell [73], by assuming trend- and season-neglecting (e.g., intra-year) measles data in urban areas of about 10 million people (provided that their population size approximation extends linearly);
 - (M2v) under the name *integer-valued Generalized Autoregressive Conditional Heteroscedastic model* INGARCH(1,0) by Ferland, Latour & Oraichi [74] (since the conditional variance is $VarP_{\bullet}[X_n|\mathcal{F}_{n-1}] = \alpha_{\bullet} + \beta_{\bullet} \cdot X_{n-1}$), see also Weiß [75]; this has been refinedly named as INARCH(1) model by Weiß [76,77], and frequently applied thereafter; for an “overlapping-generation type” interpretation of the INARCH(1) model, which is an adequate description for the time-evolution of overdispersed counts with an autoregressive serial dependence structure, see Weiß & Testik [78]; for a corresponding comprehensive recent survey (also to more general count time series), the reader is referred to the book of Weiß [48];

Moreover, according to the general considerations of Grunwald et al. [45], the Poissonian Galton-Watson model with immigration may possibly be “distributionally equal” to an integer-valued autoregressive model with random coefficient (thinning).

Nowadays, besides the name *homogeneous Galton-Watson model with immigration* GWI, the name *INARCH(1)* seems to be the most used one, and we follow this terminology (with emphasis on GWI). Typical features of the above-mentioned models (M1) to (M2v), are the use of \mathbb{Z} as the set of times, and the assumptions $\alpha_{\bullet} > 0$ as well as $\beta_{\bullet} \in]0, 1[$, which guarantee stationarity and ergodicity (see above). In contrast, we employ \mathbb{N}_0 as the set of times, degenerate (and thus, non-equilibrium) starting distribution, and arbitrary $\alpha_{\bullet} \geq 0$ as well as $\beta_{\bullet} > 0$. For such a situation, as explained above, we quantitatively compare two competing GWI models \mathcal{H} and \mathcal{A} with respective parameter sets $(\beta_{\mathcal{H}}, \alpha_{\mathcal{H}})$ and $(\beta_{\mathcal{A}}, \alpha_{\mathcal{A}})$. Since—as can be seen e.g., in (29) below—we basically employ only (conditionally) distributional ingredients, such as the corresponding likelihood ratio (see e.g., (13) to (15), (27) to (29) below), *all the results of the Sections 3–6 can be immediately carried over to the above-mentioned time-series contexts* (where we even allow for non-stationarities, in fact we start with a one-point/Dirac distribution); for the sake of brevity, in the rest of the paper this will not be mentioned explicitly anymore.

Notice that a Poissonian GWI as well as all models (M1) and (M2) are—despite of their *conditional* Poisson law—typically overdispersed since

$$EP_{\bullet}[X_n] = \alpha_{\bullet} + \beta_{\bullet} \cdot EP_{\bullet}[X_{n-1}] \leq \alpha_{\bullet} + \beta_{\bullet} \cdot EP_{\bullet}[X_{n-1}] + \beta_{\bullet}^2 \cdot VarP_{\bullet}[X_{n-1}] = VarP_{\bullet}[X_n], \quad n \in \mathbb{N} \setminus \{1\},$$

with equality iff (i.e., if and only if) $\alpha_{\bullet} = 0$ (NI) and $X_{n-2} = 0$ (extinction at $n - 2$ with $n \geq 3$).

2.3. Applicability to Epidemiology

The above-mentioned framework can be used for any of the numerous fields of applications of discrete-time branching processes, and of the closely related INARCH(1) models. For the sake of brevity, we explain this—as a kind of running-example—in detail for the currently highly important context of the epidemiology of infectious diseases. For insightful non-mathematical introductions to the latter, see e.g., Kaslow & Evans [79], Osterholm & Hedberg [80]; for a first entry as well as overviews on modelling, the reader is referred to e.g., Grassly & Fraser [81], Keeling & Rohani [82], Yan [83,84], Britton [85], Diekmann, Heesterbeek & Britton [86], Cummings & Lessler [87], Just et al. [88], Britton & Giardina [89], Britton & Pardoux [43]. A survey on the particular role of branching processes in epidemiology can be found e.g., in Jacob [41].

Undoubtedly, by nature, the spreading of an infectious disease through a (human, animal, plant) population is a branching process with possible immigration. Indeed, typically one has the following mechanism:

- (D1) at some time t_k^E —called the time of exposure (moment of infection)—an individual k of a specified population is infected in a wide sense, i.e., entered/invaded/colonized by a number of transmissible disease-causative pathogens (etiologic agents such as viruses, bacteria, protozoans and other parasites, subviruses (e.g., prions and plant viroids), etc.); the individual is then a *host* (of pathogens);
- (D2) depending on the level of immunity and some other factors, these pathogens may multiply/replicate within the host to an extent (over a threshold number) such that at time t_k^I some of the pathogens start to leave their host (*shedding of pathogens*); in other words, the individual k becomes *infectious* at the time t_k^I of *onset of infectiousness*. Ex post, one can then say that the individual became infected in the narrow sense at earlier time t_k^E and call it a *primary case*. The time interval $[t_k^E, t_k^I[$ is called the *latent/latency/pre-infectious period* of k , and $t_k^I - t_k^E$ its duration (in some literature, there is no verbal distinction between them); notice that t_k^I may differ from the time t_k^{OS} of *onset (first appearance) of symptoms*, which leads to the so-called *incubation period* $[t_k^E, t_k^{OS}[$; if $t_k^I < t_k^{OS}$ then $[t_k^I, t_k^{OS}[$ is called the *pre-symptomatic period*;
- (D3) as long as the individual k stays infectious, by shedding of pathogens it may infect in a narrow sense a random number $Y_k \in \mathbb{N}_0$ of other individuals which are *susceptible* (i.e., neither immune nor already infected in a narrow sense), where the distribution of Y_k depends on the individual's (natural, voluntary, forced) behaviour, its environment, as well as some other factors e.g., connected with the type of pathogen transmission; the newly infected individuals are called *offspring of k* , and *secondary cases* if they are from the same specified population or *exportations* if they are from a different population; from the view of the latter, these infections are *imported cases* and thus can be viewed as *immigrants*;
- (D4) at the time t_k^R of *cessation of infectiousness*, the individual stops being infectious (e.g., because of recovery, death, or total isolation); the time interval $[t_k^I, t_k^R[$ is called the *period of infectiousness (also period of communicability, infectious/infective/shedding/contagious period)* of k , and $t_k^R - t_k^I$ its duration (in some literature, there is no verbal distinction between them); notice that t_k^R may differ from the time t_k^{CS} of *cessation (last appearance) of symptoms* which leads to the so-called *sickness period* $[t_k^{OS}, t_k^{CS}[$;
- (D5) this branching mechanism continues within the specified population until there are no infectious individuals and also no importations anymore (eradication, full extinction, total elimination)—up to a specified final time (which may be large or even infinite);

All the above-mentioned times t_k and time intervals are random, by nature. Two further connected quantities are also important for modelling (see e.g., Yan & Chowell [84] (p. 241ff), including a history of corresponding terminology). Firstly, the *generation interval* (generation time, transmission interval)

is the time interval from the onset of infectiousness in a primary case (called the infector) to the onset of infectiousness in a secondary case (called the infectee) infected by the primary case; clearly, the generation interval is random, and so is its duration (often, the (population-)mean of the latter is also called generation interval). Typically, generation intervals are important ingredients of branching process models of infectious diseases. Secondly, the *serial interval* describes time interval from the onset of symptoms in a primary case to the onset of symptoms in a secondary case infected by the primary case. By nature, the serial interval is random, and so is its duration (often, the (population-)mean of the latter is also called serial interval). Typically, the serial interval is easier to observe than the generation interval, and thus, the latter is often approximately estimated from data of the former. For further investigations on generation and serial intervals, the reader is referred to e.g., Fine [90], Svensson [91,92], Wallinga & Lipsitch [93], Forsberg White & Pagano [94], Nishiura [95], Scalia Tomba et al. [96], Trichereau et al. [97], Vink, Bootsma & Wallinga [98], Champredon & Dushoff [99], Just et al. [88], and—especially for the novel COVID-19 pandemics—An der Heiden & Hamouda [100], Ferretti et al. [101], Ganyani et al. [102], Li et al. [103], Nishiura, Linton & Akhmetzhanov [104], Park et al. [105].

With the help of the above-mentioned *individual* ingredients, one can aggregatedly build numerous different *population-wide* models of infectious diseases in discrete time as well as in continuous time; the latter are typically observed only in discrete-time steps (discrete-time sampling), and hence in the following we concentrate on discrete-time modelling (of the real or the observational process). In fact, we confine ourselves to the important task of modelling the evolution $n \mapsto X_n$ of the number of *incidences* at “stage” n , where *incidence* refers to the number of *new* infected/infectious individuals. Here, n may be a generation number where, inductively, $n = 0$ refers to the generation of the first appearing primary cases in the population (also called *initial importations*), and n refers to the generation of offsprings of all individuals of generation $n - 1$. Alternatively, n may be the index of a physical (“calender”) point of time t_n , which may be deterministic or random; e.g., $(t_n)_{n \in \mathbb{N}}$ may be a strictly increasing series of (i) equidistant deterministic time points (and thus, one can identify $t_n = n$ in appropriate time units such as days, weeks, bi-weeks, months), or (ii) non-equidistant deterministic time points, or (iii) random time points (as a side remark, let us mention that in some situations, X_n may alternatively denote the number of *prevalences* at “stage” n , where *prevalence* refers to the total number of infected/infectious individuals (e.g., through some methodical tricks like “self-infection”).

In the light of this, one can loosely define an *epidemic* as the rapid spread of an infectious disease within a specified population, where the numbers X_n of incidences are high (or much higher than expected) for that kind of population. A *pandemic* is a geographically large-scale (e.g., multicontinental or worldwide) epidemic. An *outbreak/onset* of an epidemic in the narrow sense is the (time of) change where an infectious disease turns into an epidemic, which is typically quantified by exceedance over an threshold; analogously, an *outbreak/onset* of a pandemic is the (time of) change where the epidemic turns into a pandemic. Of course, one goal of infectious-disease modelling is to quantify “early enough” the potential danger of an emerging outbreak of an epidemic or a pandemic.

Returning to possible models of the incidence-evolution $n \mapsto X_n$, its description may be theoretically derived from more detailed, time-finer, highly sophisticated, individual-based “mechanistic” infectious-disease models such as e.g., continuous-time susceptible-exposed-infectious-recovered (SEIR) models (see the above-mentioned introductory texts); however, as e.g., pointed out in Held et al. [72], the estimation of the correspondingly involved numerous parameters may be too ambitious for routinely collected, non-detailed disease data, such as e.g., daily/weekly counts X_n of incidences—especially in decisive emerging/early phases of a novel disease (such as the current COVID-19 pandemic). Accordingly, in the following we assume that X_n can be approximately described by a Poissonian Galton-Watson process with immigration respectively a (“distributionally equal”) Poissonian autoregressive Generalized Linear Model in the sense of (M2). Depending on the situation, this can be quite reasonable, for the following arguments (apart from the usual “if the data say so”). Firstly, it is well known (see e.g., Bartoszynski [33], Ludwig [34],

Becker [35,36], Metz [37], Heyde [38], von Bahr & Martin-Löf [39], Ball [40], Jacob [41], Barbour & Reinert [42], Section 1.2 of Britton & Pardoux [43]) that in populations with a relatively high number of susceptible individuals and a relatively low number of infectious individuals (e.g., in a large population and in decisive emerging/early phases of the disease spreading), the incidence-evolution $n \mapsto X_n$ can be well approximated by a (e.g., Poissonian) Galton-Watson process with possible immigration where n plays the role of a *generation number*. If the above-mentioned generation interval is “nearly” deterministic (leading to nearly synchronous, non-overlapping generations)—which is the case e.g., for (phases of) Influenza A(H1N1)pdm09, Influenza A(H3N2), Rubella (cf. Vink, Bootsma & Wallinga [98]), and COVID-19 (cf. Ferretti et al. [101])—and the length of the generation interval is approximated by its mean length and the latter is tuned to be equal to the unit time between consecutive observations, then n plays the role of an *observation (surveillance) time*. This effect is even more realistic if the period of infectiousness is nearly deterministic and relatively short. Secondly, as already mentioned above, the spreading of an infectious disease is intrinsically a (not necessarily Poissonian Galton-Watson) branching mechanism, which may be blurred by other effects in a way that a Poissonian autoregressive Generalized Linear Model is still a reasonably fitting model for the observational process in disease surveillance. The latter have been used e.g., by Finkenstädt, Bjornstad & Grenfell [73], Held, Höhle & Hofmann [47], and Held et al. [72]; they all use non-constant parameters (e.g., to describe seasonal effects, which are however unknown in early phases of a novel infectious disease such as COVID-19). In contrast, we employ different new—namely divergence-based—statistical techniques, for which we assume constant parameters but also indicate procedures for the detection of changes; the extension to non-constant parameters is straightforward.

Returning to Galton-Watson processes, let us mention as a *side remark* that they can be also used to model the above-mentioned within-host replication dynamics (D2) (e.g., in the time-interval $[t_k^E, t_k^I[$ and beyond) on a sub-cellular level, see e.g., Spouge [106], as well as Taneyhill, Dunn & Hatcher [107] for parasitic pathogens; on the other hand, one can also employ Galton-Watson processes for quantifying snowball-effect (avalanche-effect, cascade-effect) type, economic-crisis triggered consequences of large epidemics and pandemics, such as e.g., the potential spread of transmissible (i) foreclosures of homes (cf. Parnes [108]), or clearly also (ii) company insolvencies, downsizings and credit-risk downgradings; moreover, the time-evolution of integer-valued indicators concerning the spread of (rational or unwarranted) fears resp. perceived threats may be modelled, too.

Summing up things, we model the evolution $n \mapsto X_n$ of the number of incidences at stage n by a Poissonian Galton Watson process with immigration GWI

$$X_0 \in \mathbb{N}; \quad N_0 \ni X_n = \sum_{k=1}^{X_{n-1}} Y_{n-1,k} + \tilde{Y}_n, \quad n \in \mathbb{N}, \quad \text{cf. (1), (GWI1)–(GWI3) with law } P_{\bullet},$$

(where $Y_{n-1,k}$ corresponds to the Y_k of (D3), equipped with an additional stage-index $n - 1$), respectively by a corresponding “distributionally equal”—possibly non-stationary—Poissonian autoregressive Generalized Linear Model in the sense of (M2); depending on the situation, we may also fix a (deterministic or random) upper time horizon other than infinity. Recall that both models are overdispersed, which is consistent with the current debate on overdispersion in connection with the current COVID-19 pandemic. In infectious-disease language, the sum $\sum_{k=1}^{X_{n-1}} Y_{n-1,k}$ can also be loosely interpreted as *epidemic component* (in a narrow sense) driven by the parameter β_{\bullet} , and \tilde{Y}_n as *endemic component* driven by the parameter α_{\bullet} . In fact, the offspring mean (here, β_{\bullet}) is called *reproduction number* and plays a major role—also e.g., in the current public debate about the COVID-19 pandemic—because it crucially determines the rapidity of the spread of the disease and—as already indicated above in the second and third paragraph after (PS3)—also the probability that the epidemic/pandemic becomes (maybe temporally) extinct or at least stationary at a low level (that is, *endemic*). For this to happen, β_{\bullet} should be subcritical, i.e., $\beta_{\bullet} < 1$, and even better, close to zero. Of course, the size of the *importation mean* $\alpha_{\bullet} \geq 0$ matters, too, in a secondary order.

Keeping this in mind, let us discuss on which factors the reproduction number β_{\bullet} and the importation mean α_{\bullet} depend upon, and how they can be influenced/controlled. To begin with, by recalling the above-mentioned points (D1) to (D5) and by adapting the considerations of e.g., Grassly & Fraser [81] to our model, one encounters the fact that the distribution of the offspring $Y_{n-1,k}$ —here driven by the reproduction number (offspring mean) β_{\bullet} —depends on the following factors:

- (B1) the *degree of infectiousness* of the individual k , with three major components:
- (B1a) degree of *biological* infectiousness; this reflects the within-host dynamics (D2) of the “representative” individual k , in particular the duration and amount of the corresponding replication and shedding/excretion of the infectious pathogens; this degree depends thus on (i) the number of host-invading pathogens (called the *initial infectious dose*), (ii) the type of the pathogen with respect to e.g., its principal capabilities of replication speed, range of spread and drug-sensitivity, (iii) features of the immune system of the host k including the level of innate or acquired immunity, and (iv) the interaction between the genetic determinants of disease progression in both the pathogen and the host;
 - (B1b) degree of *behavioural* infectiousness; this depends on the contact patterns of an infected/infectious individual (and, if relevant, the contact patterns of intermediate hosts or vectors), in relation to the disease-specific type of route(s) of transmission of the infectious pathogens (for an overview of the latter, see e.g., Table 3 of Kaslow & Evans [79]); a long-distance-travel behaviour may also lead to the disease exportation to another, outside population (and thus, for the latter to a disease importation);
 - (B1c) degree of *environmental* infectiousness; this depends on the location and environment of the host k , which influences the duration of outside-host survival of the pathogens (and, if relevant, of the intermediate hosts or vectors) as well as the speed and range of their outside-host spread; for instance, high temperature may kill the pathogens, high airflow or rainfall dynamics may ease their spread, etc.
- (B2) the *degree of susceptibility* of uninfected individuals who have contact with k , with the following three major components (with similar background as their infectiousness counterparts):
- (B2a) degree of *biological* susceptibility;
 - (B2b) degree of *behavioural* susceptibility;
 - (B2c) degree of *environmental* susceptibility.

All these factors (B1a) to (B2c) can be principally influenced/controlled to a certain—respective—extent. Let us briefly discuss this for *human* infectious diseases, where one major goal of epidemic risk management is to operate countermeasures/interventions in order to slow down the disease transmission (e.g., by reducing the reproduction number β_{\bullet} to less than 1) and eventually even break the chain of transmission, for the sake of containment or mitigation; preparedness and preparation are motives, too, for instance as a part of governmental pandemic risk management.

For instance, (B1a) can be reduced or even erased through pharmaceutical interventions such as medication (if available), and preventive strengthening of the immune system through non-extreme sports activities and healthy food.

Moreover, the following exemplary control measures for (B2) can be either put into action by common-sense self-behaviour, or by large-scale public recommendations (e.g., through mass media), or by rules/requirements from authorities:

- (i) personal preventive measures such as frequent washing and disinfecting of hands; keeping hands away from face; covering coughs; avoidance of handshakes and hugs with non-family-members; maintaining physical distance (e.g., of two meters) from non-family-members; wearing a

- face-mask of respective security degree (such as homemade cloth face mask, particulate-filtering face-piece respirator, medical (non-surgical) mask, surgical mask); self-quarantine;
- (ii) environmental measures, such as e.g., cleaning of surfaces;
 - (iii) community measures aimed at mild or stringent social distancing, such as e.g., prohibiting/cancelling/banning gatherings of more than z non-family members (e.g., $z = 2, 5, 10, 100, 1000$ in various different phases and countries during the current COVID-19 pandemic); mask-wearing (see above); closing of schools, universities, some or even all nonessential (“system-irrelevant”) businesses and venues; home-officing/work ban; home isolation of disease cases; isolation of homes for the elderly/aged (nursing homes); stay-at-home orders with exemptions, household or even general quarantine; testing & tracing; lockdown of entire cities and beyond; restricting the degrees of travel freedom/allowed mobility (e.g., local, union-state, national, international including border and airport closure). The latter also affects the mean importation rate α_{\bullet} , which can be controlled by vaccination programs in “outside populations”, too.

As far as the degree of *biological* susceptibility (B2a) is concerned, one obvious therapeutic countermeasure is a mass vaccination program/campaign (if available).

In case of *highly virulent* infectious diseases causing epidemics and pandemics with substantial *fatality rates*, some of the above-mentioned control strategies and countermeasures may (have to) be “drastic” (e.g., lockdown), and thus imply considerable social and economic costs, with a huge impact and potential danger of triggering severe social, economic and political disruptions.

In order to prepare corresponding suggestions for decisions about appropriate control measures (e.g., public policies), it is therefore important—especially for a novel infectious disease such as the current COVID-19 pandemic—to have a model for the time-evolution of the incidences in (i) a natural (basically uncontrolled) set-up, as well as in (ii) the control set-ups under consideration. As already mentioned above, we assume that all these situations can be distilled into an incidence evolution $n \mapsto X_n$ which follows a Poissonian Galton-Watson process with respectively different parameter pairs $(\beta_{\bullet}, \alpha_{\bullet})$. Correspondingly, we always compare two alternative models (\mathcal{H}) and (\mathcal{A}) with parameter pairs $(\beta_{\mathcal{H}}, \alpha_{\mathcal{H}})$ and $(\beta_{\mathcal{A}}, \alpha_{\mathcal{A}})$ which reflect either a “pure” statistical uncertainty (under the *same* uncontrolled or controlled set-up), or the uncertainty between two *different* potential control set-ups (for the sake of assessing the potential impact/efficiency of some planned interventions, compared with alternative ones); the economic impact can be also taken into account, within a Bayesian decision framework discussed in Section 2.5 below. As will be explained in the next subsections, we achieve such comparisons by means of density-based dissimilarity distances/divergences and related quantities thereof.

From the above-mentioned detailed explanations, it is immediately clear that for the described epidemiological context one should investigate *all* types of criticality and importation means for the therein involved two Poissonian Galton-Watson processes with/without immigration (respectively the equally distributed INARCH(1) models); in particular, this motivates (or even “justifies”) the necessity of the very lengthy detailed studies in the Sections 3–7 below.

2.4. Information Measures

Having two competing models (\mathcal{H}) and (\mathcal{A}) at stake, it makes sense to study questions such as “how far are they apart?” and thus “how dissimilar are they?”. This can be quantified in terms of divergences in the sense of directed (i.e., not necessarily symmetric) distances, where usually the triangular inequality fails. Let us first discuss our employed divergence subclasses in a *general* set-up of two *equivalent* probability measures $P_{\mathcal{H}}, P_{\mathcal{A}}$ on a measurable space (Ω, \mathcal{F}) . In terms of the parameter

$\lambda \in \mathbb{R}$, the *power divergences*—also known as Cressie-Read divergences, relative Tsallis entropies, or generalized cross-entropy family—are defined as (see e.g., Liese & Vajda [1,10])

$$0 \leq I_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}) := \begin{cases} I(P_{\mathcal{A}}||P_{\mathcal{H}}), & \text{if } \lambda = 1, \\ \frac{1}{\lambda(\lambda-1)} (H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}) - 1), & \text{if } \lambda \in \mathbb{R} \setminus \{0, 1\}, \\ I(P_{\mathcal{H}}||P_{\mathcal{A}}), & \text{if } \lambda = 0, \end{cases} \quad (2)$$

where

$$I(P_{\mathcal{A}}||P_{\mathcal{H}}) := \int_{\Omega} p_{\mathcal{A}} \log \frac{p_{\mathcal{A}}}{p_{\mathcal{H}}} d\mu \geq 0 \quad (3)$$

is the *Kullback-Leibler information divergence* (also known as *relative entropy*) and

$$H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}) := \int_{\Omega} p_{\mathcal{A}}^\lambda p_{\mathcal{H}}^{1-\lambda} d\mu \geq 0 \quad (4)$$

is the *Hellinger integral of order* $\lambda \in \mathbb{R} \setminus \{0, 1\}$; for this, we assume as usual without loss of generality that the probability measures $P_{\mathcal{H}}, P_{\mathcal{A}}$ are dominated by some σ -finite measure μ , with densities

$$p_{\mathcal{A}} = \frac{dP_{\mathcal{A}}}{d\mu} \quad \text{and} \quad p_{\mathcal{H}} = \frac{dP_{\mathcal{H}}}{d\mu} \quad (5)$$

defined on Ω (the zeros of $p_{\mathcal{H}}, p_{\mathcal{A}}$ are handled in (3) and (4) with the usual conventions). Clearly, for $\lambda \in \{0, 1\}$ one trivially gets

$$H_0(P_{\mathcal{A}}||P_{\mathcal{H}}) = H_1(P_{\mathcal{A}}||P_{\mathcal{H}}) = 1.$$

The Kullback-Leibler information divergences (relative entropies) in (2) and (3) can alternatively be expressed as (see, e.g., Liese & Vajda [1])

$$I(P_{\mathcal{A}}||P_{\mathcal{H}}) = \lim_{\lambda \nearrow 1} \frac{1 - H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})}{\lambda(1-\lambda)}, \quad I(P_{\mathcal{H}}||P_{\mathcal{A}}) = \lim_{\lambda \searrow 0} \frac{1 - H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})}{\lambda(1-\lambda)}. \quad (6)$$

Apart from the Kullback-Leibler information divergence (relative entropy), other prominent examples of power divergences are the squared Hellinger distance $\frac{1}{2} I_{1/2}(P_{\mathcal{A}}||P_{\mathcal{H}})$ and Pearson's χ^2 -divergence $2 I_2(P_{\mathcal{A}}||P_{\mathcal{H}})$; the Hellinger integral $H_{1/2}(P_{\mathcal{A}}||P_{\mathcal{H}})$ is also known as (multiple of) the *Bhattacharyya coefficient*. Extensive studies about basic and advanced general facts on power divergences, Hellinger integrals and the related Renyi divergences of order $\lambda \in \mathbb{R} \setminus \{0, 1\}$

$$0 \leq R_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}) := \frac{1}{\lambda(\lambda-1)} \log H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}), \quad \text{with } \log 0 = -\infty, \quad (7)$$

can be found e.g., in Liese & Vajda [1,10], Jacod & Shiryaev [24], van Erven & Harremoës [20] (as a side remark, $R_{1/2}(P_{\mathcal{A}}||P_{\mathcal{H}})$ is also known as (multiple of) *Bhattacharyya distance*). For instance, the integrals in (3) and (4) do not depend on the choice of μ . Furthermore, one has the skew symmetries

$$H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}) = H_{1-\lambda}(P_{\mathcal{H}}||P_{\mathcal{A}}), \quad \text{as well as} \quad I_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}) = I_{1-\lambda}(P_{\mathcal{H}}||P_{\mathcal{A}}), \quad (8)$$

for all $\lambda \in \mathbb{R}$ (see e.g., Liese & Vajda [1]). As far as finiteness is concerned, for $\lambda \in]0, 1[$ one gets the rudimentary bounds

$$0 < H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}) \leq 1, \quad \text{and equivalently,} \quad (9)$$

$$0 \leq I_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}}) = \frac{1 - H_\lambda(P_{\mathcal{A}}||P_{\mathcal{H}})}{\lambda(1-\lambda)} < \frac{1}{\lambda(1-\lambda)}, \quad (10)$$

where the lower bound in (10) (upper bound in (9)) is achieved iff $P_A = P_H$. For $\lambda \in \mathbb{R} \setminus]0, 1[$, one gets the bounds

$$0 \leq I_\lambda(P_A || P_H) \leq \infty, \quad \text{and equivalently,} \quad 1 \leq H_\lambda(P_A || P_H) \leq \infty, \quad (11)$$

where, in contrast to above, both the lower bound of $H_\lambda(P_A || P_H)$ and the lower bound of $I_\lambda(P_A || P_H)$ is achieved iff $P_A = P_H$; however, the power divergence $I_\lambda(P_A || P_H)$ and Hellinger integral $H_\lambda(P_A || P_H)$ might be infinite, depending on the particular setup.

The Hellinger integrals can be also used for bounds of the well-known *total variation*

$$0 \leq V(P_A || P_H) := 2 \sup_{A \in \mathcal{F}} \{P_A(A) - P_H(A)\} = \int_{\Omega} |p_A - p_H| d\mu,$$

with p_A and p_H defined in (5). Certainly, the total variation is one of the best known statistical distances, see e.g., Le Cam [109]. For arbitrary $\lambda \in]0, 1[$ there holds (cf. Liese & Vajda [1])

$$1 - \frac{V(P_A || P_H)}{2} \leq H_\lambda(P_A || P_H) \leq \left(1 + \frac{V(P_A || P_H)}{2}\right)^{\max\{\lambda, 1-\lambda\}} \left(1 - \frac{V(P_A || P_H)}{2}\right)^{\min\{\lambda, 1-\lambda\}}.$$

From this together with the particular choice $\lambda = \frac{1}{2}$, we can derive the fundamental universal bounds

$$2(1 - H_{\frac{1}{2}}(P_A || P_H)) \leq V(P_A || P_H) \leq 2\sqrt{1 - (H_{\frac{1}{2}}(P_A || P_H))^2}. \quad (12)$$

We apply these concepts to our setup of Section 2.1 with two competing models (\mathcal{H}) and (\mathcal{A}) of Galton-Watson processes with immigration, where one can take $\Omega \subset \mathbb{N}_0^{\mathbb{N}_0}$ to be the space of all paths of $(X_n)_{n \in \mathbb{N}}$. More detailed, in terms of the extinction set $\mathcal{B} := \{\tau < \infty\}$ and the parameter-set notation (PS1) to (PS3), it is known that for \mathcal{P}_{SP} the two laws P_H and P_A are equivalent, whereas for \mathcal{P}_{NI} the two restrictions $P_H|_{\mathcal{B}}$ and $P_A|_{\mathcal{B}}$ are equivalent (see e.g., Lemma 1.1.3 of Guttorp [52]); with a slight abuse of notation we shall henceforth omit $|_{\mathcal{B}}$. Consistently, for fixed time $n \in \mathbb{N}_0$ we introduce $P_{A,n} := P_A|_{\mathcal{F}_n}$ and $P_{H,n} := P_H|_{\mathcal{F}_n}$ as well as the corresponding Radon-Nikodym-derivative (likelihood ratio)

$$Z_n := \frac{dP_{A,n}}{dP_{H,n}}, \quad (13)$$

where $(\mathcal{F}_n)_{n \in \mathbb{N}}$ denotes the corresponding canonical filtration generated by $X := (X_n)_{n \in \mathbb{N}}$; in other words, \mathcal{F}_n reflects the “process-intrinsic” information known at stage n . Clearly, $Z_0 = 1$. By choosing the reference measure $\mu = P_{H,n}$ one obtains from (4) the Hellinger integral $H_\lambda(P_{A,0} || P_{H,0}) = 1$, as well as and for all $n \in \mathbb{N}$

$$H_\lambda(P_{A,n} || P_{H,n}) = EP_{H,n}[(Z_n)^\lambda], \quad (14)$$

$$I(P_{A,n} || P_{H,n}) = EP_{A,n}[\log Z_n], \quad (15)$$

from which one can immediately build $I_\lambda(P_{A,n} || P_{H,n})$ ($\lambda \in \mathbb{R}$) respectively $R_\lambda(P_{A,n} || P_{H,n})$ ($\lambda \in \mathbb{R} \setminus \{0, 1\}$) respectively bounds of $V(P_{A,n} || P_{H,n})$ via (2) respectively (7) respectively (12).

The outcoming values (respectively bounds) of $H_\lambda(P_{A,n} || P_{H,n})$ are quite diverse and depend on the choice of the involved parameter pairs (β_H, α_H) , (β_A, α_A) as well as λ ; the exact details will be given in the Sections 3 and 6 below.

Before we achieve this, in the following we explain how the outcoming dissimilarity results can be applied to Bayesian testing and more general Bayesian decision making, as well as to Neyman-Pearson testing.

2.5. Decision Making under Uncertainty

Within the above-mentioned context of two competing models (\mathcal{H}) and (\mathcal{A}) of Galton-Watson processes with immigration, let us briefly discuss how knowledge about the time-evolution of the Hellinger integrals $H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ —or equivalently, of the power divergences $I_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n})$, cf. (2)—can be used in order to take decisions under uncertainty, within a framework of Bayesian decision making BDM, or alternatively, of Neyman-Pearson testing NPT.

In our context of BDM, we decide between an action $d_{\mathcal{H}}$ “associated with” the (say) hypothesis law $P_{\mathcal{H}}$ and an action $d_{\mathcal{A}}$ “associated with” the (say) alternative law $P_{\mathcal{A}}$, based on the sample path observation $\mathcal{X}_n := \{X_l : l \in \{0, 1, \dots, n\}\}$ of the GWI-generation-sizes (e.g., infectious-disease incidences, cf. Section 2.3) up to observation horizon $n \in \mathbb{N}$. Following the lines of Stummer & Vajda [15] (adapted to our branching process context), for our BDM let us consider as admissible decision rules $\delta_n : \Omega_n \mapsto \{d_{\mathcal{H}}, d_{\mathcal{A}}\}$ the ones generated by all path sets $G_n \in \Omega_n$ (where Ω_n denotes the space of all possible paths of $(X_k)_{k \in \{1, \dots, n\}}$) through

$$\delta_n(\mathcal{X}_n) := \delta_{G_n}(\mathcal{X}_n) := \begin{cases} d_{\mathcal{A}}, & \text{if } \mathcal{X}_n \in G_n, \\ d_{\mathcal{H}}, & \text{if } \mathcal{X}_n \notin G_n, \end{cases}$$

as well as loss functions of the form

$$\begin{pmatrix} L(d_{\mathcal{H}}, \mathcal{H}) & L(d_{\mathcal{H}}, \mathcal{A}) \\ L(d_{\mathcal{A}}, \mathcal{H}) & L(d_{\mathcal{A}}, \mathcal{A}) \end{pmatrix} := \begin{pmatrix} 0 & L_{\mathcal{A}} \\ L_{\mathcal{H}} & 0 \end{pmatrix} \tag{16}$$

with pre-given constants $L_{\mathcal{A}} > 0, L_{\mathcal{H}} > 0$ (e.g., arising as bounds from quantities in worst-case scenarios); notice that in (16), $d_{\mathcal{H}}$ is assumed to be a zero-loss action under \mathcal{H} and $d_{\mathcal{A}}$ a zero-loss action under \mathcal{A} . Per definition, the *Bayes decision rule* $\delta_{G_{n,\min}}$ minimizes—over G_n —the *mean decision loss*

$$\begin{aligned} \mathcal{L}(\delta_{G_n}) &:= p_{\mathcal{H}}^{\text{prior}} \cdot \mathcal{L}_{\mathcal{H}} \cdot Pr(\delta_{G_n}(\mathcal{X}_n) = d_{\mathcal{A}} | \mathcal{H}) + p_{\mathcal{A}}^{\text{prior}} \cdot \mathcal{L}_{\mathcal{A}} \cdot Pr(\delta_{G_n}(\mathcal{X}_n) = d_{\mathcal{H}} | \mathcal{A}) \\ &= p_{\mathcal{H}}^{\text{prior}} \cdot \mathcal{L}_{\mathcal{H}} \cdot P_{\mathcal{H},n}(G_n) + p_{\mathcal{A}}^{\text{prior}} \cdot \mathcal{L}_{\mathcal{A}} \cdot P_{\mathcal{A},n}(\Omega_n - G_n) \end{aligned} \tag{17}$$

for given prior probabilities $p_{\mathcal{H}}^{\text{prior}} = Pr(\mathcal{H}) \in]0, 1[$ for \mathcal{H} and $p_{\mathcal{A}}^{\text{prior}} := Pr(\mathcal{A}) = 1 - p_{\mathcal{H}}^{\text{prior}}$ for \mathcal{A} . As a side remark let us mention that, in a certain sense, the involved model (parameter) uncertainty expressed by the “superordinate” Bernoulli-type law $Pr = Bin(1, p_{\mathcal{H}}^{\text{prior}})$ can also be reinterpreted as a rudimentary static random environment caused e.g., by a random Bernoulli-type external static force. By straightforward calculations, one gets with (13) the minimizing path set $G_{n,\min} = \left\{ Z_n \geq \frac{p_{\mathcal{H}}^{\text{prior}} L_{\mathcal{H}}}{p_{\mathcal{A}}^{\text{prior}} L_{\mathcal{A}}} \right\}$ leading to the *minimal mean decision loss*, i.e., the *Bayes risk*,

$$\mathcal{R}_n := \min_{G_n} \mathcal{L}(\delta_{G_n}) = \mathcal{L}(\delta_{G_{n,\min}}) = \int_{\Omega_n} \min \left\{ p_{\mathcal{H}}^{\text{prior}} L_{\mathcal{H}}, p_{\mathcal{A}}^{\text{prior}} L_{\mathcal{A}} Z_n \right\} dP_{\mathcal{H},n}. \tag{18}$$

Notice that—by straightforward standard arguments—the *alternative* decision procedure

$$\text{take action } d_{\mathcal{A}} \text{ (resp. } d_{\mathcal{H}}) \text{ if } L_{\mathcal{H}} \cdot p_{\mathcal{H}}^{\text{post}}(\mathcal{X}_n) \leq \text{(resp. } >) L_{\mathcal{A}} \cdot p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n)$$

with posterior probabilities $p_{\mathcal{H}}^{\text{post}}(\mathcal{X}_n) := \frac{p_{\mathcal{H}}^{\text{prior}}}{(1 - p_{\mathcal{H}}^{\text{prior}}) \cdot Z_n(\mathcal{X}_n) + p_{\mathcal{H}}^{\text{prior}}} =: 1 - p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n)$, leads exactly to the same actions as $\delta_{G_{n,\min}}$. By adapting the Lemma 6.5 of Stummer & Vajda [15]—which on general probability spaces gives *fundamental universal* inequalities relating Hellinger integrals (or equivalently, power divergences) and Bayes risks—one gets for all $L_{\mathcal{H}} > 0, L_{\mathcal{A}} > 0, p_{\mathcal{H}}^{\text{prior}} \in]0, 1[, \lambda \in]0, 1[$ and $n \in \mathbb{N}$ the upper bound

$$\mathcal{R}_n \leq \Lambda_{\mathcal{A}}^\lambda \Lambda_{\mathcal{H}}^{1-\lambda} H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}), \quad \text{with } \Lambda_{\mathcal{H}} := p_{\mathcal{H}}^{\text{prior}} L_{\mathcal{H}}, \Lambda_{\mathcal{A}} := (1 - p_{\mathcal{H}}^{\text{prior}}) L_{\mathcal{A}}, \tag{19}$$

as well as the lower bound

$$(\mathcal{R}_n)^{\min\{\lambda, 1-\lambda\}} \cdot (\Lambda_{\mathcal{H}} + \Lambda_{\mathcal{A}} - \mathcal{R}_n)^{\max\{\lambda, 1-\lambda\}} \geq \Lambda_{\mathcal{A}}^{\lambda} \Lambda_{\mathcal{H}}^{1-\lambda} H_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$$

which implies in particular the “direct” lower bound

$$\mathcal{R}_n \geq \frac{\Lambda_{\mathcal{A}}^{\max\{1, \frac{\lambda}{1-\lambda}\}} \Lambda_{\mathcal{H}}^{\max\{1, \frac{1-\lambda}{\lambda}\}}}{(\Lambda_{\mathcal{A}} + \Lambda_{\mathcal{H}})^{\max\{\frac{\lambda}{1-\lambda}, \frac{1-\lambda}{\lambda}\}}} \cdot (H_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n}))^{\max\{\frac{1}{\lambda}, \frac{1}{1-\lambda}\}}. \quad (20)$$

By using (19) (respectively (20)) together with the exact values and the upper (respectively lower) bounds of the Hellinger integrals $H_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ derived in the following sections, we end up with upper (respectively lower) bounds of the Bayes risk \mathcal{R}_n . Of course, with the help of (2) the bounds (19) and (20) can be (i) immediately rewritten in terms of the power divergences $I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ and (ii) thus be *directly* interpreted in terms of dissimilarity-size arguments. As a side-remark, in such a Bayesian context the λ -order Hellinger integral $H_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = EP_{\mathcal{H},n}[(Z_n)^{\lambda}]$ (cf. (14)) can be also interpreted as λ -order Bayes-factor moment (with respect to $P_{\mathcal{H},n}$), since $Z_n = Z_n(\mathcal{X}_n) = \frac{p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n)}{p_{\mathcal{H}}^{\text{post}}(\mathcal{X}_n)} \bigg/ \frac{p_{\mathcal{A}}^{\text{prior}}}{p_{\mathcal{H}}^{\text{prior}}}$ is the Bayes factor (i.e., the posterior odds ratio of (\mathcal{A}) to (\mathcal{H}) , divided by the prior odds ratio of (\mathcal{A}) to (\mathcal{H})).

At this point, the potential applicant should be warned about the *usual way of* asynchronous decision making, where one first *tests* (\mathcal{A}) versus (\mathcal{H}) (i.e., $L_{\mathcal{A}} = L_{\mathcal{H}} = 1$ which leads to 0–1 losses in (16)) and afterwards, based on the outcoming result (e.g., in favour of (\mathcal{A})), takes the attached economic decision (e.g., $d_{\mathcal{A}}$); this can lead to distortions compared with synchronous decision making with “full” monetary losses $L_{\mathcal{A}}$ and $L_{\mathcal{H}}$, as is shown in Stummer & Lao [16] within an economic context in connection with discrete approximations of financial diffusion processes (they call this distortion effect a *non-commutativity between Bayesian statistical and investment decisions*).

For different types of—mainly parameter estimation (squared-error type loss function) concerning—Bayesian analyses based on GW(I) generation size observations, see e.g., Jagers [56], Heyde [38], Heyde & Johnstone [110], Johnson et al. [111], Basawa & Rao [60], Basawa & Scott [61], Scott [112], Guttorp [52], Yanév & Tsokos [113], Mendoza & Gutierrez-Pena [114], and the references therein.

Within our running-example epidemiological context of Section 2.3, let us briefly discuss the role of the above-mentioned losses $L_{\mathcal{A}}$ and $L_{\mathcal{H}}$. To begin with, as mentioned above the *unit-free* choice $L_{\mathcal{A}} = L_{\mathcal{H}} = 1$ corresponds to *Bayesian testing*. Recall that this concerns with two alternative infectious-disease models (\mathcal{H}) and (\mathcal{A}) with parameter pairs (recall the interpretation of β_{\bullet} as reproduction number and α_{\bullet} as importation mean) $(\beta_{\mathcal{H}}, \alpha_{\mathcal{H}})$ and $(\beta_{\mathcal{A}}, \alpha_{\mathcal{A}})$ which reflect either a “pure” statistical uncertainty (under the *same* uncontrolled or controlled set-up), or the uncertainty between two *different* potential control set-ups (for the sake of assessing the potential impact/efficiency of some planned interventions, compared with alternative ones). As far as *non-unit-free*—e.g., macroeconomic or monetary—losses is concerned, recall that some of the above-mentioned control strategies (countermeasures, public policies, governmental pandemic risk management plans) may imply considerable social and economic costs, with a huge impact and potential danger of triggering severe social, economic and political disruptions; a corresponding tradeoff between health and economic issues can be incorporated by choosing $L_{\mathcal{A}}$ and $L_{\mathcal{H}}$ to be (e.g., monetary) values which reflect estimates or upper bounds of losses due to wrong decisions, e.g., if at stage n due to the observed data one erroneously thinks (reinforced by fear) that a novel infectious disease (e.g., COVID-19) will lead (or re-emerge) to a severe pandemic and consequently decides for a lockdown with drastic future economic consequences, versus, if one erroneously thinks (reinforced by carelessness) that the infectious disease is (or stays) non-severe and consequently eases some/all control measures which will lead to extremely devastating future economic consequences. For the estimates/bounds of $L_{\mathcal{A}}$ and $L_{\mathcal{H}}$, one can e.g., employ (i) the comprehensive stochastic studies of Feicht & Stummer [115] on the

quantitative degree of elasticity and speed of recovery of economies after a sudden macroeconomic disaster, or (ii) the more short-term, German-specific, scenario-type (basically non-stochastic) studies of Dorn et al. [116,117] in connection with the current COVID-19 pandemic.

Of course, the above-mentioned Bayesian decision procedure can be also operated in *sequential way*. For instance, suppose that we are encountered with a novel infectious disease (e.g., COVID-19) of non-negligible fatality rate and let (\mathcal{A}) reflect a “potentially dangerous” infectious-disease-transmission situation (e.g., a reproduction number of substantially supercritical case $\beta_{\mathcal{A}} = 2$, and an importation mean of $\alpha_{\mathcal{A}} = 10$, for *weekly* appearing new incidence-generations) whereas (\mathcal{H}) describes a “relatively harmless/mild” situation (e.g., a substantially subcritical $\beta_{\mathcal{H}} = 0.5$, $\alpha_{\mathcal{H}} = 0.2$). Moreover, let $d_{\mathcal{A}}$ respectively $d_{\mathcal{H}}$ denote (non-quantitatively) the decision/action to accept (\mathcal{A}) respectively (\mathcal{H}). It can then be reasonable to decide to stop the observation process $n \mapsto \mathcal{X}_n$ (also called *surveillance* or *online-monitoring*) of incidence numbers at the first time at which $n \mapsto Z_n = Z_n(\mathcal{X}_n)$ exceeds the threshold $p_{\mathcal{H}}^{\text{prior}}/p_{\mathcal{A}}^{\text{prior}}$; if this happens, one takes $d_{\mathcal{A}}$ as decision (and e.g., declare the situation as *occurrence of an epidemic outbreak* and start with control/intervention measures (however, as explained above, one should synchronously involve also the potential economic losses)) whereas as long as this does not happen, one continues the observation (and implicitly takes $d_{\mathcal{H}}$ as decision). This can be modelled in terms of the pair $(\tilde{\tau}, d_{\mathcal{A}})$ with (random) stopping time $\tilde{\tau} := \inf \{n \in \mathcal{N} : Z_n \geq \frac{p_{\mathcal{H}}^{\text{prior}}}{p_{\mathcal{A}}^{\text{prior}}}\}$ (with the usual convention that the infimum of the empty set is infinity), and the corresponding decision $d_{\mathcal{A}}$. After the time $\tilde{\tau} < \infty$ and e.g., immediate subsequent employment of some control/counter measures, one can e.g., take the old model (\mathcal{A}) as new (\mathcal{H}), declare a new target (\mathcal{A}) for the desired quantification of the effectiveness of the employed control measures (e.g., a mitigation to a slightly subcritical case of $\beta_{\mathcal{A}} = 0.95$, $\alpha_{\mathcal{H}} = 0.8$), and starts to observe the new incidence numbers until the new target (\mathcal{A}) has been reached. This can be interpreted as online-detection of a distributional change; a related comprehensive new framework for the use of divergences (even much beyond power divergences) for distributional change detection can be found e.g., in the recent work of Kißlinger & Stummer [118]. A completely different, SIR-model based, approach for the detection of change points in the spread of COVID-19 is given in Dehning et al. [119]. Moreover, other different surveillance methods can be also found e.g., in the corresponding overview of Frisen [120] and the Swedish epidemics outbreak investigations of Friesen & Andersson & Schiöler [121].

One can refine the above-mentioned sequential procedure via two (instead of one) appropriate thresholds $c_1 < c_2$ and the pair $(\tilde{\tau}, \delta_{\tilde{\tau}})$, with the stopping time $\tilde{\tau} := \inf \{n \in \mathcal{N} : Z_n \notin [c_1, c_2]\}$ as well as corresponding decision rule

$$\delta_{\tilde{\tau}} := \begin{cases} d_{\mathcal{A}}, & \text{if } Z_{\tilde{\tau}} > c_2, \\ d_{\mathcal{H}}, & \text{if } Z_{\tilde{\tau}} < c_1. \end{cases}$$

An exact optimized treatment on the two above-mentioned sequential procedures, and their connection to Hellinger integrals (and power divergences) of Galton-Watson processes with immigration, is beyond the scope of this paper.

As a side remark, let us mention that our above-mentioned suggested method of Bayesian decision making with Hellinger integrals of GWIs differs completely from the very recent work of Brauner et al. [122] who use a Bayesian hierarchical model for the concrete, very comprehensive study on the effectiveness and burden of non-pharmaceutical interventions against COVID-19 transmission.

The power divergences $I_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ ($\lambda \in \mathbb{R}$) can be employed also in other ways within Bayesian decision making, of statistical nature. Namely, by adapting the general lines of Österreicher & Vajda [123] (see also Liese & Vajda [10], as well as diffusion-process applications in Stummer [5,31,32]) to our context of Galton-Watson processes with immigration, we can proceed as follows. For the sake of comfortable notations, we first attach the value $\theta := 1$ to the GWI model (\mathcal{A}) (which has prior probability $p_{\mathcal{A}}^{\text{prior}} \in]0, 1[$) and $\theta := 0$ to (\mathcal{H}) (which has prior probability $1 - p_{\mathcal{A}}^{\text{prior}}$). Suppose we

want to decide, in an optimal Bayesian way, which *degree of evidence* $\text{deg} \in [0, 1]$ we should attribute (according to a pre-given *loss function* \mathcal{LO}) to the model (\mathcal{A}) . In order to achieve this goal, we choose a nonnegatively-valued loss function $\mathcal{LO}(\theta, \text{deg})$ defined on $\{0, 1\} \times [0, 1]$, of two types which will be specified below. The risk at stage 0 (i.e., prior to the GWI-path observations \mathcal{X}_n), from the optimal decision about the degree of evidence deg concerning the decision parameter θ , is defined as

$$\mathcal{BR}_{\mathcal{LO}}(p_{\mathcal{A}}^{\text{prior}}) := \min_{\text{deg} \in [0,1]} \left\{ (1 - p_{\mathcal{A}}^{\text{prior}}) \cdot \mathcal{LO}(0, \text{deg}) + p_{\mathcal{A}}^{\text{prior}} \cdot \mathcal{LO}(1, \text{deg}) \right\},$$

which can be thus interpreted as a *minimal prior expected loss* (the minimum will always exist). The corresponding risk *posterior* to the GWI-path observations \mathcal{X}_n , from the optimal decision about the degree of evidence deg concerning the parameter θ , is given by

$$\mathcal{BR}_{\mathcal{LO}}^{\text{post}}(p_{\mathcal{A}}^{\text{prior}}) := \int_{\Omega_n} \mathcal{BR}_{\mathcal{LO}}(p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n)) (p_{\mathcal{A}}^{\text{prior}} dP_{\mathcal{A},n} + (1 - p_{\mathcal{A}}^{\text{prior}}) dP_{\mathcal{H},n}),$$

which is achieved by the optimal decision rule (about the degree of evidence)

$$\mathfrak{D}^*(\mathcal{X}_n) := \arg \min_{\text{deg} \in [0,1]} \left\{ (1 - p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n)) \cdot \mathcal{LO}(0, \text{deg}) + p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n) \cdot \mathcal{LO}(1, \text{deg}) \right\}.$$

The corresponding *statistical information measure* (in the sense of De Groot [124])

$$\Delta \mathcal{BR}_{\mathcal{LO}}(p_{\mathcal{A}}^{\text{prior}}) := \mathcal{BR}_{\mathcal{LO}}(p_{\mathcal{A}}^{\text{prior}}) - \mathcal{BR}_{\mathcal{LO}}^{\text{post}}(p_{\mathcal{A}}^{\text{prior}}) \geq 0$$

represents the *reduction of the decision risk* about the degree of evidence deg concerning the parameter θ , that can be attained by observing the GWI-path \mathcal{X}_n until stage n . For the first-type loss function $\widetilde{\mathcal{LO}}(\theta, \text{deg}) := \text{deg} - (2 \text{deg} - 1) \cdot \mathbf{1}_{\{1\}}(\theta)$, defined on $\{0, 1\} \times [0, 1]$ with the help of the indicator function $\mathbf{1}_A(\cdot)$ on the set A , one can show that

$$\mathfrak{D}^*(\mathcal{X}_n) := \begin{cases} 0, & \text{if } p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n) \in [0, \frac{1}{2}[\\ 1, & \text{if } p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n) \in]\frac{1}{2}, 1[\\ \text{any number in } [0, 1], & \text{if } p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n) = \frac{1}{2}, \end{cases}$$

as well as the representation formula

$$I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = \int_0^1 \Delta \mathcal{BR}_{\widetilde{\mathcal{LO}}}(p_{\mathcal{A}}^{\text{prior}}) \cdot (1 - p_{\mathcal{A}}^{\text{prior}})^{\lambda-2} \cdot (p_{\mathcal{A}}^{\text{prior}})^{-1-\lambda} dp_{\mathcal{A}}^{\text{prior}}, \quad \lambda \in \mathbb{R}, \quad (21)$$

(cf. Österreicher & Vajda [123], Liese & Vajda [10], adapted to our GWI context); in other words, the power divergence $I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ can be regarded as a *weighted-average statistical information measure* (*weighted-average decision risk reduction*). One can also use other weights of $p_{\mathcal{A}}^{\text{prior}}$ in order to get bounds of $I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ (analogously to Stummer [5]).

For the second-type loss function $\mathcal{LO}_{\lambda,\chi}(\theta, \text{deg}) := \frac{\lambda^{\theta-1} \text{deg}^{\lambda-\theta}}{\chi^{\lambda} (1-\chi)^{1-\lambda} (1-\lambda)^{\theta} (1-\text{deg})^{\lambda-\theta}}$ defined on $\{0, 1\} \times [0, 1]$ with parameters $\lambda \in]0, 1[$ and $\chi \in]0, 1[$, one can derive the optimal decision rule

$$\mathfrak{D}^*(\mathcal{X}_n) = p_{\mathcal{A}}^{\text{post}}(\mathcal{X}_n)$$

as well as the representation formula as a *limit statistical information measure* (*limit decision risk reduction*)

$$I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = \lim_{\chi \rightarrow p_{\mathcal{A}}^{\text{prior}}} \Delta \mathcal{BR}_{\mathcal{LO}_{\lambda,\chi}}(p_{\mathcal{A}}^{\text{prior}}) =: \Delta \mathcal{BR}_{\mathcal{LO}_{\lambda,p_{\mathcal{A}}^{\text{prior}}}}(p_{\mathcal{A}}^{\text{prior}}) \quad (22)$$

(cf. Österreicher & Vajda [123], Stummer [5], adapted to our GWI context).

As an alternative to the above-mentioned Bayesian-decision-making applications of Hellinger integrals $H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n})$, let us now briefly discuss the use of the latter for the corresponding Neyman-Pearson (NPT) framework with randomized tests $\mathcal{T}_n : \Omega_n \mapsto [0, 1]$ of the hypothesis $P_{\mathcal{H}}$ against the alternative $P_{\mathcal{A}}$, based on the GWI-generation-size sample path observations $\mathcal{X}_n := \{X_l : l \in \{0, 1, \dots, n\}\}$. In contrast to (17) and (18) a Neyman-Pearson test minimizes—over \mathcal{T}_n —the type II error probability $\int_{\Omega_n} (1 - \mathcal{T}_n) dP_{\mathcal{A},n}$ in the class of the tests for which the type I error probability $\int_{\Omega_n} \mathcal{T}_n dP_{\mathcal{H},n}$ is at most $\zeta \in]0, 1[$. The corresponding minimal type II error probability

$$\mathcal{E}_\zeta(P_{\mathcal{A},i}||P_{\mathcal{H},i}) := \inf_{\mathcal{T}_i: \int_{\Omega_i} \mathcal{T}_i dP_{\mathcal{H},i} \leq \zeta} \int_{\Omega_i} (1 - \mathcal{T}_i) dP_{\mathcal{A},i}$$

can for all $\zeta \in]0, 1[$, $\lambda \in]0, 1[$, $i \in \mathcal{I}$ be bounded from above by

$$\mathcal{E}_\zeta(P_{\mathcal{A},i}||P_{\mathcal{H},i}) \leq \mathcal{E}_\zeta^U(P_{\mathcal{A},i}||P_{\mathcal{H},i}) := \min \left\{ (1 - \lambda) \cdot \left(\frac{\lambda}{\zeta}\right)^{\lambda/(1-\lambda)} \cdot \left(H_\lambda(P_{\mathcal{A},i}||P_{\mathcal{H},i})\right)^{1/(1-\lambda)}, 1 \right\}, \quad (23)$$

and for all $\lambda > 1$, $i \in \mathcal{I}$ it can be bounded from below by

$$\mathcal{E}_\zeta(P_{\mathcal{A},i}||P_{\mathcal{H},i}) \geq \mathcal{E}_\zeta^L(P_{\mathcal{A},i}||P_{\mathcal{H},i}) := (1 - \zeta)^{\lambda/(\lambda-1)} \cdot \left(H_\lambda(P_{\mathcal{A},i}||P_{\mathcal{H},i})\right)^{1/(1-\lambda)}, \quad (24)$$

which is an adaption of a general result of Krafft & Plachky [125], see also Liese & Vajda [1] as well as Stummer & Vajda [15]. Hence, by combining (23) and (24) with the exact values respectively upper bounds of the Hellinger integrals $H_{1-\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ from the following sections, we obtain for our context of Galton-Watson processes with Poisson offspring and Poisson immigration (including the non-immigration case) some upper bounds of $\mathcal{E}_\zeta(P_{\mathcal{A},n}||P_{\mathcal{H},n})$, which can also be immediately rewritten as lower bounds for the power $1 - \mathcal{E}_\zeta(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ of a most powerful test at level ζ . In contrast to such finite-time-horizon results, for the (to our context) incompatible setup of Galton-Watson processes with Poisson offspring but nonstochastic immigration of constant value 1, the asymptotic rates of decrease as $n \rightarrow \infty$ of the unconstrained type II error probabilities as well as the type I error probabilities were studied in Linkov & Lunyova [53] by a different approach employing also Hellinger integrals. Some other types of Galton-Watson-process concerning Neyman-Pearson testing investigations different to ours can be found e.g., in Basawa & Scott [126], Feigin [127], Sweeting [128], Basawa & Scott [61], and the references therein.

2.6. Asymptotical Distinguishability

The next two concepts deal with two general families $(P_{\mathcal{A},i})_{i \in \mathcal{I}}$ and $(P_{\mathcal{H},i})_{i \in \mathcal{I}}$ of probability measures on the measurable spaces $(\Omega_i, \mathcal{F}_i)_{i \in \mathcal{I}}$, where the index set \mathcal{I} is either \mathbb{N}_0 or \mathbb{R}_+ . For them, the following two general types of asymptotical distinguishability are well known (see e.g., LeCam [109], Liese & Vajda [1], Jacod & Shiryaev [24], Linkov [129], and the references therein).

Definition 1. The family $(P_{\mathcal{A},i})_{i \in \mathcal{I}}$ is contiguous to the family $(P_{\mathcal{H},i})_{i \in \mathcal{I}}$ —in symbols, $(P_{\mathcal{A},i}) \triangleleft (P_{\mathcal{H},i})$ —if for all sets $A_i \in \mathcal{F}_i$ with $\lim_{i \rightarrow \infty} P_{\mathcal{H},i}(A_i) = 0$ there holds $\lim_{i \rightarrow \infty} P_{\mathcal{A},i}(A_i) = 0$.

Definition 2. Families of measures $(P_{\mathcal{A},i})_{i \in \mathcal{I}}$ and $(P_{\mathcal{H},i})_{i \in \mathcal{I}}$ are called entirely separated (completely asymptotically distinguishable)—in symbols, $(P_{\mathcal{A},i}) \triangle (P_{\mathcal{H},i})$ —if there exist a sequence $i_m \uparrow \infty$ as $m \uparrow \infty$ and for each $m \in \mathbb{N}_0$ an $A_{i_m} \in \mathcal{F}_{i_m}$ such that $\lim_{m \rightarrow \infty} P_{\mathcal{A},i_m}(A_{i_m}) = 1$ and $\lim_{m \rightarrow \infty} P_{\mathcal{H},i_m}(A_{i_m}) = 0$.

It is clear that the notion of contiguity is the attempt to carry the concept of absolute continuity over to families of measures. Loosely speaking, $(P_{\mathcal{A},i})$ is contiguous to $(P_{\mathcal{H},i})$, if the limit $\lim_{i \rightarrow \infty} (P_{\mathcal{A},i})$ (existence preconditioned) is absolute continuous to the limit $\lim_{i \rightarrow \infty} (P_{\mathcal{H},i})$. However, for the definition

of contiguity, we do not need to require the probability measures to converge to limiting probability measures. On the other hand, entire separation is the generalization of singularity to families of measures.

The corresponding negations will be denoted by $\overleftarrow{\triangleleft}$ and $\overleftarrow{\triangle}$. One can easily check that a family $(P_{\mathcal{A},i})$ cannot be both contiguous and entirely separated to a family $(P_{\mathcal{H},i})$. In fact, as shown in Linkov [129], the relation between the families $(P_{\mathcal{A},i})$ and $(P_{\mathcal{H},i})$ can be uniquely classified into the following *distinguishability types*:

- (a) $(P_{\mathcal{A},i}) \triangleleft \triangleright (P_{\mathcal{H},i})$;
- (b) $(P_{\mathcal{A},i}) \triangleleft (P_{\mathcal{H},i}), (P_{\mathcal{H},i}) \overleftarrow{\triangleleft} (P_{\mathcal{A},i})$;
- (c) $(P_{\mathcal{A},i}) \overleftarrow{\triangleleft} (P_{\mathcal{H},i}), (P_{\mathcal{H},i}) \triangleleft (P_{\mathcal{A},i})$;
- (d) $(P_{\mathcal{A},i}) \overleftarrow{\triangleleft} \overleftarrow{\triangleright} (P_{\mathcal{H},i}), (P_{\mathcal{A},i}) \overleftarrow{\triangle} (P_{\mathcal{H},i})$;
- (e) $(P_{\mathcal{A},i}) \triangle (P_{\mathcal{H},i})$.

As demonstrated in the above-mentioned references for a general context, one can conclude the type of distinguishability from the time-evolution of Hellinger integrals. Indeed, the following assertions can be found e.g., in Linkov [129], where part (c) was established in Liese & Vajda [1] and (f), (g) in Vajda [3].

Proposition 1. *The following assertions are equivalent:*

- (a) $(P_{\mathcal{A},i}) \triangle (P_{\mathcal{H},i})$,
- (b) $\liminf_{i \rightarrow \infty} H_\lambda(P_{\mathcal{A},i} || P_{\mathcal{H},i}) = 0$ for all $\lambda \in]0, 1[$,
- (c) there exists a $\lambda \in]0, 1[$: $\liminf_{i \rightarrow \infty} H_\lambda(P_{\mathcal{A},i} || P_{\mathcal{H},i}) = 0$,
- (d) there exists a $\pi \in]0, 1[$: $\liminf_{i \rightarrow \infty} e_\pi(P_{\mathcal{A},i} || P_{\mathcal{H},i}) = 0$,
- (e) $\limsup_{i \rightarrow \infty} V(P_{\mathcal{A},i} || P_{\mathcal{H},i}) = 2$,
- (f) there exists a $\lambda \in]0, 1[$: $\limsup_{i \rightarrow \infty} I_\lambda(P_{\mathcal{A},i} || P_{\mathcal{H},i}) = \frac{1}{\lambda \cdot (1 - \lambda)}$,
- (g) $\limsup_{i \rightarrow \infty} I_\lambda(P_{\mathcal{A},i} || P_{\mathcal{H},i}) = \frac{1}{\lambda \cdot (1 - \lambda)}$, for all $\lambda \in]0, 1[$.

In combination with the discussion after Definition 2, one can thus interpret the λ -order Hellinger integral $H_\lambda(P_{\mathcal{A},i} || P_{\mathcal{H},i})$ as a “measure” for the distinctness of the two families $P_{\mathcal{A},i}$ and $P_{\mathcal{H},i}$ up to a fixed finite time horizon $i \in \mathcal{I}$.

Furthermore, for the contiguity we obtain the equivalence (see e.g., Liese & Vajda [1], Linkov [129])

$$\begin{aligned}
 (P_{\mathcal{A},i}) \triangleleft (P_{\mathcal{H},i}) &\iff \liminf_{\lambda \nearrow 1} \left\{ \liminf_{i \rightarrow \infty} H_\lambda(P_{\mathcal{A},i} || P_{\mathcal{H},i}) \right\} = 1 \\
 &\iff \limsup_{\lambda \nearrow 1} \left\{ \limsup_{i \rightarrow \infty} \lambda \cdot (1 - \lambda) \cdot I_\lambda(P_{\mathcal{A},i} || P_{\mathcal{H},i}) \right\} = 0.
 \end{aligned}
 \tag{26}$$

All the above-mentioned general results can be applied to our context of two competing Poissonian Galton-Watson processes with immigration (GWI) (\mathcal{H}) and (\mathcal{A}) (reflected by the two different laws $P_{\mathcal{H}}$ resp. $P_{\mathcal{A}}$ with parameter pairs $(\beta_{\mathcal{H}}, \alpha_{\mathcal{H}})$ resp. $(\beta_{\mathcal{A}}, \alpha_{\mathcal{A}})$), by taking $P_{\mathcal{A},i} := P_{\mathcal{A}}|_{\mathcal{F}_i}$ and $P_{\mathcal{H},i} := P_{\mathcal{H}}|_{\mathcal{F}_i}$. Recall from the preceding subsections (by identifying i with n) that the latter two describe the stochastic dynamics of the respective GWI within the restricted time-/stage-frame $\{0, 1, \dots, i\}$.

In the following, we study in detail the evolution of Hellinger integrals between two competing models of Galton-Watson processes with immigration, which turns out to be quite extensive.

3. Detailed Recursive Analyses of Hellinger Integrals

3.1. A First Basic Result

In terms of our notations (PS1) to (PS3), a typical situation for applications in our mind is that one particular constellation $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}$ (e.g., obtained from theoretical or previous statistical investigations) is fixed, whereas—in contrast—the parameter $\lambda \in \mathbb{R} \setminus \{0, 1\}$ for the Hellinger integral or the power divergence might be chosen freely, e.g., depending on which (transform of a) dissimilarity measure one decides to choose for further analysis. At this point, let us emphasize that *in general* we will not make assumptions of the form $\beta_{\bullet} \gtrless 1$, i.e., upon the type of criticality.

To start with our investigations, in order to justify for all $n \in \mathbb{N}_0$

$$Z_n := \frac{dP_{\mathcal{A},n}}{dP_{\mathcal{H},n}} \quad (\text{cf. (13)}),$$

(14) and (15) (as well as $I_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ for $\lambda \in \mathbb{R}$ respectively $R_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ for $\lambda \in \mathbb{R} \setminus \{0, 1\}$), we first mention the following straightforward facts: (i) if $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{\text{NI}}$, then $P_{\mathcal{A},n}$ and $P_{\mathcal{H},n}$ are equivalent (i.e., $P_{\mathcal{A},n} \sim P_{\mathcal{H},n}$), as well as (ii) if $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{\text{SP}}$, then $P_{\mathcal{A},n}$ and $P_{\mathcal{H},n}$ are equivalent (i.e., $P_{\mathcal{A},n} \sim P_{\mathcal{H},n}$). Moreover, by recalling $Z_0 = 1$ and using the “rate functions” $f_{\bullet}(x) = \beta_{\bullet}x + \alpha_{\bullet}$ ($x \in [0, \infty[$), a version of (13) can be easily determined by calculating for each $\vec{x} := (x_0, x_1, x_2, \dots) \in \Omega := \mathbb{N} \times \mathbb{N}_0 \times \mathbb{N}_0 \times \dots$

$$Z_n(\vec{x}) = \prod_{k=1}^n Z_{n,k}(\vec{x}) \quad \text{with } Z_{n,k}(\vec{x}) := \exp \left\{ - (f_{\mathcal{A}}(x_{k-1}) - f_{\mathcal{H}}(x_{k-1})) \right\} \left[\frac{f_{\mathcal{A}}(x_{k-1})}{f_{\mathcal{H}}(x_{k-1})} \right]^{x_k},$$

where for the last term we use the convention $\binom{0}{0}^x = 1$ for all $x \in \mathbb{N}_0$. Furthermore, we define for each $\vec{x} \in \Omega$

$$Z_{n,k}^{(\lambda)}(\vec{x}) := \exp \left\{ - (\lambda f_{\mathcal{A}}(x_{k-1}) + (1 - \lambda)f_{\mathcal{H}}(x_{k-1})) \right\} \frac{[(f_{\mathcal{A}}(x_{k-1}))^\lambda (f_{\mathcal{H}}(x_{k-1}))^{1-\lambda}]^{x_k}}{x_k!} \quad (27)$$

with the convention $\binom{0}{0}^0 = 1$ for the last term. Accordingly, one obtains from (14) the Hellinger integral $H_\lambda(P_{\mathcal{A},0}||P_{\mathcal{H},0}) = 1$, as well as for all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P} \times (\mathbb{R} \setminus \{0, 1\})$

$$H_\lambda(P_{\mathcal{A},1}||P_{\mathcal{H},1}) = \exp \left\{ (f_{\mathcal{A}}(x_0))^\lambda (f_{\mathcal{H}}(x_0))^{(1-\lambda)} - (\lambda f_{\mathcal{A}}(x_0) + (1 - \lambda)f_{\mathcal{H}}(x_0)) \right\} \quad (28)$$

for $x_0 = X_0 \in \mathbb{N}$, and for all $n \in \mathbb{N} \setminus \{1\}$

$$\begin{aligned} H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) &= EP_{\mathcal{H},n}[(Z_n)^\lambda] = \sum_{x_1=0}^\infty \dots \sum_{x_n=0}^\infty \prod_{k=1}^n Z_{n,k}^{(\lambda)}(\vec{x}) \\ &= \sum_{x_1=0}^\infty \dots \sum_{x_{n-1}=0}^\infty \prod_{k=1}^{n-1} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot e^{-(\lambda f_{\mathcal{A}}(x_{n-1}) + (1-\lambda)f_{\mathcal{H}}(x_{n-1}))} \sum_{x_n=0}^\infty \frac{[(f_{\mathcal{A}}(x_{n-1}))^\lambda (f_{\mathcal{H}}(x_{n-1}))^{1-\lambda}]^{x_n}}{x_n!} \\ &= \sum_{x_1=0}^\infty \dots \sum_{x_{n-1}=0}^\infty \prod_{k=1}^{n-1} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot \exp\{(f_{\mathcal{A}}(x_{n-1}))^\lambda (f_{\mathcal{H}}(x_{n-1}))^{1-\lambda} - (\lambda f_{\mathcal{A}}(x_{n-1}) + (1 - \lambda)f_{\mathcal{H}}(x_{n-1}))\}. \end{aligned} \quad (29)$$

From (29), one can see that a crucial role for the exact calculation (respectively the derivation of bounds) of the Hellinger integral is played by the functions defined for $x \in [0, \infty[$

$$\varphi_\lambda(x) := \varphi(x, \beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) := \varphi_\lambda(x) - f_\lambda(x), \quad \text{with} \quad (30)$$

$$\varphi_\lambda(x) := \varphi(x, \beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) := (f_{\mathcal{A}}(x))^\lambda (f_{\mathcal{H}}(x))^{1-\lambda} \quad \text{and} \quad (31)$$

$$f_\lambda(x) := f(x, \beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) := \lambda f_A(x) + (1 - \lambda) f_H(x) = \alpha_\lambda + \beta_\lambda x, \tag{32}$$

where we have used the λ -weighted-averages

$$\alpha_\lambda := \alpha(\alpha_A, \alpha_H, \lambda) := \lambda \cdot \alpha_A + (1 - \lambda) \cdot \alpha_H \quad \text{and} \quad \beta_\lambda := \beta(\beta_A, \beta_H, \lambda) := \lambda \cdot \beta_A + (1 - \lambda) \cdot \beta_H.$$

Since λ plays a special role, henceforth we typically use it as index and often omit $(\beta_A, \beta_H, \alpha_A, \alpha_H)$. According to Lemma A1 in the Appendix A.1, it follows that for $\lambda \in]0, 1[$ (respectively $\lambda \in \mathbb{R} \setminus]0, 1[$) one gets $\phi_\lambda(x) \leq 0$ (respectively $\phi_\lambda(x) \geq 0$) for all $x \in [0, \infty[$. Furthermore, in both cases there holds $\phi_\lambda(x) = 0$ iff $f_A(x) = f_H(x)$, i.e., for $x = x^* := \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} \geq 0$. This is consistent with the corresponding generally valid upper and lower bounds (cf. (9) and (11)) $0 < H_\lambda(P_{A,n} || P_{H,n}) \leq 1$, for $\lambda \in]0, 1[$, $1 \leq H_\lambda(P_{A,n} || P_{H,n}) \leq \infty$, for $\lambda \in \mathbb{R} \setminus]0, 1[$.

As a first indication for our proposed method, let us start by illuminating the simplest case $\lambda \in \mathbb{R} \setminus \{0, 1\}$ and $\gamma := \alpha_H \beta_A - \alpha_A \beta_H = 0$. This means that $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$, where $\mathcal{P}_{SP,1}$ is the set of all (componentwise) strictly positive $(\beta_A, \beta_H, \alpha_A, \alpha_H)$ with $\beta_A \neq \beta_H, \alpha_A \neq \alpha_H$ and $\frac{\beta_A}{\beta_H} = \frac{\alpha_A}{\alpha_H} \neq 1$ ("the equal-fraction-case"). In this situation, all the three functions (30) to (32) are linear. Indeed,

$$\varphi_\lambda(x) = p_\lambda^E + q_\lambda^E x \tag{33}$$

with $p_\lambda^E := \alpha_A^\lambda \alpha_H^{1-\lambda}$ and $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$ (where the index E stands for exact linearity). Clearly, $q_\lambda^E > 0$ on $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$, as well as $p_\lambda^E > 0$ on $\mathcal{P}_{SP,1}$ and $p_\lambda^E = 0$ on \mathcal{P}_{NI} . Furthermore,

$$\phi_\lambda(x) = r_\lambda^E + s_\lambda^E x$$

with $r_\lambda^E := p_\lambda^E - \alpha_\lambda = \alpha_A^\lambda \alpha_H^{1-\lambda} - (\lambda \alpha_A + (1 - \lambda) \alpha_H)$ and $s_\lambda^E := q_\lambda^E - \beta_\lambda = \beta_A^\lambda \beta_H^{1-\lambda} - (\lambda \beta_A + (1 - \lambda) \beta_H)$. Due to Lemma A1 one knows that on $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$ one gets $s_\lambda^E < 0$ for $\lambda \in]0, 1[$ and $s_\lambda^E > 0$ for $\lambda \in \mathbb{R} \setminus]0, 1[$. Furthermore, on $\mathcal{P}_{SP,1}$ one gets $r_\lambda^E < 0$ (resp. $r_\lambda^E > 0$) for $\lambda \in]0, 1[$ (resp. $\lambda \in \mathbb{R} \setminus]0, 1[$), whereas on \mathcal{P}_{NI} , the no-immigration setup, we get for all $\lambda \in \mathbb{R} \setminus \{0, 1\}$ $r_\lambda^E = 0$.

As it will be seen later on, such kind of linearity properties are useful for the recursive handling of the Hellinger integrals. However, only on the parameter set $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$ the functions φ_λ and ϕ_λ are linear. Hence, in the general case $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P} \times \mathbb{R} \setminus \{0, 1\}$ we aim for linear lower and upper bounds

$$\varphi_\lambda^L(x) := p_\lambda^L + q_\lambda^L x \leq \varphi_\lambda(x) \leq \varphi_\lambda^U(x) := p_\lambda^U + q_\lambda^U x, \tag{34}$$

$x \in [0, \infty[$ (ultimately, $x \in \mathbb{N}_0$), which by (30) and (31) leads to

$$\phi_\lambda(x) \begin{cases} \leq \phi_\lambda^U(x) := r_\lambda^U + s_\lambda^U \cdot x := (p_\lambda^U - \alpha_\lambda) + (q_\lambda^U - \beta_\lambda) \cdot x, \\ \geq \phi_\lambda^L(x) := r_\lambda^L + s_\lambda^L \cdot x := (p_\lambda^L - \alpha_\lambda) + (q_\lambda^L - \beta_\lambda) \cdot x, \end{cases} \tag{35}$$

$x \in [0, \infty[$ (ultimately, $x \in \mathbb{N}_0$). Of course, the involved slopes and intercepts should satisfy reasonable restrictions. Later on, we shall impose further restrictions on the involved slopes and intercepts, in order to guarantee nice properties of the general Hellinger integral bounds given in Theorem 1 below (for instance, in consistency with the nonnegativity of φ_λ we could require $p_\lambda^U \geq p_\lambda^L \geq 0, q_\lambda^U \geq q_\lambda^L \geq 0$ which nontrivially implies that these bounds possess certain monotonicity properties). For the formulation of our first assertions on Hellinger integrals, we make use of the following notation:

Definition 3. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P} \times \mathbb{R} \setminus \{0, 1\}$ and all $p, q \in \mathbb{R}$ let us define the sequences $(a_n^{(q)})_{n \in \mathbb{N}_0}$ and $(b_n^{(p,q)})_{n \in \mathbb{N}_0}$ recursively by

$$a_0^{(q)} := 0 \quad ; \quad a_n^{(q)} := \zeta_\lambda^{(q)} \left(a_{n-1}^{(q)} \right) := q \cdot e^{a_{n-1}^{(q)}} - \beta_\lambda, \quad n \in \mathbb{N}, \tag{36}$$

$$b_0^{(p,q)} := 0 \quad ; \quad b_n^{(p,q)} := p \cdot e^{a_{n-1}^{(q)}} - \alpha_\lambda, \quad n \in \mathbb{N}. \tag{37}$$

Notice the interrelation $a_1^{(q^A)} = s_\lambda^A$ and $b_1^{(p^A, q^A)} = r_\lambda^A$ for $A \in \{E, L, U\}$. Clearly, for all $q \in \mathbb{R} \setminus \{0\}$ and $p \in \mathbb{R}$ one has the linear interrelation

$$b_n^{(p,q)} = \frac{p}{q} a_n^{(q)} + \frac{p}{q} \beta_\lambda - \alpha_\lambda, \quad n \in \mathbb{N}. \tag{38}$$

Accordingly, we obtain fundamental Hellinger integral evaluations:

Theorem 1.

- (a) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}) \times \mathbb{R} \setminus \{0, 1\}$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ one can recursively compute the **exact value**

$$H_\lambda(P_{A,n} || P_{H,n}) = \exp \left\{ a_n^{(q_\lambda^E)} X_0 + \frac{\alpha_A}{\beta_A} \sum_{k=1}^n a_k^{(q_\lambda^E)} \right\} =: V_{\lambda, X_0, n}, \tag{39}$$

where $\frac{\alpha_A}{\beta_A}$ can be equivalently replaced by $\frac{\alpha_H}{\beta_H}$. Recall that $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$. Notice that on $\mathcal{P}_{NI} \times (\mathbb{R} \setminus \{0, 1\})$ the formula (39) simplifies significantly, since $\alpha_A = \alpha_H = 0$.

- (b) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus \{0, 1\})$, all coefficients $p_\lambda^L, p_\lambda^U, q_\lambda^L, q_\lambda^U \in \mathbb{R}$ which satisfy (35) for all $x \in \mathbb{N}_0$ (and thus in particular $p_\lambda^L \leq p_\lambda^U, q_\lambda^L \leq q_\lambda^U$), all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ one gets the following **recursive** (i.e., recursively computable) **bounds** for the Hellinger integrals:

$$\text{for } \lambda \in]0, 1[: \quad B_{\lambda, X_0, n}^L := \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} < H_\lambda(P_{A,n} || P_{H,n}) \leq \min \left\{ \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)}, 1 \right\} =: B_{\lambda, X_0, n}^U, \tag{40}$$

$$\text{for } \lambda \in \mathbb{R} \setminus]0, 1[: \quad B_{\lambda, X_0, n}^L := \max \left\{ \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)}, 1 \right\} \leq H_\lambda(P_{A,n} || P_{H,n}) < \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} =: B_{\lambda, X_0, n}^U, \tag{41}$$

where for general $\lambda \in \mathbb{R} \setminus \{0, 1\}, p \in \mathbb{R}, q \in \mathbb{R} \setminus \{0\}$ we use the definitions

$$\tilde{B}_{\lambda, X_0, n}^{(p,q)} := \exp \left\{ a_n^{(q)} \cdot X_0 + \sum_{k=1}^n b_k^{(p,q)} \right\} = \exp \left\{ a_n^{(q)} \cdot X_0 + \frac{p}{q} \sum_{k=1}^n a_k^{(q)} + n \cdot \left(\frac{p}{q} \beta_\lambda - \alpha_\lambda \right) \right\}, \tag{42}$$

as well as

$$\tilde{B}_{\lambda, X_0, n}^{(p,0)} := \exp \left\{ -\beta_\lambda \cdot X_0 + (p \cdot e^{-\beta_\lambda} - \alpha_\lambda) \cdot n \right\}.$$

Remark 1.

- (a) Notice that the expression $\tilde{B}_{\lambda, X_0, n}^{(p,q)}$ can analogously be defined on the parameter set $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$. For the choices $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda} > 0$ and $p_\lambda^E := \alpha_A^\lambda \alpha_H^{1-\lambda} = q_\lambda^E \cdot \frac{\alpha_A}{\beta_A} = q_\lambda^E \cdot \frac{\alpha_H}{\beta_H} \geq 0$ one gets $(p_\lambda^E / q_\lambda^E) \cdot \beta_\lambda - \alpha_\lambda = 0$, and thus the characterization $\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^E, q_\lambda^E)} = V_{\lambda, X_0, n}$ as the exact value (rather than a lower/upper bound (component)).
- (b) In the case $q = \beta_\lambda$ one gets the explicit representation $\tilde{B}_{\lambda, X_0, n}^{(p,q)} = \exp \{ (p - \alpha_\lambda) \cdot n \}$.
- (c) Using the skew symmetry (8), one can derive alternative bounds of the Hellinger integral by switching to the transformed parameter setup $(\tilde{\beta}_A, \tilde{\beta}_H, \tilde{\alpha}_A, \tilde{\alpha}_H, \tilde{\lambda}) := (\beta_H, \beta_A, \alpha_H, \alpha_A, 1 - \lambda)$. However, this does not lead to different bounds: define $\tilde{\phi}_\lambda, \tilde{\varphi}_\lambda$ and \tilde{f}_λ analogously to (30), (31) and (32) by

replacing the parameters $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ with $(\overleftrightarrow{\beta}_A, \overleftrightarrow{\beta}_H, \overleftrightarrow{\alpha}_A, \overleftrightarrow{\alpha}_H, \overleftrightarrow{\lambda})$. Then, there holds $\overleftrightarrow{f} \overleftrightarrow{\leftarrow} (x) = f_\lambda(x)$, $\overleftrightarrow{\varphi} \overleftrightarrow{\leftarrow} (x) = \varphi_\lambda(x)$ and $\overleftrightarrow{\phi} \overleftrightarrow{\leftarrow} (x) = \phi_\lambda(x)$, and the set of (lower and upper bound) parameters $p_\lambda^L, q_\lambda^L, p_\lambda^U, q_\lambda^U$ satisfying (35) does not change under this transformation.

- (d) If there are no other restrictions on $p_\lambda^L, p_\lambda^U, q_\lambda^L, q_\lambda^U$ than (35), the bounds in (40) and (41) can have some inconvenient features, e.g., being 1 for all (large enough) $n \in \mathbb{N}$, having oscillating n -behaviour, being suboptimal in certain (other) senses. For a detailed discussion, the reader is referred to Section 3.16 ff. below.
- (e) For the (to our context) incompatible setup of GWI with Poisson offspring but nonstochastic immigration of constant value 1, the exact values of the corresponding Hellinger integrals (i.e., an “analogue” of part (a)) was established in Linkov & Lunyova [53].

Proof of Theorem 1. Let us fix $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}$ as well as $x_0 := X_0 \in \mathbb{N}$, and start with arbitrary $\lambda \in]0, 1[$. We first prove the upper bound $B_{\lambda, X_0, n}^U$ of part (b). Correspondingly, we suppose that the coefficients p_λ^U, q_λ^U satisfy (35) for all $x \in \mathbb{N}_0$. From (28), (30), (31), (32) and (35) one gets immediately $B_{\lambda, X_0, 1}^U$ in terms of the first sequence-element $a_1^{(q_\lambda^U)}$ (cf. (36)). With the help of (29) for all observation horizons $n \in \mathbb{N} \setminus \{1\}$ we get (with the obvious shortcut for $n = 2$)

$$\begin{aligned}
 H_\lambda(P_{\mathcal{A}, n} || P_{\mathcal{H}, n}) &= \sum_{x_1=0}^\infty \cdots \sum_{x_{n-1}=0}^\infty \prod_{k=1}^{n-1} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot \exp \left\{ \varphi_\lambda(x_{n-1}) - f_\lambda(x_{n-1}) \right\} \\
 &< \sum_{x_1=0}^\infty \cdots \sum_{x_{n-1}=0}^\infty \prod_{k=1}^{n-1} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot \exp \left\{ (p_\lambda^U - \alpha_\lambda) + (q_\lambda^U - \beta_\lambda) x_{n-1} \right\} \\
 &= \sum_{x_1=0}^\infty \cdots \sum_{x_{n-1}=0}^\infty \prod_{k=1}^{n-1} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot \exp \left\{ b_1^{(p_\lambda^U, q_\lambda^U)} + a_1^{(q_\lambda^U)} x_{n-1} \right\} \\
 &= \exp \left\{ b_1^{(p_\lambda^U, q_\lambda^U)} \right\} \sum_{x_1=0}^\infty \cdots \sum_{x_{n-2}=0}^\infty \prod_{k=1}^{n-2} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot \exp \left\{ \exp \left\{ a_1^{(q_\lambda^U)} \right\} \varphi_\lambda(x_{n-2}) - f_\lambda(x_{n-2}) \right\} \\
 &< \exp \left\{ b_1^{(p_\lambda^U, q_\lambda^U)} \right\} \sum_{x_1=0}^\infty \cdots \sum_{x_{n-2}=0}^\infty \prod_{k=1}^{n-2} Z_{n,k}^{(\lambda)}(\vec{x}) \\
 &\quad \cdot \exp \left\{ \left(\exp \left\{ a_1^{(q_\lambda^U)} \right\} p_\lambda^U - \alpha_\lambda \right) + \left(\exp \left\{ a_1^{(q_\lambda^U)} \right\} q_\lambda^U - \beta_\lambda \right) \cdot x_{n-2} \right\} \\
 &< \exp \left\{ b_1^{(p_\lambda^U, q_\lambda^U)} \right\} \sum_{x_1=0}^\infty \cdots \sum_{x_{n-2}=0}^\infty \prod_{k=1}^{n-2} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot \exp \left\{ b_2^{(p_\lambda^U, q_\lambda^U)} + a_2^{(q_\lambda^U)} x_{n-2} \right\} \\
 &< \cdots < \exp \left\{ a_n^{(q_\lambda^U)} x_0 + \sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)} \right\}. \tag{43}
 \end{aligned}$$

Notice that for the strictness of the above inequalities we have used the fact that $\phi_\lambda(x) < \phi_\lambda^U(x)$ for some (in fact, all but at most two) $x \in \mathbb{N}_0$ (cf. Properties 3(P19) below). Since for some admissible choices of p_λ^U, q_λ^U and some $n \in \mathbb{N}$ the last term in (43) can become larger than 1, one needs to take into account the cutoff-point 1 arising from (9). The lower bound $B_{\lambda, X_0, n}^L$ of part (b), as well as the exact value of part (a) follow from (29) in an analogous manner by employing p_λ^L, q_λ^L and p_λ^E, q_λ^E respectively. Furthermore, we use the fact that for $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}) \times]0, 1[$ one gets from (38) the relation $b_n^{(p_\lambda^E, q_\lambda^E)} = \frac{\alpha_A}{\beta_A} a_n^{(q_\lambda^E)}$. For the sake of brevity, the corresponding straightforward details are omitted here. Although we take the minimum of the upper bound derived in (43) and 1, the inequality $B_{\lambda, X_0, n}^L < B_{\lambda, X_0, n}^U$ is nevertheless valid: the reason is that for constituting a lower bound, the parameters p_λ^L, q_λ^L must fulfill either the conditions $[p_\lambda^L - \alpha_\lambda < 0 \text{ and } q_\lambda^L - \beta_\lambda \leq 0]$ or $[p_\lambda^L - \alpha_\lambda \leq 0 \text{ and } q_\lambda^L - \beta_\lambda < 0]$ (or both), which guarantees that $B_{\lambda, X_0, n}^L < 1$. The proof for all $\lambda \in \mathbb{R} \setminus]0, 1[$ works

out completely analogous, by taking into account the generally valid lower bound $H_\lambda(P_{A,n}||P_{\mathcal{H},n}) \geq 1$ (cf. (11)). \square

3.2. Some Useful Facts for Deeper Analyses

Theorem 1(b) and Remark 1(a) indicate the crucial role of the expression $\tilde{B}_{\lambda, X_0, n}^{(p,q)}$ and that the choice of the quantities p, q depends on the underlying (e.g., fixed) offspring-immigration parameter constellation $(\beta_A, \beta_{\mathcal{H}}, \alpha_A, \alpha_{\mathcal{H}})$ as well as on the (e.g., selectable) value of λ , i.e., $p_\lambda^A = p^A(\beta_A, \beta_{\mathcal{H}}, \alpha_A, \alpha_{\mathcal{H}}, \lambda)$ and $q_\lambda^A = q^A(\beta_A, \beta_{\mathcal{H}}, \alpha_A, \alpha_{\mathcal{H}}, \lambda)$ with $A \in \{E, L, U\}$. In order to study the desired time-behaviour $n \mapsto \tilde{B}_{\lambda, X_0, n}^{(\cdot, \cdot)}$ of the Hellinger integral bounds resp. exact values, one therefore faces a six-dimensional (and thus highly non-obvious) detailed analysis, including the search for criteria (in addition to (35)) on good/optimal choices of $p_\lambda^L, q_\lambda^L, p_\lambda^U, q_\lambda^U$. Since these criteria will (almost) always imply the nonnegativity of p_λ^A, q_λ^A ($A \in \{L, U\}$) and $p_\lambda^E \geq 0, q_\lambda^E > 0$ (cf. Remark 1(a)), let us first present some fundamental properties of the underlying crucial sequences $(a_n^{(q)})_{n \in \mathbb{N}}$ and $(b_n^{(p,q)})_{n \in \mathbb{N}}$ for general $p \geq 0, q \geq 0$.

Properties 1. For all $\lambda \in \mathbb{R}$ the following holds:

(P1) If $0 < q < \beta_\lambda$, then the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ is strictly negative, strictly decreasing and converges to the unique negative solution $x_0^{(q)} \in]-\beta_\lambda, q - \beta_\lambda[$ of the equation

$$\xi_\lambda^{(q)}(x) = q \cdot e^x - \beta_\lambda = x. \tag{44}$$

(P2) If $0 < q = \beta_\lambda$, then $a_n^{(q)} \equiv 0$.

(P3) If $q > \max\{0, \beta_\lambda\}$, then the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ is strictly positive and strictly increasing. Notice that in this setup, $q = 1$ implies $\min\{1, e^{\beta_\lambda - 1}\} = e^{\beta_\lambda - 1} < q$.

(P3a) If additionally $q \leq \min\{1, e^{\beta_\lambda - 1}\}$, then the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ converges to the smallest positive solution $x_0^{(q)} \in]0, -\log q]$ of the Equation (44).

(P3b) If additionally $q > \min\{1, e^{\beta_\lambda - 1}\}$, then the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ diverges to ∞ , faster than exponentially (i.e., there do not exist constants $c_1, c_2 \in \mathbb{R}$ such that $a_n^{(q)} \leq e^{c_1 + c_2 n}$ for all $n \in \mathbb{N}$).

(P4) If $q = 0$, then one gets $a_n^{(0)} \equiv -\beta_\lambda$.

Due to the linear interrelation (38), these results directly carry over to the behaviour of the sequence $(b_n^{(p,q)})_{n \in \mathbb{N}}$:

(P5) If $p > 0$ and $0 < q < \beta_\lambda$, then the sequence $(b_n^{(p,q)})_{n \in \mathbb{N}}$ is strictly decreasing and converges to $p \cdot e^{x_0^{(q)}} - \alpha_\lambda$. Trivially, $b_1^{(p,q)} = p - \alpha_\lambda$.

(P5a) If additionally $p < \alpha_\lambda$, then $(b_n^{(p,q)})_{n \in \mathbb{N}}$ is strictly negative for all $n \in \mathbb{N}$.

(P5b) If additionally $p = \alpha_\lambda$, then $(b_n^{(p,q)})_{n \in \mathbb{N}}$ is strictly negative for all $n \in \mathbb{N} \setminus \{1\}$.

(P5c) If additionally $p > \alpha_\lambda$, then $(b_n^{(p,q)})_{n \in \mathbb{N}}$ is strictly positive for some (and possibly for all) $n \in \mathbb{N}$.

(P6) If $0 < q = \beta_\lambda$, then $b_n^{(p,q)} \equiv p - \alpha_\lambda$.

(P7) If $p > 0$ and $q > \max\{0, \beta_\lambda\}$, then the sequence $(b_n^{(p,q)})_{n \in \mathbb{N}}$ is strictly increasing.

(P7a) If additionally $q \leq \min \{1, e^{\beta_\lambda - 1}\}$, then the sequence $(b_n^{(p,q)})_{n \in \mathbb{N}}$ converges to $p \cdot e^{x_0^{(q)}} - \alpha_\lambda \in]p - \alpha_\lambda, p/q - \alpha_\lambda]$; this limit can take any sign, depending on the parameter constellation.

(P7b) If additionally $q > \min \{1, e^{\beta_\lambda - 1}\}$, then the sequence $(b_n^{(p,q)})_{n \in \mathbb{N}}$ diverges to ∞ , faster than exponentially.

(P8) For the remaining cases we get: $b_n^{(0,q)} \equiv -\alpha_\lambda$ and $b_n^{(p,0)} \equiv p \cdot e^{-\beta_\lambda} - \alpha_\lambda$ ($p \in \mathbb{R}, q \in \mathbb{R}$).

Moreover, in our investigations we will repeatedly make use of the function $\xi_\lambda^{(q)}(\cdot)$ from the definition (36) of $a_n^{(q)}$ (see also (44)), which has the following properties:

(P9) For $q \in]0, \infty[$ and all $\lambda \in \mathbb{R} \setminus \{0, 1\}$ the function $\xi_\lambda^{(q)}(\cdot)$ is strictly increasing, strictly convex and smooth, and there holds

$$(P9a) \quad \xi_\lambda^{(q)}(0) \begin{cases} < 0, & \text{if } q < \beta_\lambda, \\ = 0, & \text{if } q = \beta_\lambda, \\ > 0, & \text{if } q > \beta_\lambda. \end{cases}$$

$$(P9b) \quad \lim_{x \rightarrow -\infty} \xi_\lambda^{(q)}(x) = -\beta_\lambda, \quad \text{and} \quad \lim_{x \rightarrow \infty} \xi_\lambda^{(q)}(x) = \infty.$$

The proof of these properties is provided in Appendix A.1. From Properties 1 (P1) to (P4) we can see, that the behaviour of the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ can be classified basically into four different types;

besides the case (P2) where $a_n^{(q)}$ is constant, the sequence can be either (i) strictly decreasing and convergent (e.g., for the NI case $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) = (0.5, 2, 0, 0, 0.5)$ leading to $\beta_\lambda = \lambda\beta_{\mathcal{A}} + (1 - \lambda)\beta_{\mathcal{H}} = 1.25$ and to $q := q_\lambda^E = \beta_{\mathcal{A}}^\lambda \beta_{\mathcal{H}}^{1-\lambda} = 1$, cf. (33) resp. Theorem 1(a)), or (ii) strictly increasing and convergent (e.g., for $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) = (0.5, 2, 0, 0, 1.5)$ leading to $\beta_\lambda = -0.25, q := q_\lambda^E = 0.25$), or (iii) strictly increasing and divergent (e.g., for $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) = (0.5, 2, 0, 0, 2.7)$ leading to $\beta_\lambda = -2.05, q := q_\lambda^E \approx 0.047366$). Within our running-example epidemiological context of Section 2.3, this corresponds to a “potentially dangerous” infectious-disease-transmission situation (\mathcal{H}) (with supercritical reproduction number $\beta_{\mathcal{H}} = 2$), whereas (\mathcal{A}) describes a “mild” situation (with “low” subcritical $\beta_{\mathcal{A}} = 0.5$).

As already mentioned before, the sequences $(a_n^{(q)})_{n \in \mathbb{N}}$ and $(b_n^{(p,q)})_{n \in \mathbb{N}}$ —whose behaviours for general $p \geq 0$ and $q \geq 0$ were described by the Properties 1—have to be evaluated at setup-dependent choices $p = p_\lambda = p(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda)$ and $q = q_\lambda = q(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda)$. Hence, for fixed $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}})$, one of the questions—which arises in the course of the desired investigations of the time-behaviour of the Hellinger integral bounds (resp. exact values)—is for which $\lambda \in \mathbb{R}$ the sequence $(a_n^{(q_\lambda)})_{n \in \mathbb{N}}$ converges. In the following, we illuminate this for the important special case $q_\lambda = \beta_{\mathcal{A}}^\lambda \beta_{\mathcal{H}}^{1-\lambda}$. Suppose at first that $\beta_{\mathcal{A}} \neq \beta_{\mathcal{H}}$. Properties 1 (P1) implies that for $\lambda \in]0, 1[$ one has $\lim_{n \rightarrow \infty} a_n^{(q_\lambda)} = x_0^{(q_\lambda)} \in]-\beta_\lambda, q_\lambda - \beta_\lambda[$, and Lemma A1 states that $q_\lambda - \beta_\lambda < 0$. For $\lambda \in \mathbb{R} \setminus]0, 1[$, there holds $q_\lambda > \max\{0, \beta_\lambda\}$, and from (P3) one can see that $(a_n^{(q_\lambda)})_{n \in \mathbb{N}}$ does not converge to $x_0^{(q_\lambda)}$ in general, but for $q_\lambda \leq \min\{1, e^{\beta_\lambda - 1}\}$ which constitutes an implicit condition on λ . This can be made explicit, with the help of the auxiliary variables

$$\lambda_- := \lambda_-(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}) := \begin{cases} \inf \left\{ \lambda \leq 0 : \beta_{\mathcal{A}}^\lambda \beta_{\mathcal{H}}^{1-\lambda} \leq \min \{1, \exp\{\lambda\beta_{\mathcal{A}} + (1-\lambda)\beta_{\mathcal{H}} - 1\}\} \right\}, & \text{in case that the set is nonempty,} \\ 0, & \text{else,} \\ \sup \left\{ \lambda \geq 1 : \beta_{\mathcal{A}}^\lambda \beta_{\mathcal{H}}^{1-\lambda} \leq \min \{1, \exp\{\lambda\beta_{\mathcal{A}} + (1-\lambda)\beta_{\mathcal{H}} - 1\}\} \right\}, & \text{in case that the set is nonempty,} \\ 1, & \text{else.} \end{cases}$$

For the constellation $\beta_A = \beta_H > 0$ we clearly obtain $q_\lambda = \beta_A^\lambda \beta_H^{1-\lambda} = \beta_A = \beta_H = \beta_\lambda$. Hence, (P2) implies that the sequence $\left(a_n^{(q_\lambda)}\right)_{n \in \mathbb{N}}$ converges for all $\lambda \in \mathbb{R} \setminus \{0, 1\}$ and we can set $\lambda_- := -\infty$ as well as $\lambda_+ := \infty$. Incorporating this and by adapting a result of Linkov & Lunyova [53] on $\lambda_-(v_1, v_2), \lambda_+(v_1, v_2)$ for $\beta_A \neq \beta_H$, we end up with

Lemma 1. (a) For all $\beta_A > 0, \beta_H > 0$ with $\beta_A \neq \beta_H$ there holds

$$\lambda_- = \lambda_-(\beta_A, \beta_H) = \begin{cases} 0, & \text{if } \beta_H \geq 1, \\ \check{\lambda}, & \text{if } \beta_H < 1 \text{ and } \beta_A \notin [\beta_H, \beta_H z(\beta_H)], \\ -\infty, & \text{if } \beta_H < 1 \text{ and } \beta_A \in]\beta_H, \beta_H z(\beta_H)], \end{cases}$$

$$\lambda_+ = \lambda_+(\beta_A, \beta_H) = \begin{cases} 1, & \text{if } \beta_A \geq 1, \\ \check{\lambda}, & \text{if } \beta_A < 1 \text{ and } \beta_H \notin [\beta_A, \beta_A z(\beta_A)], \\ \infty, & \text{if } \beta_A < 1 \text{ and } \beta_H \in]\beta_A, \beta_A z(\beta_A)], \end{cases}$$

where

$$\check{\lambda} := \check{\lambda}(\beta_A, \beta_H) := \frac{\beta_H - 1 - \log(\beta_H)}{\beta_H - \beta_A + \log\left(\frac{\beta_A}{\beta_H}\right)} \begin{cases} < 0, & \text{if } \beta_H < 1 \text{ and } \beta_A \notin [\beta_H, \beta_H z(\beta_H)], \\ > 1, & \text{if } \beta_A < 1 \text{ and } \beta_H \notin [\beta_A, \beta_A z(\beta_A)]. \end{cases}$$

Here, for fixed $\beta \in]0, \infty[\setminus \{1\}$ we denote by $z(\beta)$ the unique solution of the equation $\log(x) - \beta(x - 1) = 0, x \in]0, \infty[\setminus \{1\}$. For $\beta = 1, z(\beta) = 1$ denotes the unique solution of $\log(x) - (x - 1) = 0, x \in]0, \infty[$.

(b) For all $\beta_A = \beta_H > 0$ one gets $\lambda_- = \lambda_-(\beta_A, \beta_H) = -\infty$ as well as $\lambda_+ = \lambda_+(\beta_A, \beta_H) = \infty$.

Notice that the relationship $\check{\lambda}(\beta_A, \beta_H) = 1 - \check{\lambda}(\beta_H, \beta_A)$ is consistent with the skew symmetry (8).

A corresponding proof is given in Appendix A.1.

With these auxiliary basic facts in hand, let us now work out our detailed investigations of the time-behaviour $n \mapsto H_\lambda(P_{A,n} || P_{H,n})$, where we start with the exactly treatable case (a) in Theorem 1.

3.3. Detailed Analyses of the Exact Recursive Values, i.e., for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$

In the no-immigration-case $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{NI}$ and in the equal-fraction-case $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP,1}$, the Hellinger integral can be calculated exactly in terms of $H_\lambda(P_{A,n} || P_{H,n}) = V_{\lambda, X_0, n}$ (cf. (39)), as proposed in part (a) of Theorem 1. This quantity depends on the behaviour of the sequence $\left(a_n^{(q_\lambda^E)}\right)_{n \in \mathbb{N}}$, with $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda} > 0$, and of the sum $\left(\frac{\alpha_A}{\beta_A} \sum_{k=1}^n a_k^{(q_\lambda^E)}\right)_{n \in \mathbb{N}}$. The last expression is equal to zero on \mathcal{P}_{NI} . On $\mathcal{P}_{SP,1}$, this sum is unequal to zero. Using Lemma A1 we conclude that $q_\lambda^E < \beta_\lambda$ (resp. $q_\lambda^E > \beta_\lambda$) iff $\lambda \in]0, 1[$ (resp. $\lambda \in \mathbb{R} \setminus [0, 1]$), since on $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$ there holds $\beta_A \neq \beta_H$. Thus, from Properties 1 (P1) we can see that the sequence $\left(a_n^{(q_\lambda^E)}\right)_{n \in \mathbb{N}}$ is strictly negative, strictly decreasing and it converges to the unique solution $x_0^{(q_\lambda^E)} \in]-\beta_\lambda, q_\lambda^E - \beta_\lambda[$ of the Equation (44) if $\lambda \in]0, 1[$. For $\lambda \in \mathbb{R} \setminus [0, 1]$, (P3) implies that the sequence $\left(a_n^{(q_\lambda^E)}\right)_{n \in \mathbb{N}}$ is strictly positive, strictly increasing and converges to the smallest positive solution $x_0^{(q_\lambda^E)} \in]0, -\log(q_\lambda^E)]$ of the Equation (44) in case that (P3a) is satisfied, otherwise it diverges to ∞ . Thus, we have shown the following detailed behaviour of Hellinger integrals:

Proposition 2. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{NI} \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds

- (a) $H_\lambda(P_{A,1}||P_{H,1}) = \exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \lambda \beta_A - (1-\lambda) \beta_H \right) X_0 \right\} < 1,$
- (b) the sequence $(H_\lambda(P_{A,n}||P_{H,n}))_{n \in \mathbb{N}}$ given by

$$H_\lambda(P_{A,n}||P_{H,n}) = \exp \left\{ a_n^{(q_\lambda^E)} X_0 \right\} =: V_{\lambda, X_0, n}$$
 is strictly decreasing,
- (c) $\lim_{n \rightarrow \infty} H_\lambda(P_{A,n}||P_{H,n}) = \exp \left\{ x_0^{(q_\lambda^E)} X_0 \right\} \in]0, 1[,$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log H_\lambda(P_{A,n}||P_{H,n}) = 0$
- (e) the map $X_0 \mapsto V_{\lambda, X_0, n}$ is strictly decreasing.

Proposition 3. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{NI} \times (\mathbb{R} \setminus]0, 1])$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $H_\lambda(P_{A,1}||P_{H,1}) = \exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda \right) \cdot X_0 \right\} > 1,$
- (b) the sequence $(H_\lambda(P_{A,n}||P_{H,n}))_{n \in \mathbb{N}}$ given by

$$H_\lambda(P_{A,n}||P_{H,n}) = \exp \left\{ a_n^{(q_\lambda^E)} \cdot X_0 \right\} =: V_{\lambda, X_0, n}$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} H_\lambda(P_{A,n}||P_{H,n}) = \begin{cases} \exp \left\{ x_0^{(q_\lambda^E)} \cdot X_0 \right\} > 1, & \text{if } \lambda \in [\lambda_-, \lambda_+] \setminus]0, 1], \\ \infty, & \text{if } \lambda \in]-\infty, \lambda_-[\cup]\lambda_+, \infty[, \end{cases}$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log H_\lambda(P_{A,n}||P_{H,n}) = \begin{cases} 0, & \text{if } \lambda \in [\lambda_-, \lambda_+] \setminus]0, 1], \\ \infty, & \text{if } \lambda \in]-\infty, \lambda_-[\cup]\lambda_+, \infty[, \end{cases}$
- (e) the map $X_0 \mapsto V_{\lambda, X_0, n}$ is strictly increasing.

In the case $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP,1}$, the sequence $\left(a_n^{(q_\lambda^E)} \right)_{n \in \mathbb{N}}$ under consideration is formally the same, with the parameter $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda} > 0$. However, in contrast to the case \mathcal{P}_{NI} , on $\mathcal{P}_{SP,1}$ both the sequence $\left(a_n^{(q_\lambda^E)} \right)_{n \in \mathbb{N}}$ and the sum $\left(\frac{\alpha_A}{\beta_A} \sum_{k=1}^n a_k^{(q_\lambda^E)} \right)_{n \in \mathbb{N}}$ are strictly decreasing in case that $\lambda \in]0, 1[$, and strictly increasing in case that $\lambda \in \mathbb{R} \setminus]0, 1]$. The respective convergence behaviours are given in Properties 1 (P1) and (P3). We thus obtain

Proposition 4. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,1} \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $H_\lambda(P_{A,1}||P_{H,1}) = \exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda \right) \cdot \left(X_0 + \frac{\alpha_A}{\beta_A} \right) \right\} < 1,$
- (b) the sequence $(H_\lambda(P_{A,n}||P_{H,n}))_{n \in \mathbb{N}}$ given by

$$H_\lambda(P_{A,n}||P_{H,n}) = \exp \left\{ a_n^{(q_\lambda^E)} \cdot X_0 + \frac{\alpha_A}{\beta_A} \sum_{k=1}^n a_k^{(q_\lambda^E)} \right\} =: V_{\lambda, X_0, n}$$
 is strictly decreasing,
- (c) $\lim_{n \rightarrow \infty} H_\lambda(P_{A,n}||P_{H,n}) = 0,$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log H_\lambda(P_{A,n}||P_{H,n}) = \frac{\alpha_A}{\beta_A} \cdot x_0^{(q_\lambda^E)} < 0,$
- (e) the map $X_0 \mapsto V_{\lambda, X_0, n}$ is strictly decreasing.

Proposition 5. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,1} \times (\mathbb{R} \setminus [0, 1])$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $H_\lambda(P_{A,1} || P_{H,1}) = \exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda \right) \cdot \left(X_0 + \frac{\alpha_A}{\beta_A} \right) \right\} > 1,$
- (b) the sequence $(H_\lambda(P_{A,n} || P_{H,n}))_{n \in \mathbb{N}}$ given by

$$H_\lambda(P_{A,n} || P_{H,n}) = \exp \left\{ a_n^{(q_\lambda^E)} \cdot X_0 + \frac{\alpha_A}{\beta_A} \sum_{k=1}^n a_k^{(q_\lambda^E)} \right\} =: V_{\lambda, X_0, n}$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}) = \infty,$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log H_\lambda(P_{A,n} || P_{H,n}) = \begin{cases} \frac{\alpha_A}{\beta_A} \cdot x_0^{(q_\lambda^E)} > 0, & \text{if } \lambda \in [\lambda_-, \lambda_+] \setminus [0, 1], \\ \infty, & \text{if } \lambda \in] - \infty, \lambda_- [\cup] \lambda_+, \infty [, \end{cases}$
- (e) the map $X_0 \mapsto V_{\lambda, X_0, n}$ is strictly increasing.

Due to the nature of the equal-fraction-case $\mathcal{P}_{SP,1}$, in the assertions (a), (b), (d) of the Propositions 4 and 5, the fraction α_A / β_A can be equivalently replaced by α_H / β_H .

Remark 2. For the (to our context) incompatible setup of GWI with Poisson offspring but nonstochastic immigration of constant value 1, an “analogue” of part (d) of the Propositions 4 resp. 5 was established in Linkov & Lunyova [53].

3.4. Some Preparatory Basic Facts for the Remaining Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$

The bounds $B_{\lambda, X_0, n}^L, B_{\lambda, X_0, n}^U$ for the Hellinger integral introduced in formula (40) in Theorem 1 can be chosen arbitrarily from a $(p_\lambda^L, q_\lambda^L, p_\lambda^U, q_\lambda^U)$ -indexed set of context-specific parameters satisfying (34), or equivalently (35).

In order to derive bounds which are optimal, with respect to goals that will be discussed later, the following monotonicity properties of the sequences $(a_n^{(q)})_{n \in \mathbb{N}}$ and $(b_n^{(p,q)})_{n \in \mathbb{N}}$ (cf. (36), (37)) for general, context-independent parameters q and p , will turn out to be very useful:

Properties 2.

- (P10) For $0 \leq q_1 < q_2 < \infty$ there holds $a_n^{(q_1)} < a_n^{(q_2)}$ for all $n \in \mathbb{N}$.
- (P11) For each fixed $q \geq 0$ and $0 \leq p_1 < p_2 < \infty$ there holds $b_n^{(p_1,q)} < b_n^{(p_2,q)}$, for all $n \in \mathbb{N}$.
- (P12) For fixed $p > 0$ and $0 \leq q_1 < q_2$ it follows $b_n^{(p,q_1)} < b_n^{(p,q_2)}$ for all $n \in \mathbb{N}$.
- (P13) Suppose that $0 \leq p_1 < p_2$ and $0 \leq q_2 < q_1$. For fixed $n \in \mathbb{N}$, no dominance assertion can be conjectured for $b_n^{(p_1,q_1)}, b_n^{(p_2,q_2)}$. As an example, consider the setup $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (0.4, 0.8, 5, 3, 0.5)$; within our running-example epidemiological context of Section 2.3, this corresponds to a “nearly dangerous” infectious-disease-transmission situation (\mathcal{H}) (with nearly critical reproduction number $\beta_H = 0.8$ and importation mean of $\alpha_H = 3$), whereas (\mathcal{A}) describes a “mild” situation (with “low” subcritical $\beta_A = 0.4$ and $\alpha_A = 5$). On the nonnegative real line, the function $\phi_\lambda(x)$ can be bounded from above by the linear functions $\phi_\lambda^{U,1}(x) := p_1 + q_1 x := 4.040 + 0.593 \cdot x$ as well as by $\phi_\lambda^{U,2}(x) := p_2 + q_2 x := 4.110 + 0.584 \cdot x$. Clearly, $p_1 < p_2$ and $q_1 > q_2$. Let us show the first eight elements and the respective limits of the corresponding sequences $b_n^{(p_1,q_1)}, b_n^{(p_2,q_2)}$:

n	1	2	3	4	5	6	7	8	...	∞
$b_n^{(p_1,q_1)}$	0.040	0.011	−0.005	−0.015	−0.021	−0.024	−0.026	−0.028	...	−0.029
$b_n^{(p_2,q_2)}$	0.110	0.045	0.007	−0.014	−0.026	−0.033	−0.036	−0.039	...	−0.041

(P14) For arbitrary $0 < p_1, p_2$ and $0 \leq q_1, q_2 \leq \min\{1, e^{\beta_\lambda - 1}\}$ suppose that $\log(p_1) + x_0^{(q_1)} < \log(p_2) + x_0^{(q_2)}$. Then there holds

$$p_1 \cdot e^{x_0^{(q_1)}} - \alpha_\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n b_k^{(p_1, q_1)} < \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n b_k^{(p_2, q_2)} = p_2 \cdot e^{x_0^{(q_2)}} - \alpha_\lambda.$$

From (P10) to (P12) one deduces that both sequences $(a_n^{(q)})_{n \in \mathbb{N}}$ and $(b_n^{(p, q)})_{n \in \mathbb{N}}$ are monotone in the general parameters $p, q \geq 0$. Thus, for the upper bound of the Hellinger integral $B_{\lambda, X_0, n}^U$ we should use nonnegative context-specific parameters $p_\lambda^U = p^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ and $q_\lambda^U = q^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ which are as small as possible, and for the lower bound $B_{\lambda, X_0, n}^L$ we should use nonnegative context-specific parameters $p_\lambda^L = p^L(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ and $q_\lambda^L = q^L(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ which are as large as possible, of course, subject to the (equivalent) restrictions (34) and (35).

To find “optimal” parameter pairs, we have to study the following properties of the function $\phi_\lambda(\cdot) = \phi(\cdot, \beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ defined on $[0, \infty[$ in (30) (which are also valid for the previous parameter context $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1})$):

Properties 3.

(P15) One has

$$\phi_\lambda(x) = (\alpha_A + \beta_A x)^\lambda (\alpha_H + \beta_H x)^{1-\lambda} - \lambda(\alpha_A + \beta_A x) + (1-\lambda)(\alpha_H + \beta_H x) \begin{cases} \leq 0, & \text{if } \lambda \in]0, 1[, \\ \geq 0, & \text{if } \lambda \in \mathbb{R} \setminus]0, 1[, \end{cases}$$

where equality holds iff $f_A(x) = f_H(x)$ for some $x \in [0, \infty[$ iff $x = x^* := \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} \in [0, \infty[$.

(P16) There holds

$$\phi_\lambda(0) = \alpha_A^\lambda \alpha_H^{1-\lambda} - \alpha_\lambda \begin{cases} \leq 0, & \text{if } \lambda \in]0, 1[, \\ \geq 0, & \text{if } \lambda \in \mathbb{R} \setminus]0, 1[, \end{cases}$$

with equality iff $\alpha_A = \alpha_H$ together with $\beta_A \neq \beta_H$ (cf. Lemma A1).

(P17) For all $\lambda \in \mathbb{R} \setminus \{0, 1\}$ one gets

$$\phi'_\lambda(x) = \lambda \beta_A (f_A(x))^{\lambda-1} (f_H(x))^{1-\lambda} + (1-\lambda) \beta_H (f_A(x))^\lambda (f_H(x))^{-\lambda} - \beta_\lambda.$$

(P18) There holds

$$\lim_{x \rightarrow \infty} \phi'_\lambda(x) = \beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda \begin{cases} \leq 0, & \text{if } \lambda \in]0, 1[, \\ \geq 0, & \text{if } \lambda \in \mathbb{R} \setminus]0, 1[, \end{cases}$$

with equality iff $\beta_A = \beta_H$ together with $\alpha_A \neq \alpha_H$ (cf. Lemma A1).

(P19) There holds

$$\phi''_\lambda(x) = -\lambda(1-\lambda) (f_A(x))^{\lambda-2} (f_H(x))^{-\lambda-1} (\alpha_A \beta_H - \alpha_H \beta_A)^2 \begin{cases} \leq 0, & \text{if } \lambda \in]0, 1[, \\ \geq 0, & \text{if } \lambda \in \mathbb{R} \setminus]0, 1[, \end{cases}$$

with equality iff $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1})$. Hence, for $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$, the function ϕ_λ is strictly concave (convex) for $\lambda \in]0, 1[$ ($\lambda \in \mathbb{R} \setminus]0, 1[$). Notice that $\phi'_\lambda(0) = \lambda \beta_A \left(\frac{\alpha_A}{\alpha_H}\right)^{\lambda-1} + (1-\lambda) \beta_H \left(\frac{\alpha_A}{\alpha_H}\right)^\lambda - \beta_\lambda$ can be either negative (e.g., for the setup $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \{(4, 2, 3, 1, 0.5), (4, 2, 5, 1, 2)\}$), or zero (e.g., for $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \{(4, 2, 4, 1, 0.5), (4, 2, 3, 1, 2)\}$), or positive (e.g., for $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \{(4, 2, 5, 1, 0.5), (4, 2, 2, 1, 2)\}$), where the exemplary parameter

constellations have concrete interpretations in our running-example epidemiological context of Section 2.3. Accordingly, for $\lambda \in]0, 1[$, due to concavity and (P17), the function $\phi_\lambda(\cdot)$ can be either strictly decreasing, or can obtain its global maximum in $]0, \infty[$, or—only in the case $\beta_A = \beta_H$ —can be strictly increasing. Analogously, for $\lambda \in \mathbb{R} \setminus]0, 1[$, the function $\phi_\lambda(\cdot)$ can be either strictly increasing, or can obtain its global minimum in $]0, \infty[$, or—only in the case $\beta_A = \beta_H$ —can be strictly decreasing.

(P20) For all $\lambda \in \mathbb{R} \setminus \{0, 1\}$ one has

$$\begin{aligned} \lim_{x \rightarrow \infty} (\phi_\lambda(x) - (\tilde{r}_\lambda + \tilde{s}_\lambda x)) &= 0, \\ \text{for } \tilde{r}_\lambda &:= \tilde{p}_\lambda - \alpha_\lambda := \lambda \alpha_A \left(\frac{\beta_A}{\beta_H}\right)^{\lambda-1} + (1-\lambda) \alpha_H \left(\frac{\beta_A}{\beta_H}\right)^\lambda - \alpha_\lambda \\ \text{and } \tilde{s}_\lambda &:= \tilde{q}_\lambda - \beta_\lambda := \beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda. \end{aligned}$$

The linear function $\tilde{\phi}_\lambda(x) := \tilde{r}_\lambda + \tilde{s}_\lambda \cdot x$ constitutes the asymptote of $\phi_\lambda(\cdot)$. Notice that if $\beta_A = \beta_H$ one has $\tilde{s}_\lambda = 0 = \tilde{r}_\lambda$; if $\beta_A \neq \beta_H$ we have $\tilde{s}_\lambda < 0$ in the case $\lambda \in]0, 1[$ and $\tilde{s}_\lambda > 0$ if $\lambda \in \mathbb{R} \setminus]0, 1[$. Furthermore, $\phi_\lambda(0) < \tilde{r}_\lambda$ if $\lambda \in]0, 1[$ and $\phi_\lambda(0) > \tilde{r}_\lambda$ if $\lambda \in \mathbb{R} \setminus]0, 1[$, (cf. Lemma A1(c1) and (c2)). If $\alpha_A = \alpha_H$ (and thus $\beta_A \neq \beta_H$), then the intercept \tilde{r}_λ is strictly positive if $\lambda \in]0, 1[$ resp. strictly negative if $\lambda \in \mathbb{R} \setminus]0, 1[$. In contrast, for the case $\alpha_A \neq \alpha_H$, the intercept \tilde{r}_λ can assume any sign, take e.g., $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \{(3.7, 0.9, 2.0, 1.0, 0.5), (4, 2, 1.6, 1, 2)\}$ for $\tilde{r}_\lambda > 0$, $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \{(3.6, 0.9, 2.0, 1.0, 0.5), (4, 2, 1.5, 1, 2)\}$ for $\tilde{r}_\lambda = 0$, and $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \{(3.5, 0.9, 2.0, 1.0, 0.5), (4, 2, 1.4, 1, 2)\}$ for $\tilde{r}_\lambda < 0$; again, the exemplary parameter constellations have concrete interpretations in our running-example epidemiological context of Section 2.3.

The properties (P15) to (P20) above describe in detail the characteristics of the function $\phi_\lambda(\cdot) = \phi(\cdot, \beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$. In the previous parameter setup $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$, this function is linear, which can be seen from (P19). In the current parameter setup $\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$, this function can basically be classified into four different types. From (P16) to (P20) it is easy to see that for all current parameter constellations the particular choices

$$p_\lambda^A := \alpha_A^\lambda \alpha_H^{1-\lambda} > 0, \quad q_\lambda^A := \beta_A^\lambda \beta_H^{1-\lambda} > 0, \tag{45}$$

which correspond to the following choices in (35)

$$r_\lambda^A := \alpha_A^\lambda \alpha_H^{1-\lambda} - \alpha_\lambda \leq 0 \quad (\text{resp. } \geq 0), \quad s_\lambda^A := \beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda \leq 0 \quad (\text{resp. } \geq 0),$$

– where $A = L$ (resp. $A = U$)—lead to the tightest lower bound $B_{\lambda, X_0, n}^L$ (resp. upper bound $B_{\lambda, X_0, n}^U$) for $H_\lambda(P_{A,n} || P_{H,n})$ in (40) in the case $\lambda \in]0, 1[$ (resp. $\lambda \in \mathbb{R} \setminus]0, 1[$). Notice that for the previous parameter setup $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1})$ these choices led to the exact values of the Hellinger integral and to the simplification $(p_\lambda^E / q_\lambda^E) \cdot \beta_\lambda - \alpha_\lambda = 0$, which implies $b_n^{(p_\lambda^E, q_\lambda^E)} = (\alpha_A / \beta_A) \cdot a_n^{(q_\lambda^E)}$. In contrast, in the current parameter setup $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$ we only derive the optimal lower (resp. upper) bound for $\lambda \in]0, 1[$ (resp. $\lambda \in \mathbb{R} \setminus]0, 1[$) by using the parameters p_λ^A, q_λ^A for $A = L$ (resp. $A = U$) and $(p_\lambda^A / q_\lambda^A) \cdot \beta_\lambda - \alpha_\lambda \neq 0$. For a better distinguishability and easier reference we thus stick to the L -notation (resp. U -notation) here.

3.5. Lower Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$

The discussion above implies that the lower bound $B_{\lambda, X_0, n}^L$ for the Hellinger integral $H_\lambda(P_{A,n} || P_{H,n})$ in (40) is optimal for the choices $p_\lambda^L, q_\lambda^L > 0$ defined in (45). If $\beta_A \neq \beta_H$, due to Properties 1 (P1) and Lemma A1, the sequence $\left(a_n^{(q_\lambda^L)}\right)_{n \in \mathbb{N}}$ is strictly negative and strictly decreasing and converges to the unique negative solution of the Equation (44). Furthermore, due to (P5),

the sequence $\left(b_n^{(p_\lambda^L, q_\lambda^L)}\right)_{n \in \mathbb{N}}$, as defined in (37), is strictly decreasing. Since $b_1^{(p_\lambda^L, q_\lambda^L)} = p_\lambda^L - \alpha_\lambda \leq 0$ by Lemma A1, with equality iff $\alpha_A = \alpha_H$, the sequence $\left(b_n^{(p_\lambda^L, q_\lambda^L)}\right)_{n \in \mathbb{N}}$ is also strictly negative (with the exception $b_1^{(p_\lambda^L, q_\lambda^L)} = 0$ for $\alpha_A = \alpha_H$) and strictly decreasing. If $\beta_A = \beta_H$ and thus $\alpha_A \neq \alpha_H$, due to (P2), (P6) and Lemma A1, there holds $a_n^{(q_\lambda^L)} \equiv 0$ and $b_n^{(q_\lambda^L)} \equiv p_\lambda^L - \alpha_\lambda < 0$. Thus, analogously to the cases $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$ we obtain

Proposition 6. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $p_\lambda^L := \alpha_A^\lambda \alpha_H^{1-\lambda}$, $q_\lambda^L := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $B_{\lambda, X_0, 1}^L = \exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda \right) \cdot X_0 + \alpha_A^\lambda \alpha_H^{1-\lambda} - \alpha_\lambda \right\} < 1$,
- (b) the sequence of lower bounds $\left(B_{\lambda, X_0, n}^L\right)_{n \in \mathbb{N}}$ for $H_\lambda(P_{A,n} || P_{H,n})$ given by
$$B_{\lambda, X_0, n}^L = \exp \left\{ a_n^{(q_\lambda^L)} \cdot X_0 + \frac{p_\lambda^L}{q_\lambda^L} \sum_{k=1}^n a_k^{(q_\lambda^L)} + n \cdot \left(\frac{p_\lambda^L}{q_\lambda^L} \cdot \beta_\lambda - \alpha_\lambda \right) \right\}$$
 is strictly decreasing,
- (c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L = 0$,
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log B_{\lambda, X_0, n}^L = \frac{p_\lambda^L}{q_\lambda^L} \cdot \left(x_0^{(q_\lambda^L)} + \beta_\lambda \right) - \alpha_\lambda = p_\lambda^L \cdot e^{x_0^{(q_\lambda^L)}} - \alpha_\lambda < 0$.
- (e) the map $X_0 \mapsto B_{\lambda, X_0, n}^L$ is strictly decreasing.

3.6. Goals for Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$

For parameter constellations $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$, in contrast to the treatment of the lower bounds (cf. the previous Section 3.5), the fine-tuning of the upper bounds of the Hellinger integrals $H_\lambda(P_{A,n} || P_{H,n})$ is much more involved. To begin with, let us mention that the monotonicity-concerning Properties 2 (P10) to (P12) imply that for a tight upper bound $B_{\lambda, X_0, n}^U$ (cf. (40)) one should choose parameters $p_\lambda^U \geq p_\lambda^L > 0$, $q_\lambda^U \geq q_\lambda^L > 0$ as small as possible. Due to the concavity (cf. Properties 3 (P19)) of the function $\phi_\lambda(\cdot)$, the linear upper bound $\phi_\lambda^U(\cdot)$ (on the ultimately relevant subdomain \mathbb{N}_0) thus must hit the function $\phi_\lambda(\cdot)$ in at least one point $x \in \mathbb{N}_0$, which corresponds to some “discrete tangent line” of $\phi_\lambda(\cdot)$ in x , or in at most two points $x, x + 1 \in \mathbb{N}_0$, which corresponds to the secant line of $\phi_\lambda(\cdot)$ across its arguments x and $x + 1$. Accordingly, there is in general no overall best upper bound; of course, one way to obtain “good” upper bounds for $H_\lambda(P_{A,n} || P_{H,n})$ is to solve the optimization problem

$$\left(\overline{p_\lambda^U}, \overline{q_\lambda^U}\right) := \arg \min_{(p_\lambda^U, q_\lambda^U)} \left\{ \exp \left\{ a_n^{(q_\lambda^U)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)} \right\} \right\}, \tag{46}$$

subject to the constraint (35). However, the corresponding result generally depends on the particular choice of the initial population $X_0 \in \mathbb{N}$ and on the observation time horizon $n \in \mathbb{N}$. Hence, there is in general no overall optimal choice of p_λ^U, q_λ^U without the incorporation of further goal-dependent constraints such as $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^U = 0$ in case of $\lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}) = 0$. By the way, mainly because of the non-explicitness of the sequence $\left(a_n^{(q_\lambda^U)}\right)_{n \in \mathbb{N}}$ (due to the generally not explicitly solvable recursion (36)) and the discreteness of the constraint (35), this optimization problem seems to be not straightforward to solve, anyway. The choice of parameters p_λ^U, q_λ^U for the upper bound $B_{\lambda, X_0, n}^U \geq H_\lambda(P_{A,n} || P_{H,n})$ can be made according to different, partially incompatible (“optimality-” resp. “goodness-”) criteria and goals, such as:

- (G1) the validity of $B_{\lambda, X_0, n}^U < 1$ *simultaneously* for all initial configurations $X_0 \in \mathbb{N}$, all observation horizons $n \in \mathbb{N}$ and all $\lambda \in]0, 1[$, which leads to a *strict* improvement of the general upper bound $H_\lambda(P_{\mathcal{A}, n} || P_{\mathcal{H}, n}) < 1$ (cf. (9));
- (G2) the determination of the long-term-limits $\lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A}, n} || P_{\mathcal{H}, n})$ respectively $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^U$ for all $X_0 \in \mathbb{N}$ and all $\lambda \in]0, 1[$; in particular, one would like to check whether $\lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A}, n} || P_{\mathcal{H}, n}) = 0$, which implies that the families of probability distributions $(P_{\mathcal{A}, n})_{n \in \mathbb{N}}$ and $(P_{\mathcal{H}, n})_{n \in \mathbb{N}}$ are *asymptotically distinguishable* (entirely separated), cf. (25);
- (G3) the determination of the time-asymptotical growth rates $\lim_{n \rightarrow \infty} \frac{1}{n} \log (H_\lambda(P_{\mathcal{A}, n} || P_{\mathcal{H}, n}))$ resp. $\lim_{n \rightarrow \infty} \frac{1}{n} \log (B_{\lambda, X_0, n}^U)$ for all $X_0 \in \mathbb{N}$ and all $\lambda \in]0, 1[$.

Further goals—with which we do not deal here for the sake of brevity—are for instance (i) a very good tightness of the upper bound $B_{\lambda, X_0, n}^U$ for $n \geq N$ for some fixed large $N \in \mathbb{N}$, or (ii) the criterion (G1) with *fixed* (rather than arbitrary) initial population size $X_0 \in \mathbb{N}$.

Let us briefly discuss the three Goals (G1) to (G3) and their challenges: due to Theorem 1, Goal (G1) can only be achieved if the sequence $(a_n^{(q_\lambda^U)})_{n \in \mathbb{N}}$ is non-increasing, since otherwise, for each fixed observation horizon $n \in \mathbb{N}$ there is a large enough initial population size X_0 such that the upper bound component $\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)}$ becomes larger than 1, and thus $B_{\lambda, X_0, n}^U = 1$ (cf. (40)). Hence, Properties 1 (P1) and (P2) imply that one should have $q_\lambda^U \leq \beta_\lambda$. Then, the sequence $(b_n^{(p_\lambda^U, q_\lambda^U)})_{n \in \mathbb{N}}$ is also non-increasing. However, since $b_n^{(p_\lambda^U, q_\lambda^U)}$ might be positive for some (even all) $n \in \mathbb{N}$, the sum $(\sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)})_{n \in \mathbb{N}}$ is not necessarily decreasing. Nevertheless, the restriction

$$q_\lambda^U - \beta_\lambda \leq 0 \quad \text{and} \quad p_\lambda^U - \alpha_\lambda \leq 0, \quad \text{where at least one of the inequalities is strict,} \quad (47)$$

ensures that both sequences $(a_n^{(q_\lambda^U)})_{n \in \mathbb{N}}$ and $(b_n^{(p_\lambda^U, q_\lambda^U)})_{n \in \mathbb{N}}$ are nonpositive and decreasing, where at least one sequence is strictly negative, implying that the sum $(\sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)})_{n \in \mathbb{N}}$ is strictly negative for $n \geq 2$ and strictly decreasing. To see this, suppose that (47) is satisfied with two strict inequalities. Then, $(a_n^{(q_\lambda^U)})_{n \in \mathbb{N}}$ as well as $(b_n^{(p_\lambda^U, q_\lambda^U)})_{n \in \mathbb{N}}$ are strictly negative and strictly decreasing. If $q_\lambda^U = \beta_\lambda$ and $p_\lambda^U < \alpha_\lambda$, we see from (P2) and (P6) that $a_n^{(q_\lambda^U)} \equiv 0$ and that $b_n^{(p_\lambda^U, q_\lambda^U)} \equiv p_\lambda^U - \alpha_\lambda < 0$ (notice that $\alpha_\lambda = 0$ is not possible in the current setup $\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP}1}$ and for $\lambda \in]0, 1[$). In the last case $q_\lambda^U < \beta_\lambda$ and $p_\lambda^U = \alpha_\lambda$, from (P1) and (P5) it follows that $(a_n^{(q_\lambda^U)})_{n \in \mathbb{N}}$ is strictly negative and strictly decreasing, as well as that $b_1^{(p_\lambda^U, q_\lambda^U)} = 0$ and $(b_n^{(p_\lambda^U, q_\lambda^U)})_{n \in \mathbb{N}}$ is strictly decreasing and strictly negative for $n \geq 2$. Thus, whenever (47) is satisfied, the sum $(\sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)})_{n \in \mathbb{N}}$ is strictly negative for $n \geq 2$ and strictly decreasing.

To achieve Goal (G2), we have to require that the sequence $(a_n^{(q_\lambda^U)})_{n \in \mathbb{N}}$ converges, which is the case if either $q_\lambda^U \leq \beta_\lambda$ or $\beta_\lambda < q_\lambda^U \leq \min\{1, e^{\beta_\lambda - 1}\}$ (cf. Properties 1 (P1) to (P3)). From the upper bound component $\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)}$ (42) we conclude that Goal (G2) is met if the sequence $(b_n^{(p_\lambda^U, q_\lambda^U)})_{n \in \mathbb{N}}$ converges to a negative limit, i.e., $\lim_{n \rightarrow \infty} b_n^{(p_\lambda^U, q_\lambda^U)} = p_\lambda^U \cdot e^{x_0^{(q_\lambda^U)}} - \alpha_\lambda < 0$. Notice that this condition holds true if (47) is satisfied: suppose that $q_\lambda^U < \beta_\lambda$, then $x_0^{(q_\lambda^U)} < 0$ and $p_\lambda^U \cdot e^{x_0^{(q_\lambda^U)}} - \alpha_\lambda < p_\lambda^U - \alpha_\lambda \leq 0$. On the other hand, if $p_\lambda^U - \alpha_\lambda < 0$, one obtains $x_0^{(q_\lambda^U)} \leq 0$ leading to $p_\lambda^U \cdot e^{x_0^{(q_\lambda^U)}} - \alpha_\lambda \leq p_\lambda^U - \alpha_\lambda < 0$.

The examination of Goal (G2) above enters into the discussion of Goal (G3): if the sequence $\left(a_n^{(q_\lambda^U)}\right)_{n \in \mathbb{N}}$ converges and $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^U = 0$, then there holds

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(B_{\lambda, X_0, n}^U \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\widehat{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} \right) = p_\lambda^U \cdot e^{x_0^{(q_\lambda^U)}} - \alpha_\lambda. \tag{48}$$

For the case $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$, let us now start with our comprehensive investigations of the upper bounds, where we focus on fulfilling the condition (47) which tackles Goals (G1) and (G2) simultaneously; then, the Goal (G3) can be achieved by (48). As indicated above, various different parameter subcases can lead to different Hellinger-integral-upper-bound details, which we work out in the following. For better transparency, we employ the following notations (where the first four are just reminders of sets which were already introduced above)

$$\begin{aligned} \mathcal{P}_{NI} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in [0, \infty[^4 : \alpha_A = \alpha_H = 0; \beta_A > 0; \beta_H > 0; \beta_A \neq \beta_H \right\}, \\ \mathcal{P}_{SP} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in]0, \infty[^4 : (\alpha_A \neq \alpha_H) \text{ or } (\beta_A \neq \beta_H) \text{ or both} \right\}, \\ \mathcal{P} &:= \mathcal{P}_{NI} \cup \mathcal{P}_{SP}, \\ \mathcal{P}_{SP,1} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H, \beta_A \neq \beta_H, \frac{\alpha_A}{\beta_A} = \frac{\alpha_H}{\beta_H} \right\}, \\ \mathcal{P}_{SP,2} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A = \alpha_H, \beta_A \neq \beta_H \right\}, \\ \mathcal{P}_{SP,3} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H, \beta_A \neq \beta_H, \frac{\alpha_A}{\beta_A} \neq \frac{\alpha_H}{\beta_H} \right\} = \mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b} \cup \mathcal{P}_{SP,3c}, \\ \mathcal{P}_{SP,3a} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H, \beta_A \neq \beta_H, \frac{\alpha_A}{\beta_A} \neq \frac{\alpha_H}{\beta_H}, \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} \in]-\infty, 0[\right\}, \\ \mathcal{P}_{SP,3b} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H, \beta_A \neq \beta_H, \frac{\alpha_A}{\beta_A} \neq \frac{\alpha_H}{\beta_H}, \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} \in]0, \infty[\setminus \mathbb{N} \right\}, \\ \mathcal{P}_{SP,3c} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H, \beta_A \neq \beta_H, \frac{\alpha_A}{\beta_A} \neq \frac{\alpha_H}{\beta_H}, \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} \in \mathbb{N} \right\}, \\ \mathcal{P}_{SP,4} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H > 0, \beta_A = \beta_H \right\} = \mathcal{P}_{SP,4a} \cup \mathcal{P}_{SP,4b}, \\ \mathcal{P}_{SP,4a} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H > 0, \beta_A = \beta_H \in]0, 1[\right\}, \\ \mathcal{P}_{SP,4b} &:= \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H > 0, \beta_A = \beta_H \in [1, \infty[\right\}; \end{aligned} \tag{49}$$

notice that because of Lemma A1 and of the Properties 3 (P15) one gets on the domain $]0, \infty[$ the relation $\phi_\lambda(x) = 0$ iff $f_A(x) = f_H(x)$ iff $x = x^* := \frac{\alpha_H - \alpha_A}{\beta_A - \beta_H} \in]0, \infty[$.

3.7. Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,2} \times]0, 1[$

For this parameter constellation, one has $\phi_\lambda(0) = 0$ and $\phi'_\lambda(0) = 0$ (cf. Properties 3 (P16), (P17)). Thus, the only admissible intercept choice satisfying (47) is $r_\lambda^U = 0 = p_\lambda^U - \alpha_\lambda$ (i.e., $p_\lambda^U = p^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = \alpha_\lambda = \alpha > 0$), and the minimal admissible slope which implies (35) for all $x \in \mathbb{N}_0$ is given by $s_\lambda^U = \frac{\phi_\lambda(1) - \phi_\lambda(0)}{1 - 0} = q_\lambda^U - \beta_\lambda = a_1^{(q_\lambda^U)} < 0$ (i.e., $q_\lambda^U = q^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (\alpha + \beta_A)^\lambda (\alpha + \beta_H)^{1-\lambda} - \alpha > 0$). Analogously to the investigation for $\mathcal{P}_{SP,1}$ in the above-mentioned Section 3.3, one can derive that $\left(a_n^{(q_\lambda^U)}\right)_{n \in \mathbb{N}}$ is strictly negative, strictly decreasing, and converges to $x_0^{(q_\lambda^U)} \in]-\beta_\lambda, q_\lambda^U - \beta_\lambda[$ as indicated in Properties 1 (P1). Moreover, in the same manner as for the case $\mathcal{P}_{SP,1}$ this leads to

Proposition 7. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,2} \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $p_\lambda^U = \alpha$, $q_\lambda^U = (\alpha + \beta_A)^\lambda (\alpha + \beta_H)^{1-\lambda} - \alpha$

- (a) $B_{\lambda, X_0, 1}^U = \exp \left\{ (q_\lambda^U - \beta_\lambda) \cdot X_0 \right\} < 1,$
- (b) the sequence $\left(B_{\lambda, X_0, n}^U \right)_{n \in \mathbb{N}}$ of upper bounds for $H_\lambda(P_{A,n} || P_{H,n})$ given by

$$B_{\lambda, X_0, n}^U = \exp \left\{ a_n^{(q_\lambda^U)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)} \right\}$$
 is strictly decreasing,
- (c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^U = 0 = \lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}),$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log B_{\lambda, X_0, n}^U = p_\lambda^U \cdot e^{x_0^{(q_\lambda^U)}} - \alpha_\lambda = \alpha \left(e^{x_0^{(q_\lambda^U)}} - 1 \right) < 0.$
- (e) the map $X_0 \mapsto B_{\lambda, X_0, n}^U$ is strictly decreasing.

3.8. Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,3a} \times]0, 1[$

From Properties 3 (P16) one gets $\phi_\lambda(0) < 0$, whereas $\phi'_\lambda(0)$ can assume any sign, take e.g., the parameters $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (1.8, 0.9, 2.7, 0.7, 0.5)$ for $\phi'_\lambda(0) < 0$, $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (1.8, 0.9, 2.8, 0.7, 0.5)$ for $\phi'_\lambda(0) = 0$ and $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (1.8, 0.9, 2.9, 0.7, 0.5)$ for $\phi'_\lambda(0) > 0$; within our running-example epidemiological context of Section 2.3, this corresponds to a “nearly dangerous” infectious-disease-transmission situation (\mathcal{H}) (with nearly critical reproduction number $\beta_H = 0.9$ and importation mean of $\alpha_H = 0.7$), whereas (\mathcal{A}) describes a “dangerous” situation (with supercritical $\beta_A = 1.8$ and $\alpha_A = 2.7, 2.8, 2.9$). However, in all three subcases there holds $\max_{x \in \mathbb{N}_0} \phi_\lambda(x) \leq \max_{x \in]0, \infty[} \phi_\lambda(x) < 0$. Thus, there clearly exist parameters $p_\lambda^U = p^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$, $q_\lambda^U = q^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ with $p_\lambda^U \in [\alpha_A^\lambda \alpha_H^{1-\lambda}, \alpha_\lambda[$ and $q_\lambda^U \in [\beta_A^\lambda \beta_H^{1-\lambda}, \beta_\lambda[$ (implying (47)) such that (35) is satisfied. As explained above, we get the following

Proposition 8. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,3a} \times]0, 1[$ there exist parameters p_λ^U, q_λ^U which satisfy $p_\lambda^U \in [\alpha_A^\lambda \alpha_H^{1-\lambda}, \alpha_\lambda[$ and $q_\lambda^U \in [\beta_A^\lambda \beta_H^{1-\lambda}, \beta_\lambda[$ as well as (35) for all $x \in \mathbb{N}_0$, and for all such pairs $(p_\lambda^U, q_\lambda^U)$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds

- (a) $B_{\lambda, X_0, 1}^U = \exp \left\{ (q_\lambda^U - \beta_\lambda) \cdot X_0 + p_\lambda^U - \alpha_\lambda \right\} < 1,$
- (b) the sequence $\left(B_{\lambda, X_0, n}^U \right)_{n \in \mathbb{N}}$ of upper bounds for $H_\lambda(P_{A,n} || P_{H,n})$ given by

$$B_{\lambda, X_0, n}^U = \exp \left\{ a_n^{(q_\lambda^U)} X_0 + \sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)} \right\}$$
 is strictly decreasing,
- (c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^U = 0 = \lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}),$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log B_{\lambda, X_0, n}^U = p_\lambda^U \cdot e^{x_0^{(q_\lambda^U)}} - \alpha_\lambda < 0,$
- (e) the map $X_0 \mapsto B_{\lambda, X_0, n}^U$ is strictly decreasing.

Notice that all parts of this proposition also hold true for parameter pairs $(p_\lambda^U, q_\lambda^U)$ satisfying (35) and additionally either $p_\lambda^U = \alpha_\lambda, q_\lambda^U < \beta_\lambda$ or $p_\lambda^U < \alpha_\lambda, q_\lambda^U = \beta_\lambda$.

Let us briefly illuminate the above-mentioned possible parameter choices, where we begin with the case of $\phi'_\lambda(0) \leq 0$, which corresponds to $\lambda \beta_A (\alpha_A / \alpha_H)^{\lambda-1} + (1 - \lambda) \beta_H (\alpha_A / \alpha_H)^\lambda - \beta_\lambda \leq 0$ (cf. (P17)); then, the function $\phi_\lambda(\cdot)$ is strictly negative, strictly decreasing, and—due to (P19)—strictly concave (and thus, the assumption $\frac{\alpha_H - \alpha_A}{\beta_A - \beta_H} < 0$ is superfluous here). One pragmatic but yet reasonable parameter

choice is the following: take any intercept $p_\lambda^U \in [\alpha_A^\lambda \alpha_H^{1-\lambda}, \alpha_\lambda]$ such that $(p_\lambda^U - \alpha_\lambda) + 2(\phi_\lambda(1) - (p_\lambda^U - \alpha_\lambda)) \geq \phi_\lambda(2)$ (i.e., $2(\alpha_A + \beta_A)^\lambda (\alpha_H + \beta_H)^{1-\lambda} - p_\lambda^U + \alpha_\lambda \geq (\alpha_A + 2\beta_A)^\lambda (\alpha_H + 2\beta_H)^{1-\lambda}$) and $q_\lambda^U := \phi_\lambda(1) - (p_\lambda^U - \alpha_\lambda) + \beta_\lambda = (\alpha_A + \beta_A)^\lambda (\alpha_H + \beta_H)^{1-\lambda} - p_\lambda^U$, which corresponds to a linear function ϕ_λ^U which is (i) nonpositive on \mathbb{N}_0 and strictly negative on \mathbb{N} , and (ii) larger than or equal to ϕ_λ on \mathbb{N}_0 , strictly larger than ϕ_λ on $\mathbb{N} \setminus \{1, 2\}$, and equal to ϕ_λ at the point $x = 1$ (“discrete tangent or secant line through $x = 1$ ”). One can easily see that (due to the restriction (34)) not all $p_\lambda^U \in [\alpha_A^\lambda \alpha_H^{1-\lambda}, \alpha_\lambda]$ might qualify for the current purpose. For the particular choice $p_\lambda^U = \alpha_A^\lambda \alpha_H^{1-\lambda}$ and $q_\lambda^U = (\alpha_A + \beta_A)^\lambda (\alpha_H + \beta_H)^{1-\lambda} - \alpha_A^\lambda \alpha_H^{1-\lambda}$ one obtains $r_\lambda^U = p_\lambda^U - \alpha_\lambda = b_1^{(p_\lambda^U, q_\lambda^U)} < 0$ (cf. Lemma A1) and $s_\lambda^U = q_\lambda^U - \beta_\lambda = \phi_\lambda(1) - \phi_\lambda(0) = a_1^{(q_\lambda^U)} < 0$ (secant line through $\phi_\lambda(0)$ and $\phi_\lambda(1)$).

For the remaining case $\phi'_\lambda(0) > 0$, which corresponds to $\lambda\beta_A(\alpha_A/\alpha_H)^{\lambda-1} + (1 - \lambda)\beta_H(\alpha_A/\alpha_H)^\lambda - \beta_\lambda > 0$, the function $\phi_\lambda(\cdot)$ is strictly negative, strictly concave and hump-shaped (cf. (P18)). For the derivation of the parameter choices, we employ $x_{\max} := \operatorname{argmax}_{x \in]0, \infty[} \phi_\lambda(x)$ which is the unique solution of

$$\lambda\beta_A \left[\left(\frac{f_A(x)}{f_H(x)} \right)^{\lambda-1} - 1 \right] + (1 - \lambda)\beta_H \left[\left(\frac{f_A(x)}{f_H(x)} \right)^\lambda - 1 \right] = 0, \quad x \in]0, \infty[, \quad (50)$$

(cf. (P17), (P19)); notice that $x = x^* := \frac{\alpha_H - \alpha_A}{\beta_A - \beta_H} \in]0, \infty[$ formally satisfies the Equation (50) but does not qualify because of the current restriction $x^* < 0$.

Let us first inspect the case $\phi_\lambda(\lfloor x_{\max} \rfloor) > \phi_\lambda(\lfloor x_{\max} \rfloor + 1)$, where $\lfloor x \rfloor$ denotes the integer part of x . Consider the subcase $\phi_\lambda(\lfloor x_{\max} \rfloor) + \lfloor x_{\max} \rfloor (\phi_\lambda(\lfloor x_{\max} \rfloor) - \phi_\lambda(\lfloor x_{\max} \rfloor + 1)) \leq 0$, which means that the secant line through $\phi_\lambda(\lfloor x_{\max} \rfloor)$ and $\phi_\lambda(\lfloor x_{\max} \rfloor + 1)$ possesses a non-positive intercept. In this situation it is reasonable to choose as *intercept* any $p_\lambda^U - \alpha_\lambda = b_1^{(p_\lambda^U, q_\lambda^U)} = r_\lambda^U \in [\phi_\lambda(\lfloor x_{\max} \rfloor), \phi_\lambda(\lfloor x_{\max} \rfloor) + \lfloor x_{\max} \rfloor (\phi_\lambda(\lfloor x_{\max} \rfloor) - \phi_\lambda(\lfloor x_{\max} \rfloor + 1))]$, and as corresponding *slope* $q_\lambda^U - \alpha_\lambda = a_1^{(q_\lambda^U)} = s_\lambda^U = \frac{\phi_\lambda(\lfloor x_{\max} \rfloor) - r_\lambda^U}{(\lfloor x_{\max} \rfloor) - 0} \leq 0$. A larger intercept would lead to a linear function ϕ_λ^U for which (35) is not valid at $\lfloor x_{\max} \rfloor + 1$. In the other subcase $\phi_\lambda(\lfloor x_{\max} \rfloor) + x_{\max} (\phi_\lambda(\lfloor x_{\max} \rfloor) - \phi_\lambda(\lfloor x_{\max} \rfloor + 1)) > 0$, one can choose any intercept $p_\lambda^U - \alpha_\lambda = b_1^{(p_\lambda^U, q_\lambda^U)} = r_\lambda^U \in [\phi_\lambda(\lfloor x_{\max} \rfloor), 0]$ and as corresponding slope $q_\lambda^U - \alpha_\lambda = a_1^{(q_\lambda^U)} = s_\lambda^U = \frac{\phi_\lambda(\lfloor x_{\max} \rfloor) - r_\lambda^U}{(\lfloor x_{\max} \rfloor) - 0} \leq 0$ (notice that the corresponding line ϕ_λ^U is on $]\lfloor x_{\max} \rfloor, \infty[$ strictly larger than the secant line through $\phi_\lambda(\lfloor x_{\max} \rfloor)$ and $\phi_\lambda(\lfloor x_{\max} \rfloor + 1)$).

If $\phi_\lambda(\lfloor x_{\max} \rfloor) \leq \phi_\lambda(\lfloor x_{\max} \rfloor + 1)$, one can proceed as above by substituting the crucial pair of points $(\lfloor x_{\max} \rfloor, \lfloor x_{\max} \rfloor + 1)$ with $(\lfloor x_{\max} \rfloor + 1, \lfloor x_{\max} \rfloor + 2)$ and examining the analogous two subcases.

3.9. Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,3b} \times]0, 1[$

The only difference to the preceding Section 3.8 is that—due to Properties 3 (P15)—the maximum value of $\phi_\lambda(\cdot)$ now achieves 0, at the positive *non-integer* point $x_{\max} = x^* = \frac{\alpha_H - \alpha_A}{\beta_A - \beta_H} \in]0, \infty[\setminus \mathbb{N}$ (take e.g., $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (1.8, 0.9, 1.1, 3.0, 0.5)$ as an example, which within our running-example epidemiological context of Section 2.3 corresponds to a “nearly dangerous” infectious-disease-transmission situation (\mathcal{H}) (with nearly critical reproduction number $\beta_H = 0.9$ and importation mean of $\alpha_H = 3$), whereas (\mathcal{A}) describes a “dangerous” situation (with supercritical $\beta_A = 1.8$ and $\alpha_A = 1.1$)); this implies that $\phi_\lambda(x) < 0$ for all x on the relevant subdomain \mathbb{N}_0 . Due to (P16), (P17) and (P19) one gets automatically $\lambda\beta_A(\alpha_A/\alpha_H)^{\lambda-1} + (1 - \lambda)\beta_H(\alpha_A/\alpha_H)^\lambda - \beta_\lambda > 0$ for all $\lambda \in]0, 1[$. Analogously to Section 3.8, there exist parameter $p_\lambda^U \in [\alpha_A^\lambda \alpha_H^{1-\lambda}, \alpha_\lambda]$ and $q_\lambda^U \in [\beta_A^\lambda \beta_H^{1-\lambda}, \beta_\lambda]$ such that (47) and (35) are satisfied. Thus, all the assertions (a) to (e) of Proposition 8 also hold true for the current parameter constellations.

3.10. Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,3c} \times]0, 1[$

The only difference to the preceding Section 3.9 is that the maximum value of $\phi_\lambda(\cdot)$ now achieves 0 at the integer point $x_{\max} = x^* = \frac{\alpha_H - \alpha_A}{\beta_A - \beta_H} \in \mathbb{N}$ (take e.g., $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (1.8, 0.9, 1.2, 3.0, 0.5)$ as an example). Accordingly, there do not exist parameters p_λ^U, q_λ^U , such that (35) and (47) are satisfied simultaneously. The only parameter pair that ensures $\exp \left\{ a_n^{(q_\lambda^U)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)} \right\} \leq 1$ for all $n \in \mathbb{N}$ and all $X_0 \in \mathbb{N}$ without further investigations, leads to the choices $p_\lambda^U = \alpha_\lambda$ as well as $q_\lambda^U = \beta_\lambda$. Consequently, $B_{\lambda, X_0, n}^U \equiv 1$, which coincides with the general upper bound (9), but violates the above-mentioned desired Goal (G1). However, there might exist parameters $p_\lambda^U < \alpha_\lambda, q_\lambda^U > \beta_\lambda$ or $p_\lambda^U > \alpha_\lambda, q_\lambda^U < \beta_\lambda$, such that at least the parts (c) and (d) of Proposition 8 are satisfied. Nevertheless, by using a conceptually different method we can prove

$$H_\lambda(P_{A,n} || P_{H,n}) < 1 \quad \forall n \in \mathbb{N} \setminus \{1\} \quad \text{as well as the convergence} \quad \lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}) = 0 \quad (51)$$

which will be used for the study of complete asymptotical distinguishability (entire separation) below. This proof is provided in Appendix A.1.

3.11. Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4a} \times]0, 1[$

This setup and the remaining setup $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4b} \times]0, 1[$ (see the next Section 3.12) are the only constellations where $\phi_\lambda(\cdot)$ is strictly negative and strictly increasing, with $\lim_{x \rightarrow \infty} \phi_\lambda(x) = \lim_{x \rightarrow \infty} \phi'_\lambda(x) = 0$, leading to the choices $p_\lambda^U = \alpha_\lambda$ as well as $q_\lambda^U = \beta_\lambda = \beta$ under the restriction that $\exp \left\{ a_n^{(q_\lambda^U)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)} \right\} \leq 1$ for all $n \in \mathbb{N}$ and all $X_0 \in \mathbb{N}$. Consequently, one has $B_{\lambda, X_0, n}^U \equiv 1$, which is consistent with the general upper bound (9) but violates the above-mentioned desired Goal (G1). Unfortunately, the proof method of (51) (cf. Appendix A.1) can't be carried over to the current setup. The following proposition states two of the above-mentioned desired assertions which can be verified by a completely different proof method, which is also given in Appendix A.1.

Proposition 9. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4a} \times]0, 1[$ there exist parameters $p_\lambda^U < \alpha_\lambda, 1 > q_\lambda^U > \beta_\lambda = \beta$ such that (35) is satisfied for all $x \in [0, \infty[$ and such that for all initial population sizes $X_0 \in \mathbb{N}$ the parts (c) and (d) of Proposition 8 hold true.

3.12. Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4b} \times]0, 1[$

The assertions preceding Proposition 9 remain valid. However, any linear upper bound of the function $\phi_\lambda(\cdot)$ on the domain \mathbb{N}_0 possesses the slope $q_\lambda^U - \beta_\lambda \geq 0$. If $q_\lambda^U = \beta_\lambda$, then the intercept is $p_\lambda^U - \alpha_\lambda = 0$ leading to $B_{\lambda, X_0, n}^U \equiv 1$ and thus Goal (G1) is violated. If we use a slope $q_\lambda^U - \beta_\lambda > 0$, then both the sequences $\left(a_n^{(q_\lambda^U)} \right)_{n \in \mathbb{N}}$ and $\left(b_n^{(p_\lambda^U, q_\lambda^U)} \right)_{n \in \mathbb{N}}$ are strictly increasing and diverge to ∞ . This comes from Properties 1 (P3b) and (P7b) since $q_\lambda^U > \beta_\lambda = \beta \geq 1$. Altogether, this implies that the corresponding upper bound component $\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)}$ (cf. (42)) diverges to ∞ as well. This leads to

Proposition 10. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4b} \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ there do not exist parameters $p_\lambda^U \geq 0, q_\lambda^U \geq 0$ such that (35) is satisfied and such that the parts (c) and (d) of Proposition 8 hold true.

3.13. Concluding Remarks on Alternative Upper Bounds for all Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$

As mentioned earlier on, starting from Section 3.6 we have principally focused on constructing upper bounds $B_{\lambda, X_0, n}^U$ of the Hellinger integrals, starting from p_λ^U, q_λ^U which fulfill (35) as well as further constraints depending on the Goals (G1) and (G2). For the setups in the Sections 3.7–3.9, we have

proved the existence of *special parameter choices* p_λ^U, q_λ^U which were consistent with (G1) and (G2). Furthermore, for the constellation in the Section 3.11 we have found parameters such that at least (G2) is satisfied. In contrast, for the setup of Section 3.12 we have not found any choices which are consistent with (G1) and (G2), leading to the “cut-off bound” $B_{\lambda, X_0, n}^U \equiv 1$ which gives no improvement over the generally valid upper bound (9).

In the following, we present some *alternative choices* of p_λ^U, q_λ^U which—depending on the parameter constellation $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$ —may or may not lead to upper bounds $B_{\lambda, X_0, n}^U$ which are consistent with Goal (G1) or with (G2) (and which are maybe weaker or better than resp. incomparable with the previous upper bounds when dealing with some relaxations of (G1), such as e.g., $H_\lambda(P_{A,n} || P_{H,n}) < 1$ for all but finitely many $n \in \mathbb{N}$).

As a first alternative choice for a linear upper bound of $\phi_\lambda(\cdot)$ (cf. (35)) one could use the asymptote $\tilde{\phi}_\lambda(\cdot)$ (cf. Properties 3 (P20)) with the parameters $p_\lambda^U := \tilde{p}_\lambda = \lambda \alpha_A (\beta_A / \beta_H)^{\lambda-1} + (1 - \lambda) \alpha_H (\beta_A / \beta_H)^\lambda$ and $q_\lambda^U := \tilde{q}_\lambda = \beta_A^\lambda \beta_H^{1-\lambda}$. Another important linear upper bound of $\phi_\lambda(\cdot)$ is the tangent line $\phi_{\lambda,y}^{\tan}(\cdot)$ on $\phi_\lambda(\cdot)$ at an arbitrarily fixed point $y \in]0, \infty[$, which amounts to

$$\phi_{\lambda,y}^{\tan}(x) := r_{\lambda,y}^{\tan} + s_{\lambda,y}^{\tan} \cdot x := (p_{\lambda,y}^{\tan} - \alpha_\lambda) + (q_{\lambda,y}^{\tan} - \beta_\lambda) \cdot x := (\phi_\lambda(y) - y \cdot \phi'_\lambda(y)) + \phi'_\lambda(y) \cdot x, \tag{52}$$

where $\phi'_\lambda(\cdot)$ is given by (P17). Notice that this upper bound is for $y \in]0, \infty[\setminus \mathbb{N}$ “not tight” in the sense that $\phi_{\lambda,y}^{\tan}(\cdot)$ does not hit the function $\phi_\lambda(\cdot)$ on \mathbb{N}_0 (where the generation sizes “live”); moreover, $\phi_{\lambda,y}^{\tan}(x)$ might take on strictly positive values for large enough points x which is counter-productive for Goal (G1). Another alternative choice of a linear upper bound for $\phi_\lambda(\cdot)$, which in contrast to the tangent line is “tight” (but not necessarily avoiding the strict positivity), is the secant line $\phi_{\lambda,k}^{\sec}(\cdot)$ across its arguments k and $k + 1$, given by

$$\begin{aligned} \phi_{\lambda,k}^{\sec}(x) &:= r_{\lambda,k}^{\sec} + s_{\lambda,k}^{\sec} \cdot x := (p_{\lambda,k}^{\sec} - \alpha_\lambda) + (q_{\lambda,k}^{\sec} - \beta_\lambda) \cdot x \\ &:= [\phi_\lambda(k) - k \cdot (\phi_\lambda(k+1) - \phi_\lambda(k))] + (\phi_\lambda(k+1) - \phi_\lambda(k)) \cdot x. \end{aligned} \tag{53}$$

Another alternative choice is the horizontal line

$$\phi_\lambda^{\text{hor}}(x) \equiv \max \{ \phi_\lambda(y), y \in \mathbb{N}_0 \}. \tag{54}$$

For $p_\lambda^U \in \{ \tilde{p}_\lambda, p_{\lambda,y}^{\tan}, p_{\lambda,y}^{\sec} \}$ and $q_\lambda^U \in \{ q_{\lambda,y}^{\tan}, q_{\lambda,y}^{\sec} \}$ it is possible that in some parameter cases $(\beta_A, \beta_H, \alpha_A, \alpha_H)$ either the intercept $r_\lambda^U = p_\lambda^U - \alpha_\lambda$ is strictly larger than zero or the slope $s_\lambda^U = q_\lambda^U - \beta_\lambda$ is strictly larger than zero. Thus, it can happen that $\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} > 1$ for some (and even for all) $n \in \mathbb{N}$, such that the corresponding upper bound $B_{\lambda, X_0, n}^U$ for the Hellinger integral $H_\lambda(P_{A,n} || P_{H,n})$ amounts to the cut-off at 1. However, due to Properties 1 (P5) and (P7a), the sequence $\left(\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} \right)_{n \in \mathbb{N}}$ may become smaller than 1 and may finally converge to zero. Due to Properties 2 (P14), this upper bound can even be tighter (smaller) than those bounds derived from parameters p_λ^U, q_λ^U fulfilling (47).

As far as our desired Hellinger integral bounds are concerned, in the setup of Section 3.11—where $\lim_{y \rightarrow \infty} \phi_{\lambda,y}^{\tan}(\cdot) \equiv 0$ —for the proof of Proposition 9 in Appendix A.1 we shall employ the mappings $y \mapsto \phi_{\lambda,y}^{\tan}$ resp. $y \mapsto p_{\lambda,y}^{\tan}$ resp. $y \mapsto q_{\lambda,y}^{\tan}$. These will also be used for the proof of the below-mentioned Theorem 4.

3.14. Intermezzo 1: Application to Asymptotical Distinguishability

The above-mentioned investigations can be applied to the context of Section 2.6 on asymptotical distinguishability. Indeed, with the help of the Definitions 1 and 2 as well as the equivalence relations (25) and (26) we obtain the following

Corollary 1.

- (a) For all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,4b}$ and all initial population sizes $X_0 \in \mathbb{N}$, the corresponding sequences $(P_{\mathcal{A},n})_{n \in \mathbb{N}_0}$ and $(P_{\mathcal{H},n})_{n \in \mathbb{N}_0}$ are entirely separated (completely asymptotically distinguishable).
- (b) For all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{NI}$ with $\beta_{\mathcal{A}} \leq 1$ and all initial population sizes $X_0 \in \mathbb{N}$, the sequence $(P_{\mathcal{A},n})_{n \in \mathbb{N}_0}$ is contiguous to $(P_{\mathcal{H},n})_{n \in \mathbb{N}_0}$.
- (c) For all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{NI}$ with $\beta_{\mathcal{A}} > 1$ and all initial population sizes $X_0 \in \mathbb{N}$, the sequence $(P_{\mathcal{A},n})_{n \in \mathbb{N}_0}$ is neither contiguous to nor entirely separated to $(P_{\mathcal{H},n})_{n \in \mathbb{N}_0}$.

The proof of Corollary 1 will be given in Appendix A.1.

Remark 3.

- (a) Assertion (c) of Corollary 1 contrasts the case of Gaussian processes with independent increments where one gets either entire separation or mutual contiguity (see e.g., Liese & Vajda [1]).
- (b) By putting Corollary 1(b) and (c) together, we obtain for different “criticality pairs” in the non-immigration case \mathcal{P}_{NI} the following asymptotical distinguishability types:
 $(P_{\mathcal{A},n}) \triangleleft \triangleright (P_{\mathcal{H},n})$ if $\beta_{\mathcal{A}} \leq 1, \beta_{\mathcal{H}} \leq 1$; $(P_{\mathcal{A},n}) \triangleleft \boxplus (P_{\mathcal{H},n})$ if $\beta_{\mathcal{A}} \leq 1, \beta_{\mathcal{H}} > 1$;
 $(P_{\mathcal{A},n}) \triangleright \boxplus (P_{\mathcal{H},n})$ if $\beta_{\mathcal{A}} > 1, \beta_{\mathcal{H}} \leq 1$; $(P_{\mathcal{A},n}) \triangleright \boxminus (P_{\mathcal{H},n})$ and $(P_{\mathcal{A},n}) \overline{\Delta} (P_{\mathcal{H},n})$ if $\beta_{\mathcal{A}} > 1, \beta_{\mathcal{H}} > 1$;
 in particular, for \mathcal{P}_{NI} the sequences $(P_{\mathcal{A},n})_{n \in \mathbb{N}_0}$ and $(P_{\mathcal{H},n})_{n \in \mathbb{N}_0}$ are not completely asymptotically inseparable (indistinguishable).
- (c) In the light of the above-mentioned characterizations of contiguity resp. entire separation by means of Hellinger integral limits, the finite-time-horizon results on Hellinger integrals given in the “ $\lambda \in]0, 1[$ parts” of Theorem 1, the Sections 3.3–3.13 and also in the below-mentioned Section 6 can loosely be interpreted as “finite-sample (rather than asymptotical) distinguishability” assertions.

3.15. Intermezzo 2: Application to Decision Making under Uncertainty

3.15.1. Bayesian Decision Making

The above-mentioned investigations can be applied to the context of Section 2.5 on dichotomous Bayesian decision making on the space of all possible path scenarios (path space) of Poissonian Galton-Watson processes without/with immigration GW(I) (e.g., in combination with our running-example epidemiological context of Section 2.3). More detailed, for the minimal mean decision loss (Bayes risk) \mathcal{R}_n defined by (18) we can derive upper (respectively lower) bounds by using (19) respectively (20) together with the exact values or the upper (respectively lower) bounds of the Hellinger integrals $H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ derived in the “ $\lambda \in]0, 1[$ parts” of Theorem 1, the Sections 3.3–3.13 (and also in the below-mentioned Section 6); instead of providing the corresponding outcoming formulas—which is merely repetitive—we give the illustrative

Example 1. Based on a sample path observation $\mathcal{X}_n := \{X_\ell : \ell = 1, \dots, n\}$ of a GWI, which is either governed by a hypothesis law $P_{\mathcal{H}}$ or an alternative law $P_{\mathcal{A}}$, we want to make a dichotomous optimal Bayesian decision described in Section 2.5, namely, decide between an action $d_{\mathcal{H}}$ “associated with” $P_{\mathcal{H}}$ and an action $d_{\mathcal{A}}$ “associated with” $P_{\mathcal{A}}$, with pre-given loss function (16) involving constants $L_{\mathcal{A}} > 0, L_{\mathcal{H}} > 0$ which e.g., arise as bounds from quantities in worst-case scenarios.

For this, let us exemplarily deal with initial population $X_0 = 5$ as well as parameter setup $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) = (1.2, 0.9, 4, 3) \in \mathcal{P}_{SP,1}$; within our running-example epidemiological context of Section 2.3, this corresponds e.g., to a setup where one is encountered with a novel infectious disease (such as COVID-19) of non-negligible fatality rate, and (\mathcal{A}) reflects a “potentially dangerous” infectious-disease-transmission situation (with supercritical reproduction number $\beta_{\mathcal{A}} = 1.2$ and importation mean of $\alpha_{\mathcal{A}} = 4$, for weekly appearing new incidence-generations) whereas (\mathcal{H}) describes a “milder” situation (with subcritical $\beta_{\mathcal{H}} = 0.9$

and $\alpha_{\mathcal{H}} = 3$). Moreover, let $d_{\mathcal{H}}$ and $d_{\mathcal{A}}$ reflect two possible sets of interventions (control measures) in the course of pandemic risk management, with respective “worst-case type” decision losses $L_{\mathcal{A}} = 600$ and $L_{\mathcal{H}} = 300$ (e.g., in units of billion Euros or U.S. Dollars). Additionally we assume the prior probabilities $\pi = Pr(\mathcal{H}) = 1 - Pr(\mathcal{A}) = 0.5$, which results in the prior-loss constants $\mathfrak{L}_{\mathcal{A}} = 300$ and $\mathfrak{L}_{\mathcal{H}} = 150$. In order to obtain bounds for the corresponding minimal mean decision loss (Bayes Risk) \mathcal{R}_n defined in (18) we can employ the general Stummer-Vajda bounds (cf. [15]) (19) and (20) in terms of the Hellinger integral $H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ (with arbitrary $\lambda \in]0, 1[$), and combine this with the appropriate detailed results on the latter from the preceding subsections. To demonstrate this, let us choose $\lambda = 0.5$ (for which $H_{1/2}(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ can be interpreted as a multiple of the Bhattacharyya coefficient between the two competing GWI) respectively $\lambda = 0.9$, leading to the parameters $p_{0.5}^E = 3.464$, $q_{0.5}^E = 1.039$ respectively $p_{0.9}^E = 3.887$, $q_{0.9}^E = 1.166$ (cf. (33)). Combining (19) and (20) with Theorem 1 (a)– which provides us with the exact recursive values of $H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ in terms of the sequence $a_n^{(q_{\lambda}^E)}$ (cf. (36))– we obtain for $\lambda = 0.5$ the bounds

$$\mathcal{R}_n \leq \mathcal{R}_n^U := 2.121 \cdot 10^2 \cdot \exp \left\{ 5 \cdot a_n^{(1.039)} + \frac{10}{3} \cdot \sum_{k=1}^n a_k^{(1.039)} \right\},$$

$$\mathcal{R}_n \geq \mathcal{R}_n^L := 100 \cdot \exp \left\{ 10 \cdot a_n^{(1.039)} + \frac{20}{3} \cdot \sum_{k=1}^n a_k^{(1.039)} \right\},$$

whereas for $\lambda = 0.9$ we get

$$\mathcal{R}_n \leq \mathcal{R}_n^U := 2.799 \cdot 10^2 \cdot \exp \left\{ 5 \cdot a_n^{(1.166)} + \frac{10}{3} \cdot \sum_{k=1}^n a_k^{(1.166)} \right\},$$

$$\mathcal{R}_n \geq \mathcal{R}_n^L := 3.902 \cdot \exp \left\{ 50 \cdot a_n^{(1.166)} + \frac{100}{3} \cdot \sum_{k=1}^n a_k^{(1.166)} \right\}.$$

Figure 1 illustrates the lower (orange resp. cyan) and upper (red resp. blue) bounds \mathcal{R}_n^L resp. \mathcal{R}_n^U of the Bayes Risk \mathcal{R}_n employing $\lambda = 0.5$ resp. $\lambda = 0.9$ on both a unit scale (left graph) and a logarithmic scale (right graph). The lightgrey/grey/black curves correspond to the (18)-based empirical evaluation of the Bayes risk sequence $(\mathcal{R}_n^{sample})_{n=1, \dots, 50}$ from three independent Monte Carlo simulations of 10000 GWI sample paths (each) up to time horizon 50.

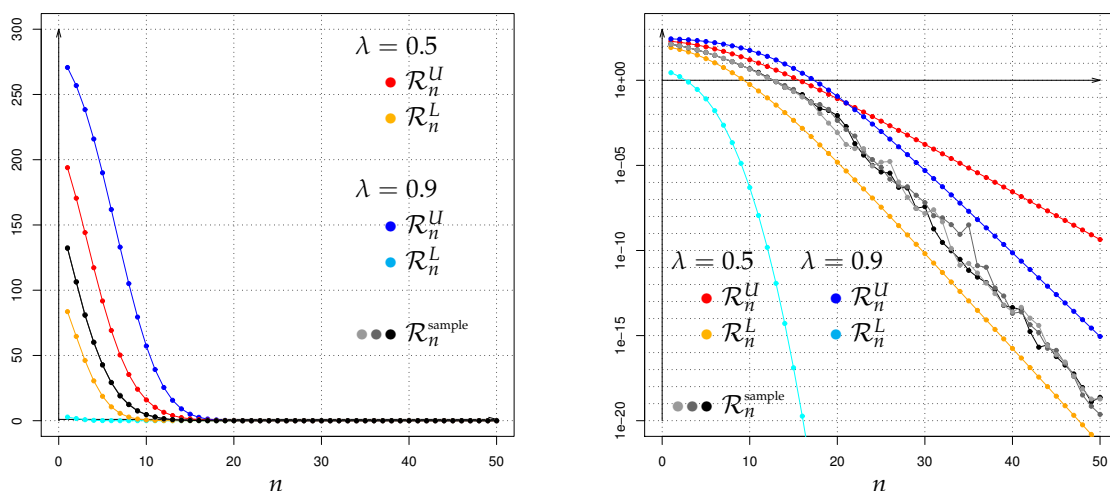


Figure 1. Bayes risk bounds (using $\lambda = 0.5$ (red/orange) resp. $\lambda = 0.9$ (blue/cyan)) and Bayes risk simulations (lightgrey/grey/black) on a unit (left graph) and logarithmic (right graph) scale in the parameter setup $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) = (1.2, 0.9, 4, 3) \in \mathcal{P}_{SP,1}$, with initial population $X_0 = 5$ and prior-loss constants $\mathfrak{L}_{\mathcal{A}} = 300$ and $\mathfrak{L}_{\mathcal{H}} = 150$.

3.15.2. Neyman-Pearson Testing

By combining (23) with the exact values resp. upper bounds of the Hellinger integrals $H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ from the preceding subsections, we obtain for our context of GW(I) with Poisson offspring and Poisson immigration (including the non-immigration case) some upper bounds of the minimal type II error probability $\mathcal{E}_\zeta(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ in the class of the tests for which the type I error probability is at most $\zeta \in]0, 1[$, which can also be immediately rewritten as lower bounds for the power $1 - \mathcal{E}_\zeta(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ of a most powerful test at level ζ . As for the Bayesian context of Section 3.15.1, instead of providing the—merely repetitive—outcoming formulas for the bounds of $\mathcal{E}_\zeta(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ we give the illustrative

Example 2. Consider the Figures 2 and 3 which deal with initial population $X_0 = 5$ and the parameter setup $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) = (0.3, 1.2, 1, 4) \in \mathcal{P}_{SP,1}$; within our running-example epidemiological context of Section 2.3, this corresponds to a “potentially dangerous” infectious-disease-transmission situation (\mathcal{H}) (with supercritical reproduction number $\beta_{\mathcal{H}} = 1.2$ and importation mean of $\alpha_{\mathcal{H}} = 4$), whereas (\mathcal{A}) describes a “very mild” situation (with “low” subcritical $\beta_{\mathcal{A}} = 0.3$ and $\alpha_{\mathcal{A}} = 1$). Figure 2 shows the lower and upper bounds of $\mathcal{E}_\zeta(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ with $\zeta = 0.05$, evaluated from the Formulas (23) and (24), together with the exact values of the Hellinger integral $H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n})$, cf. Theorem 1 (recall that we are in the setup $\mathcal{P}_{SP,1}$) on both a unit scale (left graph) and a logarithmic scale (right graph). The orange resp. red resp. purple curves correspond to the outcoming upper bounds $\mathcal{E}_n^U := \mathcal{E}_n^U(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ (cf. (23)) with parameters $\lambda = 0.3$ resp. $\lambda = 0.5$ resp. $\lambda = 0.7$. The green resp. cyan resp. blue curves correspond to the lower bounds $\mathcal{E}_n^L := \mathcal{E}_n^L(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ (cf. (24)) with parameters $\lambda = 2$ resp. $\lambda = 1.5$ resp. $\lambda = 1.1$. Notice the different λ -ranges in (23) and (24). In contrast, Figure 3 compares the lower bound \mathcal{E}_n^L (for fixed $\lambda = 1.1$) with the upper bound \mathcal{E}_n^U (for fixed $\lambda = 0.5$) of the minimal type II error probability $\mathcal{E}_\zeta(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ for different levels $\zeta = 0.1$ (orange for the lower and cyan for the upper bound), $\zeta = 0.05$ (green and magenta) and $\zeta = 0.01$ (blue and purple) on both a unit scale (left graph) and a logarithmic scale (right graph).

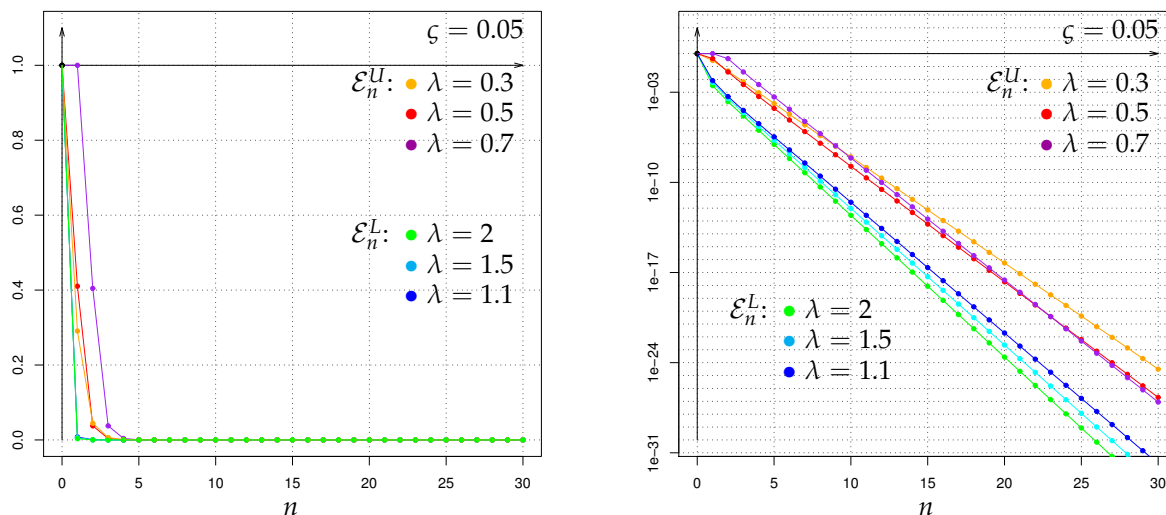


Figure 2. Different lower bounds \mathcal{E}_n^L (using $\lambda \in \{1.1, 1.5, 2\}$) and upper bounds \mathcal{E}_n^U (using $\lambda \in \{0.3, 0.5, 0.7\}$) of the minimal type II error probability $\mathcal{E}_\zeta(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ for fixed level $\zeta = 0.05$ in the parameter setup $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) = (0.3, 1.2, 1, 4) \in \mathcal{P}_{SP,1}$ together with initial population $X_0 = 5$ on both a unit scale (left graph) and a logarithmic scale (right graph).

3.16. Goals for Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0, 1])$

Recall from (49) the set $\mathcal{P}_{SP} := \{(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in]0, \infty[^4 : (\alpha_{\mathcal{A}} \neq \alpha_{\mathcal{H}}) \text{ or } (\beta_{\mathcal{A}} \neq \beta_{\mathcal{H}}) \text{ or both}\}$ and the “equal-fraction-case” set $\mathcal{P}_{SP,1} := \{(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{SP} : \alpha_{\mathcal{A}} \neq \alpha_{\mathcal{H}}, \beta_{\mathcal{A}} \neq \beta_{\mathcal{H}}, \frac{\alpha_{\mathcal{A}}}{\beta_{\mathcal{A}}} = \frac{\alpha_{\mathcal{H}}}{\beta_{\mathcal{H}}}\}$, where for the latter we have derived in Theorem 1(a) and in Proposition 5 the exact recursive values for

the time-behaviour of the Hellinger integrals $H_\lambda(P_{A,1}||P_{H,1})$ of order $\lambda \in \mathbb{R} \setminus [0, 1]$. Moreover, recall that for the case $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$ we have obtained in the Sections 3.4 and 3.5 some “optimal” linear lower bounds $\phi_\lambda^L(\cdot)$ for the strictly concave function $\phi_\lambda(x) := \phi(x, \beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ on the domain $x \in [0, \infty[$; due to the monotonicity Properties 2 (P10) to (P12) of the sequences $\left(a_n^{(q_\lambda^L)}\right)_{n \in \mathbb{N}}$ and $\left(b_n^{(p_\lambda^L, q_\lambda^L)}\right)_{n \in \mathbb{N}}$, these bounds have led to the “optimal” recursive lower bound $B_{\lambda, X_0, n}^L$ of the Hellinger integral $H_\lambda(P_{A,n}||P_{H,n})$ in (40) of Theorem 1(b)).

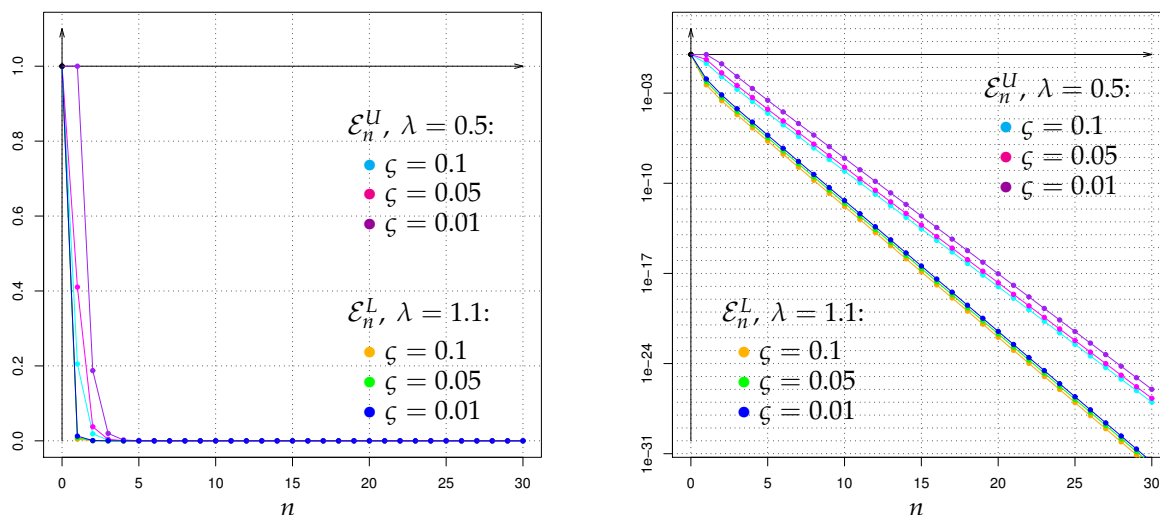


Figure 3. The lower bound \mathcal{E}_n^L (using $\lambda = 1.1$) and the upper bound \mathcal{E}_n^U (using $\lambda = 0.5$) of the minimal type II error probability $\mathcal{E}_\zeta(P_{A,n}||P_{H,n})$ for different levels $\zeta \in \{0.01, 0.05, 0.1\}$ in the parameter setup $(\beta_A, \beta_H, \alpha_A, \alpha_H) = (0.3, 1.2, 1, 4) \in \mathcal{P}_{SP,1}$ together with initial population $X_0 = 5$ on both a unit scale (left graph) and a logarithmic scale (right graph).

In contrast, the strict convexity of the function $\phi_\lambda(\cdot)$ in the case $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0, 1])$ implies that we cannot maximize both parameters $p_\lambda^L, q_\lambda^L \in \mathbb{R}$ simultaneously subject to the constraint (35). This effect carries over to the lower bounds $B_{\lambda, X_0, n}^L$ of the Hellinger integrals $H_\lambda(P_{A,n}||P_{H,n})$ (cf. (41)); in general, these bounds cannot be maximized simultaneously for all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$.

Analogously to (46), one way to obtain “good” recursive lower bounds for $H_\lambda(P_{A,n}||P_{H,n})$ from (41) in Theorem 1 (b) is to solve the optimization problem,

$$\left(\overline{p_\lambda^L}, \overline{q_\lambda^L}\right) := \arg \max_{(p_\lambda^L, q_\lambda^L) \in \mathbb{R}^2} \left\{ \exp \left\{ a_n^{(q_\lambda^L)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^L, q_\lambda^L)} \right\} \right\} \quad \text{such that (35) is satisfied,} \quad (55)$$

for each fixed initial population size $X_0 \in \mathbb{N}$ and observation horizon $n \in \mathbb{N}$. But due to the same reasons as explained right after (46), the optimization problem (55) seems to be not straightforward to solve explicitly. In a congeneric way as in the discussion of the upper bounds for the case $\lambda \in]0, 1[$ above, we now have to look for suitable parameters p_λ^L, q_λ^L for the lower bound $B_{\lambda, X_0, n}^L \leq H_\lambda(P_{A,n}||P_{H,n})$ that fulfill (35) and that guarantee certain reasonable criteria and goals; these are similar to the goals (G1) to (G3) from Section 3.6, and are therefore supplemented by an additional “'”:

- (G1') the validity of $B_{\lambda, X_0, n}^L > 1$ simultaneously for all initial configurations $X_0 \in \mathbb{N}$, all observation horizons $n \in \mathbb{N}$ and all $\lambda \in \mathbb{R} \setminus [0, 1]$, which leads to a strict improvement of the general upper bound $H_\lambda(P_{A,n}||P_{H,n}) > 1$ (cf. (11));

- (G2') the determination of the long-term-limits $\lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ respectively $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L$ for all $X_0 \in \mathbb{N}$ and all $\lambda \in \mathbb{R} \setminus [0, 1]$; in particular, one would like to check whether $\lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = \infty$;
- (G3') the determination of the time-asymptotical growth rates $\lim_{n \rightarrow \infty} \frac{1}{n} \log (H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n}))$ resp. $\lim_{n \rightarrow \infty} \frac{1}{n} \log (B_{\lambda, X_0, n}^L)$ for all $X_0 \in \mathbb{N}$ and all $\lambda \in \mathbb{R} \setminus [0, 1]$.

In the following, let us briefly discuss how these three goals can be achieved in principle, where we confine ourselves to parameters p_λ^L, q_λ^L which—in addition to (35)—fulfill the requirement

$$\left\{ q_\lambda^L \geq \max\{0, \beta_\lambda\} \quad \wedge \quad p_\lambda^L > \max\{0, \alpha_\lambda\} \right\} \quad \vee \quad \left\{ q_\lambda^L > \max\{0, \beta_\lambda\} \quad \wedge \quad p_\lambda^L \geq \max\{0, \alpha_\lambda\} \right\}, \quad (56)$$

where \wedge is the logical “AND” and \vee the logical “OR” operator. This is sufficient to tackle all three Goals (G1') to (G3'). To see this, assume that p_λ^L, q_λ^L satisfy (35). Let us begin with the two “extremal” cases in (56), i.e., with (i) $q_\lambda^L = \max\{0, \beta_\lambda\}, p_\lambda^L > \max\{0, \alpha_\lambda\}$, respectively (ii) $q_\lambda^L > \max\{0, \beta_\lambda\}, p_\lambda^L = \max\{0, \alpha_\lambda\}$.

Suppose in the first extremal case (i) that $\beta_\lambda \leq 0$. Then, $q_\lambda^L = 0$ and Properties 1 (P4) implies that $a_n^{(q_\lambda^L)} = -\beta_\lambda \geq 0$ and hence $b_n^{(p_\lambda^L, q_\lambda^L)} = p_\lambda^L e^{-\beta_\lambda} - \alpha_\lambda \geq p_\lambda^L - \alpha_\lambda > 0$ for all $n \in \mathbb{N}$. This enters into (41) as follows: the Hellinger integral lower bound becomes $B_{\lambda, X_0, n}^L \geq \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} = \exp\{-\beta_\lambda \cdot X_0 + (p_\lambda^L e^{-\beta_\lambda} - \alpha_\lambda) \cdot n\} > 1$. Furthermore, one clearly has $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L = \infty$ as well as $\lim_{n \rightarrow \infty} \frac{1}{n} \log (B_{\lambda, X_0, n}^L) = p_\lambda^L e^{-\beta_\lambda} - \alpha_\lambda > 0$. Assume now that $\beta_\lambda > 0$. Then, $q_\lambda^L = \beta_\lambda > 0, a_n^{(q_\lambda^L)} = 0$ (cf. (P2)), $b_n^{(p_\lambda^L, q_\lambda^L)} = p_\lambda^L - \alpha_\lambda > 0$ and thus $B_{\lambda, X_0, n}^L = \exp\{(p_\lambda^L - \alpha_\lambda) \cdot n\} > 1$ for all $n \in \mathbb{N}$. Furthermore, one gets $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L = \infty$ as well as $\lim_{n \rightarrow \infty} \frac{1}{n} \log (B_{\lambda, X_0, n}^L) = p_\lambda^L - \alpha_\lambda > 0$.

Let us consider the other above-mentioned extremal case (ii). Suppose that $q_\lambda^L > \max\{0, \beta_\lambda\}$ together with $q_\lambda^L > \min\{1, e^{\beta_\lambda - 1}\}$ which implies that the sequence $\left(a_n^{(q_\lambda^L)}\right)_{n \in \mathbb{N}}$ is strictly positive, strictly increasing and grows to infinity faster than exponentially, cf. (P3b). Hence, $B_{\lambda, X_0, n}^L \geq \exp\{a_n^{(q_\lambda^L)} \cdot X_0\} > 1, \lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L = \infty$ as well as $\lim_{n \rightarrow \infty} \frac{1}{n} \log (B_{\lambda, X_0, n}^L) = \infty$. If $\max\{0, \beta_\lambda\} < q_\lambda^L \leq \min\{1, e^{\beta_\lambda - 1}\}$, then $\left(a_n^{(q_\lambda^L)}\right)_{n \in \mathbb{N}}$ is strictly positive, strictly increasing and converges to $x_0^{(q_\lambda^L)} \in]0, -\log(q_\lambda^L)]$ (cf. (P3a)). This carries over to the sequence $\left(b_n^{(p_\lambda^L, q_\lambda^L)}\right)_{n \in \mathbb{N}}$: one gets $b_1^{(p_\lambda^L, q_\lambda^L)} = p_\lambda^L - \alpha_\lambda \geq 0$ and $b_n^{(p_\lambda^L, q_\lambda^L)} > 0$ for all $n \geq 2$. Furthermore, $b_n^{(p_\lambda^L, q_\lambda^L)}$ is strictly increasing and converges to $p_\lambda^L \cdot e^{x_0^{(q_\lambda^L)}} - \alpha_\lambda > 0$, leading to $B_{\lambda, X_0, n}^L > 1$ for all $n \in \mathbb{N}$, to $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L = \infty$ as well as to $\lim_{n \rightarrow \infty} \frac{1}{n} \log (B_{\lambda, X_0, n}^L) = p_\lambda^L \cdot e^{x_0^{(q_\lambda^L)}} - \alpha_\lambda > 0$.

It remains to look at the cases where p_λ^L, q_λ^L satisfy (35), and (56) with two strict inequalities. For this situation, one gets

- $\left(a_n^{(q_\lambda^L)}\right)_{n \in \mathbb{N}}$ is strictly positive, strictly increasing and—iff $q_\lambda^L \leq \min\{1, e^{\beta_\lambda - 1}\}$ —convergent (namely to the smallest positive solution $x_0^{(q_\lambda^L)} \in]0, -\log(q_\lambda^L)]$ of (44)), cf. (P3);
- $\left(b_n^{(p_\lambda^L, q_\lambda^L)}\right)_{n \in \mathbb{N}}$ is strictly increasing, strictly positive (since $b_1^{(p_\lambda^L, q_\lambda^L)} = p_\lambda^L - \alpha_\lambda > 0$) and—iff $q_\lambda^L \leq \min\{1, e^{\beta_\lambda - 1}\}$ —convergent (namely to $p_\lambda^L e^{x_0^{(q_\lambda^L)}} - \alpha_\lambda \in [p_\lambda^L - \alpha_\lambda, p_\lambda^L / q_\lambda^L - \alpha_\lambda]$), cf (P7).

Hence, under the assumptions (35) and $(p_\lambda^L > \max\{0, \alpha_\lambda\}) \wedge (q_\lambda^L > \max\{0, \beta_\lambda\})$ the corresponding lower bounds $B_{\lambda, X_0, n}^L$ of the Hellinger integral $H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ fulfill for all $X_0 \in \mathbb{N}$

- $B_{\lambda, X_0, n}^L > 1$ for all $n \in \mathbb{N}$,

- $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L = \infty,$
- $\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(B_{\lambda, X_0, n}^L \right) = p_{\lambda}^L e^{x_0^{(q_{\lambda}^L)}} - \alpha_{\lambda} > 0$ for the case $q_{\lambda}^L \in] \max\{0, \beta_{\lambda}\}, \min\{1, e^{\beta_{\lambda}-1}\}],$ respectively $\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(B_{\lambda, X_0, n}^L \right) = \infty$ for the remaining case $q_{\lambda}^L > \min\{1, e^{\beta_{\lambda}-1}\}.$

Putting these considerations together we conclude that the constraints (35) and (56) are sufficient to achieve the Goals (G1') to (G3'). Hence, for fixed parameter constellation $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda),$ we aim for finding $p_{\lambda}^L = p^L(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda)$ and $q_{\lambda}^L = q^L(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda)$ which satisfy (35) and (56). This can be achieved mostly, but not always, as we shall show below. As an auxiliary step for further investigations, it is useful to examine the set of all $\lambda \in \mathbb{R} \setminus [0, 1]$ for which $\alpha_{\lambda} \leq 0$ or $\beta_{\lambda} \leq 0$ (or both). By straightforward calculations, we see that

$$\alpha_{\lambda} \leq 0 \iff \lambda \begin{cases} \leq \frac{-\alpha_{\mathcal{H}}}{\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}}}, & \text{if } \alpha_{\mathcal{A}} > \alpha_{\mathcal{H}}, \\ \geq \frac{\alpha_{\mathcal{H}}}{\alpha_{\mathcal{H}} - \alpha_{\mathcal{A}}}, & \text{if } \alpha_{\mathcal{A}} < \alpha_{\mathcal{H}}, \end{cases} \quad \text{and} \quad \beta_{\lambda} \leq 0 \iff \lambda \begin{cases} \leq \frac{-\beta_{\mathcal{H}}}{\beta_{\mathcal{A}} - \beta_{\mathcal{H}}}, & \text{if } \beta_{\mathcal{A}} > \beta_{\mathcal{H}}, \\ \geq \frac{\beta_{\mathcal{H}}}{\beta_{\mathcal{H}} - \beta_{\mathcal{A}}}, & \text{if } \beta_{\mathcal{A}} < \beta_{\mathcal{H}}. \end{cases} \tag{57}$$

Furthermore, recall that (35) implies the general bounds $p_{\lambda}^L \leq \alpha_{\mathcal{A}}^{\lambda} \alpha_{\mathcal{H}}^{1-\lambda} = \varphi_{\lambda}(0)$ (being equivalent to the requirement $\phi_{\lambda}^L(0) = \phi_{\lambda}(0)$) and $q_{\lambda}^L \leq \beta_{\mathcal{A}}^{\lambda} \beta_{\mathcal{H}}^{1-\lambda} = \tilde{q}_{\lambda}$ (the latter being the maximal slope due to Properties 3 (P19), (P20)).

Let us now undertake the desired *detailed* investigations on lower and upper bounds of the Hellinger integrals $H_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ of order $\lambda \in \mathbb{R} \setminus [0, 1],$ for the various different subclasses of $\mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}.$

3.17. Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},2} \times (\mathbb{R} \setminus [0, 1])$

In such a constellation, where $\mathcal{P}_{\text{SP},2} := \{ (\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{\text{SP}} : \alpha_{\mathcal{A}} = \alpha_{\mathcal{H}}, \beta_{\mathcal{A}} \neq \beta_{\mathcal{H}} \}$ (cf. (49)), one gets $\phi_{\lambda}(0) = 0$ (cf. Properties 3 (P16)), $\phi'_{\lambda}(0) = 0$ (cf. (P17)). Thus, the only choice for the intercept and the slope of the linear lower bound $\phi_{\lambda}^L(\cdot)$ for $\phi_{\lambda}(\cdot),$ which satisfies (35) for all $x \in \mathbb{N}$ and (potentially) (56), is $r_{\lambda}^L = 0 = p_{\lambda}^L - \alpha_{\lambda}$ (i.e., $p_{\lambda}^L = \alpha_{\lambda} = \alpha > 0$) and $s_{\lambda}^L = \frac{\phi_{\lambda}(1) - \phi_{\lambda}(0)}{1 - 0} = q_{\lambda}^L - \beta_{\lambda} = a_1^{(q_{\lambda}^L)} > 0$ (i.e., $q_{\lambda}^L = (\alpha + \beta_{\mathcal{A}})^{\lambda} (\alpha + \beta_{\mathcal{H}})^{1-\lambda} - \alpha$). However, since $p_{\lambda}^L = \alpha_{\lambda} = \alpha > 0,$ the restriction (56) is fulfilled iff $q_{\lambda}^L > 0,$ which is equivalent to

$$\lambda \in \mathcal{I}_{\text{SP},2} := \begin{cases} \left] \frac{\log\left(\frac{\alpha}{\alpha + \beta_{\mathcal{H}}}\right)}{\log\left(\frac{\alpha + \beta_{\mathcal{A}}}{\alpha + \beta_{\mathcal{H}}}\right)}, 0 \left[\cup \right] 1, \infty \left[, & \text{if } \beta_{\mathcal{A}} > \beta_{\mathcal{H}}, \\ \right] - \infty, 0 \left[\cup \right] 1, \frac{\log\left(\frac{\alpha}{\alpha + \beta_{\mathcal{H}}}\right)}{\log\left(\frac{\alpha + \beta_{\mathcal{A}}}{\alpha + \beta_{\mathcal{H}}}\right)} \left[, & \text{if } \beta_{\mathcal{A}} < \beta_{\mathcal{H}}. \end{cases} \tag{58}$$

Suppose that $\lambda \in \mathcal{I}_{\text{SP},2}.$ As we have seen above, from Properties 1 (P3a) and (P3b) one can derive that $\left(a_n^{(q_{\lambda}^L)} \right)_{n \in \mathbb{N}}$ is strictly positive, strictly increasing, and converges to $x_0^{(q_{\lambda}^L)} \in]0, -\log(q_{\lambda}^L)]$ iff $q_{\lambda}^L \leq \min\{1, e^{\beta_{\lambda}-1}\},$ and otherwise it diverges to $\infty.$ Notice that both cases can occur: consider the parameter setup $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) = (1.5, 0.5, 0.5, 0.5) \in \mathcal{P}_{\text{SP},2},$ which leads to $\mathcal{I}_{\text{SP},2} =] - 1, 0[\cup] 1, \infty[;$ within our running-example epidemiological context of Section 2.3, this corresponds to a “mild” infectious-disease-transmission situation (\mathcal{H}) (with “low” reproduction number $\beta_{\mathcal{H}} = 0.5$ and importation mean of $\alpha_{\mathcal{H}} = 0.5$), whereas (\mathcal{A}) describes a “dangerous” situation (with supercritical $\beta_{\mathcal{A}} = 1.5$ and $\alpha_{\mathcal{A}} = 0.5$). For $\lambda = -0.5 \in \mathcal{I}_{\text{SP},2}$ one obtains $q_{\lambda}^L \approx 0.207 \leq \min\{1, e^{\beta_{\lambda}-1}\} \approx 0.368,$ whereas for $\lambda = 2 \in \mathcal{I}_{\text{SP},2}$ one gets $q_{\lambda}^L = 3.5 > \min\{1, e^{\beta_{\lambda}-1}\} = 1.$ Altogether, this leads to

Proposition 11. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP2} \times \mathcal{I}_{SP2}$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $p_\lambda^L = \alpha_A = \alpha_H = \alpha$, $q_\lambda^L = (\alpha + \beta_A)^\lambda (\alpha + \beta_H)^{1-\lambda} - \alpha$

- (a) $B_{\lambda, X_0, 1}^L = \tilde{B}_{\lambda, X_0, 1}^{(p_\lambda^L, q_\lambda^L)} = \exp \left\{ (q_\lambda^L - \beta_\lambda) \cdot X_0 \right\} > 1,$
- (b) the sequence $\left(B_{\lambda, X_0, n}^L \right)_{n \in \mathbb{N}}$ of lower bounds for $H_\lambda(P_{A,n} || P_{H,n})$ given by

$$B_{\lambda, X_0, n}^L = \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} = \exp \left\{ a_n^{(q_\lambda^L)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^L, q_\lambda^L)} \right\}$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L = \infty = \lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}),$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log B_{\lambda, X_0, n}^L = \begin{cases} p_\lambda^L \cdot \exp \left\{ x_0^{(q_\lambda^L)} \right\} - \alpha > 0, & \text{if } q_\lambda^L \leq \min \{1, e^{\beta_\lambda - 1}\}, \\ \infty, & \text{if } q_\lambda^L > \min \{1, e^{\beta_\lambda - 1}\}, \end{cases}$
- (e) the map $X_0 \mapsto B_{\lambda, X_0, n}^L = \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)}$ is strictly increasing.

Nevertheless, for the remaining constellations $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP2} \times \mathbb{R} \setminus (\mathcal{I}_{SP2} \cup [0, 1])$, all observation time horizons $n \in \mathbb{N}$ and all initial population sizes $X_0 \in \mathbb{N}$ one can still prove

$$1 < H_\lambda(P_{A,n} || P_{H,n}) \quad \text{and} \quad \lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}) = \infty, \tag{59}$$

(i.e., the achievement of the Goals (G1'), (G2')), which is done by a conceptually different method (without involving p_λ^L, q_λ^L) in Appendix A.1.

3.18. Lower Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP3a} \times (\mathbb{R} \setminus [0, 1])$

In the current setup, where $\mathcal{P}_{SP3a} := \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H, \beta_A \neq \beta_H, \frac{\alpha_A}{\beta_A} \neq \frac{\alpha_H}{\beta_H}, \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} \in] - \infty, 0[\right\}$ (cf. (49)), we always have either $(\alpha_A > \alpha_H) \wedge (\beta_A > \beta_H)$ or $(\alpha_A < \alpha_H) \wedge (\beta_A < \beta_H)$. Furthermore, from Properties 3 (P16) we obtain $\phi_\lambda(0) > 0$. As in the case $\lambda \in]0, 1[$, the derivative $\phi'_\lambda(0)$ can assume any sign on \mathcal{P}_{SP3a} , take e.g., $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (2.2, 4.5, 1, 3, 2)$ for $\phi'_\lambda(0) < 0$, $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (2.25, 4.5, 1, 3, 2)$ for $\phi'_\lambda(0) = 0$ and $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (2.3, 4.5, 1, 3, 2)$ for $\phi'_\lambda(0) > 0$ (these parameter constellations reflect “dangerous” (A) versus “highly dangerous” (H) situations within our running-example epidemiological context of Section 2.3). Nevertheless, in all three subcases one gets $\min_{x \in \mathbb{N}_0} \phi_\lambda(x) \geq \min_{x \geq 0} \phi_\lambda(x) > 0$. Thus, there exist parameters $p_\lambda^L \in]\alpha_\lambda, \alpha_A^\lambda \alpha_H^{1-\lambda}]$ and $q_\lambda^L \in]\beta_\lambda, \beta_A^\lambda \beta_H^{1-\lambda}]$ which satisfy (35) (in particular, $p_\lambda^L - \alpha_\lambda > 0, q_\lambda^L - \beta_\lambda > 0$). We now have to look for a condition which guarantees that these parameters additionally fulfill (56); such a condition is clearly that both $\alpha_\lambda \geq 0$ and $\beta_\lambda \geq 0$ hold, which is equivalent (cf. (57)) with

$$\lambda \in \mathcal{I}_{SP3a}^{(\geq)} := \begin{cases} \left[\max \left\{ \frac{-\alpha_H}{\alpha_A - \alpha_H}, \frac{-\beta_H}{\beta_A - \beta_H} \right\}, 0[\cup \right] 1, \infty[, & \text{if } (\alpha_A > \alpha_H) \wedge (\beta_A > \beta_H), \\ \left[-\infty, 0[\cup \right] 1, \min \left\{ \frac{\alpha_H}{\alpha_H - \alpha_A}, \frac{\beta_H}{\beta_H - \beta_A} \right\} \right], & \text{if } (\alpha_A < \alpha_H) \wedge (\beta_A < \beta_H); \end{cases}$$

recall that $\alpha_\lambda = 0$ and $\beta_\lambda = 0$ cannot occur simultaneously in the current setup. If $\alpha_\lambda \leq 0$ and $\beta_\lambda \leq 0$, i.e., if

$$\lambda \in \mathcal{I}_{SP3a}^{(<)} := \begin{cases} \left[-\infty, \min \left\{ \frac{-\alpha_H}{\alpha_A - \alpha_H}; \frac{-\beta_H}{\beta_A - \beta_H} \right\} \right], & \text{if } (\alpha_A > \alpha_H) \wedge (\beta_A > \beta_H), \\ \left[\max \left\{ \frac{\alpha_H}{\alpha_H - \alpha_A}; \frac{\beta_H}{\beta_H - \beta_A} \right\}, \infty[, & \text{if } (\alpha_A < \alpha_H) \wedge (\beta_A < \beta_H), \end{cases}$$

then—due to the strict positivity of the function $\varphi_\lambda(\cdot)$ (cf. (31))—there exist parameters $p_\lambda^L > 0 = \max\{0, \alpha_\lambda\}$ and $q_\lambda^L > 0 = \max\{0, \beta_\lambda\}$ which satisfy (56) and (34) (where the latter implies (35) and thus $p_\lambda^L \leq \alpha_A^\lambda \alpha_H^{1-\lambda}$, $q_\lambda^L \leq \beta_A^\lambda \beta_H^{1-\lambda}$). With

$$\mathcal{I}_{SP,3a} := \mathcal{I}_{SP,3a}^{(\geq)} \cup \mathcal{I}_{SP,3a}^{(<)} \tag{60}$$

and with the discussion below (56), we thus derive the following

Proposition 12. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,3a} \times \mathcal{I}_{SP,3a}$ there exist parameters p_λ^L, q_λ^L which satisfy $\max\{0, \alpha_\lambda\} < p_\lambda^L \leq \alpha_A^\lambda \alpha_H^{1-\lambda}$, $\max\{0, \beta_\lambda\} < q_\lambda^L \leq \beta_A^\lambda \beta_H^{1-\lambda}$ as well as (35) for all $x \in \mathbb{N}_0$, and for all such pairs $(p_\lambda^L, q_\lambda^L)$ and all initial population sizes $X_0 \in \mathbb{N}$ one gets

- (a) $B_{\lambda, X_0, 1}^L = \tilde{B}_{\lambda, X_0, 1}^{(p_\lambda^L, q_\lambda^L)} = \exp\left\{(q_\lambda^L - \beta_\lambda) \cdot X_0 + p_\lambda^L - \alpha_\lambda\right\} > 1,$
- (b) the sequence $\left(B_{\lambda, X_0, n}^L\right)_{n \in \mathbb{N}}$ of lower bounds for $H_\lambda(P_{A,n} || P_{H,n})$ given by

$$B_{\lambda, X_0, n}^L = \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} = \exp\left\{a_n^{(q_\lambda^L)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^L, q_\lambda^L)}\right\}$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^L = \infty = \lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}),$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log B_{\lambda, X_0, n}^L = \begin{cases} p_\lambda^L \cdot \exp\left\{x_0^{(q_\lambda^L)}\right\} - \alpha_\lambda > 0, & \text{if } q_\lambda^L \leq \min\{1, e^{\beta_\lambda - 1}\}, \\ \infty, & \text{if } q_\lambda^L > \min\{1, e^{\beta_\lambda - 1}\}, \end{cases}$
- (e) the map $X_0 \mapsto B_{\lambda, X_0, n}^L = \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)}$ is strictly increasing.

Notice that the assertions (a) to (e) of Proposition 12 hold true for parameter pairs $(p_\lambda^L, q_\lambda^L)$ whenever they satisfy (35) and (56); in particular, we may allow either $p_\lambda^L = \max\{0, \alpha_\lambda\}$ or $q_\lambda^L = \max\{0, \beta_\lambda\}$. Let us furthermore mention that in part (d) both asymptotical behaviours can occur: consider e.g., the parameter setup $(\beta_A, \beta_H, \alpha_A, \alpha_H) = (0.3, 0.2, 4, 3) \in \mathcal{P}_{SP,3a}$, leading to $]1, \infty[\subsetneq \mathcal{I}_{SP,3a}^{(\geq)} \subsetneq \mathcal{I}_{SP,3a}$. For $\lambda = 2 \in \mathcal{I}_{SP,3a}$, the parameters $p_\lambda^L := \tilde{p}_\lambda := 5.25$, $q_\lambda^L := \tilde{q}_\lambda := 0.45$ (corresponding to the asymptote $\tilde{\varphi}_\lambda(\cdot)$, cf. (P20)) fulfill (35), (56) and additionally $q_\lambda^L = 0.45 < \min\{1, e^{\beta_\lambda - 1}\} \approx 0.549$. Analogously, in the setup $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (3, 2, 4, 3, 2) \in \mathcal{P}_{SP,3a} \times \mathcal{I}_{SP,3a}$, the choices $p_\lambda^L := \tilde{p}_\lambda := 5.25$, $q_\lambda^L := \tilde{q}_\lambda := 4.5$ satisfy (35), (56) and there holds $q_\lambda^L = 4.5 > \min\{1, e^{\beta_\lambda - 1}\} = 1$.

For the remaining two cases $(\alpha_\lambda \leq 0) \wedge (\beta_\lambda > 0)$ (e.g., $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (6, 5, 3, 2, -3)$) and $(\alpha_\lambda > 0) \wedge (\beta_\lambda \leq 0)$ (e.g., $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) = (3, 2, 6, 5, -3)$), one has to proceed differently. Indeed, for all parameter constellations $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,3a} \times \mathbb{R} \setminus (\mathcal{I}_{SP,3a} \cup [0, 1])$, all observation time horizons $n \in \mathbb{N}$ and all initial population sizes $X_0 \in \mathbb{N}$ one can still prove

$$1 < H_\lambda(P_{A,n} || P_{H,n}), \quad \text{and} \quad \lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}) = \infty, \tag{61}$$

which is done in Appendix A.1, using a similar method as in the proof of assertion (59).

3.19. Lower Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,3b} \times (\mathbb{R} \setminus [0, 1])$

Within such a constellation, where $\mathcal{P}_{SP,3b} := \left\{ (\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H, \beta_A \neq \beta_H, \frac{\alpha_A}{\beta_A} \neq \frac{\alpha_H}{\beta_H}, \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} \in]0, \infty[\setminus \mathbb{N} \right\}$ (cf. (49)), one always has either $(\alpha_A < \alpha_H) \wedge (\beta_A > \beta_H)$ or $(\alpha_A > \alpha_H) \wedge (\beta_A < \beta_H)$. Moreover, from Properties 3 (P15) one can see that $\phi_\lambda(x) = 0$ for $x = x^* = \frac{\alpha_H - \alpha_A}{\beta_A - \beta_H} > 0$. However, $x^* \notin \mathbb{N}_0$, which implies $\phi_\lambda(x) > 0$ for all x on the relevant subdomain \mathbb{N}_0 .

Again, we incorporate (57) and consider the set of all $\lambda \in \mathbb{R} \setminus [0, 1]$ such that $\alpha_\lambda \geq 0$ and $\beta_\lambda \geq 0$ (where $\alpha_\lambda = 0 \wedge \beta_\lambda = 0$ cannot appear), i.e.,

$$\lambda \in \mathcal{I}_{SP,3b}^{(\geq)} := \begin{cases} \left[\frac{-\beta_{\mathcal{H}}}{\beta_{\mathcal{A}} - \beta_{\mathcal{H}}}, 0 \left[\cup \right] 1, \frac{\alpha_{\mathcal{H}}}{\alpha_{\mathcal{H}} - \alpha_{\mathcal{A}}} \right], & \text{if } (\alpha_{\mathcal{A}} < \alpha_{\mathcal{H}}) \wedge (\beta_{\mathcal{A}} > \beta_{\mathcal{H}}), \\ \left[\frac{-\alpha_{\mathcal{H}}}{\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}}}, 0 \left[\cup \right] 1, \frac{\beta_{\mathcal{H}}}{\beta_{\mathcal{H}} - \beta_{\mathcal{A}}} \right], & \text{if } (\alpha_{\mathcal{A}} > \alpha_{\mathcal{H}}) \wedge (\beta_{\mathcal{A}} < \beta_{\mathcal{H}}). \end{cases} \tag{62}$$

As above in Section 3.18, if $\lambda \in \mathcal{I}_{SP,3b}^{(\geq)}$ then there exist parameters $p_\lambda^L \in]\alpha_\lambda, \alpha_{\mathcal{A}}^\lambda \alpha_{\mathcal{H}}^{1-\lambda}]$, $q_\lambda^L \in]\beta_\lambda, \beta_{\mathcal{A}}^\lambda \beta_{\mathcal{H}}^{1-\lambda}]$ (which thus fulfill (56)) such that (35) is satisfied for all $x \in \mathbb{N}_0$. Hence, for all $\lambda \in \mathcal{I}_{SP,3b} := \mathcal{I}_{SP,3b}^{(\geq)}$, all assertions (a) to (e) of Proposition 12 hold true. Notice that for the current setup $\mathcal{P}_{SP,3b}$ one cannot have $\alpha_\lambda \leq 0$ and $\beta_\lambda \leq 0$ simultaneously. Furthermore, in each of the two remaining cases $(\alpha_\lambda < 0) \wedge (\beta_\lambda > 0)$ respectively $(\alpha_\lambda > 0) \wedge (\beta_\lambda < 0)$ it can happen that there do not exist parameters $p_\lambda^L, q_\lambda^L > 0$ which satisfy both (35) and (56). However, as in the case $\mathcal{P}_{SP,3a}$ above, for all $\lambda \notin \mathcal{I}_{SP,3b}$ we prove in Appendix A.1 (by a method without p_λ^L, q_λ^L) that for all observation times $n \in \mathbb{N}$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds

$$1 < H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n}) \quad \text{and} \quad \lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = \infty. \tag{63}$$

3.20. Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{SP,3c} \times (\mathbb{R} \setminus [0, 1])$

Since in this subcase one has $\mathcal{P}_{SP,3c} := \left\{ (\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{SP} : \alpha_{\mathcal{A}} \neq \alpha_{\mathcal{H}}, \beta_{\mathcal{A}} \neq \beta_{\mathcal{H}}, \frac{\alpha_{\mathcal{A}}}{\beta_{\mathcal{A}}} \neq \frac{\alpha_{\mathcal{H}}}{\beta_{\mathcal{H}}}, \frac{\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}}}{\beta_{\mathcal{H}} - \beta_{\mathcal{A}}} \in \mathbb{N} \right\}$ (cf. (49)) and thus $\phi_\lambda(x^*) = 0$ for $x^* \in \mathbb{N}$, there do not exist parameters p_λ^L, q_λ^L such that (35) and (56) are satisfied. The only parameter pair that ensures $\exp \left\{ a_n^{(q_\lambda^L)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^L, q_\lambda^L)} \right\} \geq 1$ for all $n \in \mathbb{N}$ and all $X_0 \in \mathbb{N}$ within our proposed method, is the choice $p_\lambda^L = \alpha_\lambda, q_\lambda^L = \beta_\lambda$. Consequently, $B_{\lambda, X_0, n}^L \equiv 1$, which coincides with the general lower bound (11) but violates the above-mentioned desired Goal (G1'). However, in some constellations there exist *nonnegative* parameters $p_\lambda^L < \alpha_\lambda, q_\lambda^L > \beta_\lambda$ or $p_\lambda^L > \alpha_\lambda, q_\lambda^L < \beta_\lambda$, such that at least the parts (c) and (d) of Proposition 12 are satisfied. As in Section 3.19 above, by using a conceptually different method (without p_λ^L, q_λ^L) we prove in Appendix A.1 that for all $\lambda \in \mathbb{R} \setminus [0, 1]$, all observation times $n \in \mathbb{N}$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds

$$1 < H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n}) \quad \text{and} \quad \lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = \infty. \tag{64}$$

3.21. Lower Bounds for the Cases $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{SP,4a} \times (\mathbb{R} \setminus [0, 1])$

In the current setup, where $\mathcal{P}_{SP,4a} := \{ (\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{SP} : \alpha_{\mathcal{A}} \neq \alpha_{\mathcal{H}} > 0, \beta_{\mathcal{A}} = \beta_{\mathcal{H}} \in]0, 1[\}$ (cf. (49)), the function $\phi_\lambda(\cdot)$ is strictly positive and strictly decreasing, with $\lim_{x \rightarrow \infty} \phi_\lambda(x) = \lim_{x \rightarrow \infty} \phi'_\lambda(x) = 0$. The only choice of parameters p_λ^L, q_λ^L which fulfill (35) and $\exp \left\{ a_n^{(q_\lambda^L)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^L, q_\lambda^L)} \right\} \geq 1$ for all $n \in \mathbb{N}$ and all $X_0 \in \mathbb{N}$, is the choice $p_\lambda^L = \alpha_\lambda$ as well as $q_\lambda^L = \beta_\lambda = \beta_\bullet$, where β_\bullet stands for both (equal) $\beta_{\mathcal{H}}$ and $\beta_{\mathcal{A}}$. Of course, this leads to $B_{\lambda, X_0, n}^L \equiv 1$, which is consistent with the general lower bound (11), but violates the above-mentioned desired Goal (G1'). Nevertheless, in Appendix A.1 we prove the following

Proposition 13. For all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{SP,4a} \times \mathbb{R} \setminus [0, 1]$ there exist parameters $p_\lambda^L > \alpha_\lambda$ (not necessarily satisfying $p_\lambda^L \geq 0$) and $0 < q_\lambda^L < \beta_\lambda = \beta_\bullet < \min\{1, e^{\beta_\bullet - 1}\} = e^{\beta_\bullet - 1}$ such that (35) holds for all $x \in [0, \infty[$ and such that for all initial population sizes $X_0 \in \mathbb{N}$ the parts (c) and (d) of Proposition 12 hold true.

3.22. Lower Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4b} \times (\mathbb{R} \setminus [0, 1])$

By recalling $\mathcal{P}_{SP,4b} := \{(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} : \alpha_A \neq \alpha_H > 0, \beta_A = \beta_H \in [1, \infty[\}$ (cf.(49)), the assertions preceding Proposition 13 remain valid. However, the proof of Proposition 13 in Appendix A.1 contains details which explain why it cannot be carried over to the current case $\mathcal{P}_{SP,4b}$. Thus, the generally valid lower bound $B_{\lambda, X_0, n}^L \equiv 1$ cannot be improved with our methods.

3.23. Concluding Remarks on Alternative Lower Bounds for all Cases

$(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0, 1])$

To achieve the Goals (G1') to (G3'), in the above-mentioned investigations about lower bounds of the Hellinger integral $H_\lambda(P_{A,n} || P_{H,n}), \lambda \in \mathbb{R} \setminus [0, 1]$, we have mainly focused on parameters p_λ^L, q_λ^L which satisfy (35) and additionally (56). Nevertheless, Theorem 1 (b) gives lower bounds $B_{\lambda, X_0, n}^L$ whenever (35) is fulfilled. However, this lower bound can be the trivial one, $B_{\lambda, X_0, n}^L \equiv 1$. Let us remark here that for the parameter constellations $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,2} \times \mathbb{R} \setminus ([0, 1] \cup \mathcal{I}_{SP,2})) \cup (\mathcal{P}_{SP,3a} \times \mathbb{R} \setminus ([0, 1] \cup \mathcal{I}_{SP,3a})) \cup (\mathcal{P}_{SP,3b} \times \mathbb{R} \setminus ([0, 1] \cup \mathcal{I}_{SP,3b}))$ one can prove that there exist p_λ^L, q_λ^L which satisfy (35) for all $x \in \mathbb{N}_0$ as well as the condition (generalizing (56))

$$p_\lambda^L \geq \alpha_\lambda, \quad q_\lambda^L \geq \beta_\lambda, \quad (\text{where at least one of the inequalities is strict}),$$

and that for such p_λ^L, q_λ^L one gets the validity of $H_\lambda(P_{A,n} || P_{H,n}) \geq B_{\lambda, X_0, n}^L = \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} > 1$ for all $X_0 \in \mathbb{N}$ and all $n \in \mathbb{N}$; consequently, Goal (G1') is achieved. However, in these parameter constellations it can unpleasantly happen that $n \mapsto B_{\lambda, X_0, n}^L$ is oscillating (in contrast to the monotone behaviour in the Propositions 11 (b), 12 (b)).

As a final general remark, let us mention that the functions $\phi_{\lambda, y}^{\tan}(\cdot), \phi_{\lambda, k}^{\sec}(\cdot), \phi_\lambda^{\text{hor}}(\cdot), \tilde{\phi}_\lambda(\cdot)$ –defined in (52)–(54) and Properties 3 (P20)–constitute linear lower bounds for $\phi_\lambda(\cdot)$ on the domain \mathbb{N}_0 in the case $\lambda \in \mathbb{R} \setminus [0, 1]$. Their parameters $p_\lambda^L \in \{p_{\lambda, y}^{\tan}, p_{\lambda, y}^{\sec}, p_{\lambda, y}^{\text{hor}}, \tilde{p}_\lambda\}$ and $q_\lambda^L \in \{q_{\lambda, y}^{\tan}, q_{\lambda, y}^{\sec}, q_{\lambda, y}^{\text{hor}}, \tilde{q}_\lambda\}$ lead to lower bounds $B_{\lambda, X_0, n}^L$ of the Hellinger integrals that may or may not be consistent with Goals (G1') to (G3'), and which may be possibly better respectively weaker respectively incomparable with the previous lower bounds when adding some relaxation of (G1'), such as e.g., the validity of $H_\lambda(P_{A,n} || P_{H,n}) > 1$ for all but finitely many $n \in \mathbb{N}$.

3.24. Upper Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0, 1])$

For the cases $\lambda \in \mathbb{R} \setminus [0, 1]$, the investigation of upper bounds for the Hellinger integral $H_\lambda(P_{A,n} || P_{H,n})$ is much easier than the above-mentioned derivations of lower bounds. In fact, we face a situation which is similar to the lower-bounds-studies for the cases $\lambda \in]0, 1[$: due to Properties 3 (P19), the function $\phi_\lambda(\cdot)$ is strictly convex on the nonnegative real line. Furthermore, it is asymptotically linear, as stated in (P20). The monotonicity Properties 2 (P10) to (P12) imply that for the tightest upper bound (within our framework) one should use the parameters $p_\lambda^U := \alpha_A^\lambda \alpha_H^{1-\lambda} > 0$ and $q_\lambda^U := \beta_A^\lambda \beta_H^{1-\lambda} > 0$. Lemma A1 states that $p_\lambda^U \geq \alpha_\lambda$ resp. $q_\lambda^U \geq \beta_\lambda$, with equality iff $\alpha_A = \alpha_H$ resp. iff $\beta_A = \beta_H$. From Properties 1 (P3a) we see that for $\beta_A \neq \beta_H$ the corresponding sequence $(a_n^{(q_\lambda^U)})_{n \in \mathbb{N}}$ is convergent to $x_0^{(q_\lambda^U)} \in]0, -\log(q_\lambda^U)]$ if $q_\lambda^U \leq \min\{1, e^{\beta_\lambda - 1}\}$ (i.e., if $\lambda \in [\lambda_-, \lambda_+]$, cf. Lemma 1 (a)), and otherwise it diverges to ∞ faster than exponentially (cf. (P3b)). If $\beta_A = \beta_H$ (i.e., if $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP,4} = \mathcal{P}_{SP,4a} \cup \mathcal{P}_{SP,4b}$), then one gets $q_\lambda^U = \beta_\lambda$ and $a_n^{(q_\lambda^U)} = 0 = x_0^{(q_\lambda^U)}$ for all $n \in \mathbb{N}$ (cf. (P2)). Altogether, this leads to

Proposition 14. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0, 1])$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $p_\lambda^U := \alpha_A^\lambda \alpha_H^{1-\lambda}$, $q_\lambda^U := \beta_A^\lambda \beta_H^{1-\lambda}$

(a) $B_{\lambda, X_0, 1}^U = \tilde{B}_{\lambda, X_0, 1}^{(p_\lambda^U, q_\lambda^U)} = \exp \left\{ (\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda) \cdot X_0 + \alpha_A^\lambda \alpha_H^{1-\lambda} - \alpha_\lambda \right\} > 1,$

(b) the sequence $(B_{\lambda, X_0, n}^U)_{n \in \mathbb{N}}$ of upper bounds for $H_\lambda(P_{A,n} || P_{H,n})$ given by

$$B_{\lambda, X_0, n}^U = \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} = \exp \left\{ a_n^{(q_\lambda^U)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)} \right\}$$

is strictly increasing,

(c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^U = \infty,$

(d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log B_{\lambda, X_0, n}^U = \begin{cases} p_\lambda^U \cdot \exp \left\{ x_0^{(q_\lambda^U)} \right\} - \alpha_\lambda > 0, & \text{if } \lambda \in [\lambda_-, \lambda_+] \setminus [0, 1], \\ \infty, & \text{if } \lambda \in] - \infty, \lambda_- [\cup] \lambda_+, \infty [, \end{cases}$

(e) the map $X_0 \mapsto B_{\lambda, X_0, n}^U = \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)}$ is strictly increasing.

4. Power Divergences of Non-Kullback-Leibler-Information-Divergence Type

4.1. A First Basic Result

For orders $\lambda \in \mathbb{R} \setminus \{0, 1\}$, all the results of the previous Section 3 carry correspondingly over from the Hellinger integrals $H_\lambda(\cdot || \cdot)$ to the total variation distance $V(\cdot || \cdot)$, by virtue of the relation (cf. (12))

$$2 (1 - H_{\frac{1}{2}}(P_{A,n} || P_{H,n})) \leq V(P_{A,n} || P_{H,n}) \leq 2 \sqrt{1 - (H_{\frac{1}{2}}(P_{A,n} || P_{H,n}))^2},$$

to the Renyi divergences $R_\lambda(\cdot || \cdot)$, by virtue of the relation (cf. (7))

$$0 \leq R_\lambda(P_{A,n} || P_{H,n}) = \frac{1}{\lambda(\lambda - 1)} \log H_\lambda(P_{A,n} || P_{H,n}), \quad \text{with } \log 0 := -\infty,$$

as well as to the power divergences $I_\lambda(\cdot || \cdot)$, by virtue of the relation (cf. (2))

$$I_\lambda(P_{A,n} || P_{H,n}) = \frac{1 - H_\lambda(P_{A,n} || P_{H,n})}{\lambda \cdot (1 - \lambda)}, \quad n \in \mathbb{N};$$

in the following, we concentrate on the latter. In particular, the above-mentioned carrying-over procedure leads to bounds on $I_\lambda(P_A || P_H)$ which are tighter than the general rudimentary bounds (cf. (10) and (11))

$$0 \leq I_\lambda(P_{A,n} || P_{H,n}) < \frac{1}{\lambda(1 - \lambda)}, \quad \text{for } \lambda \in]0, 1[, \quad 0 \leq I_\lambda(P_{A,n} || P_{H,n}) \leq \infty, \quad \text{for } \lambda \in \mathbb{R} \setminus [0, 1].$$

Because power divergences have a *very insightful interpretation* as “directed distances” between two probability distributions (e.g., within our running-example epidemiological context), and function as important tools in statistics, information theory, machine learning, and artificial intelligence, we present explicitly the outcoming exact values respectively bounds of $I_\lambda(P_A || P_H)$ ($\lambda \in \mathbb{R} \setminus \{0, 1\}$, $n \in \mathbb{N}$), in the current and the following subsections. For this, recall the case-dependent parameters $p^A = p_\lambda^A = p^A(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ and $q^A = q_\lambda^A = q^A(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ ($A \in \{E, L, U\}$). To begin with, we can deduce from Theorem 1

Theorem 2.

- (a) For all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1})$, all initial population sizes $X_0 \in \mathbb{N}_0$, all observation horizons $n \in \mathbb{N}$ and all $\lambda \in \mathbb{R} \setminus \{0, 1\}$ one can recursively compute the **exact value**

$$I_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n}) = \frac{1}{\lambda(\lambda - 1)} \cdot \left[\exp \left\{ a_n^{(q_{\lambda}^E)} \cdot X_0 + \frac{\alpha_{\mathcal{A}}}{\beta_{\mathcal{A}}} \sum_{k=1}^n a_k^{(q_{\lambda}^E)} \right\} - 1 \right] =: V_{\lambda, X_0, n}^I, \tag{65}$$

where $\frac{\alpha_{\mathcal{A}}}{\beta_{\mathcal{A}}}$ can be equivalently replaced by $\frac{\alpha_{\mathcal{H}}}{\beta_{\mathcal{H}}}$ and $q_{\lambda}^E := \beta_{\mathcal{A}}^{\lambda} \beta_{\mathcal{H}}^{1-\lambda}$. Notice that on \mathcal{P}_{NI} the formula (65) simplifies significantly, since $\alpha_{\mathcal{A}} = \alpha_{\mathcal{H}} = 0$.

- (b) For general parameters $p \in \mathbb{R}, q \neq 0$ recall the general expression (cf. (42))

$$\tilde{B}_{\lambda, X_0, n}^{(p,q)} := \exp \left\{ a_n^{(q)} \cdot X_0 + \frac{p}{q} \sum_{k=1}^n a_k^{(q)} + n \cdot \left(\frac{p}{q} \beta_{\lambda} - \alpha_{\lambda} \right) \right\}$$

as well as

$$\tilde{B}_{\lambda, X_0, n}^{(p,0)} := \exp \left\{ -\beta_{\lambda} \cdot X_0 + (p \cdot e^{-\beta_{\lambda}} - \alpha_{\lambda}) \cdot n \right\}.$$

Then, for all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$, all $\lambda \in \mathbb{R} \setminus \{0, 1\}$, all coefficients $p_{\lambda}^L, p_{\lambda}^U, q_{\lambda}^L, q_{\lambda}^U \in \mathbb{R}$ which satisfy (35) for all $x \in \mathbb{N}_0$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ one gets the following recursive bounds for the power divergences: for $\lambda \in]0, 1[$ there holds

$$I_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n}) \begin{cases} < \frac{1}{\lambda(1-\lambda)} \cdot (1 - B_{\lambda, X_0, n}^L) = \frac{1}{\lambda(1-\lambda)} \cdot (1 - \tilde{B}_{\lambda, X_0, n}^{(p_{\lambda}^L, q_{\lambda}^L)}) =: B_{\lambda, X_0, n}^{L,U} \\ & \geq \frac{1}{\lambda(1-\lambda)} \cdot (1 - B_{\lambda, X_0, n}^U) = \frac{1}{\lambda(1-\lambda)} \cdot (1 - \min \{ \tilde{B}_{\lambda, X_0, n}^{(p_{\lambda}^U, q_{\lambda}^U)}, 1 \}) =: B_{\lambda, X_0, n}^{L,L} \end{cases}$$

whereas for $\lambda \in \mathbb{R} \setminus [0, 1]$ there holds

$$I_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n}) \begin{cases} < \frac{1}{\lambda(\lambda-1)} \cdot (B_{\lambda, X_0, n}^U - 1) = \frac{1}{\lambda(\lambda-1)} \cdot (\tilde{B}_{\lambda, X_0, n}^{(p_{\lambda}^U, q_{\lambda}^U)} - 1) =: B_{\lambda, X_0, n}^{U,U} \\ & \geq \frac{1}{\lambda(\lambda-1)} \cdot (B_{\lambda, X_0, n}^L - 1) = \frac{1}{\lambda(\lambda-1)} \cdot (\max \{ \tilde{B}_{\lambda, X_0, n}^{(p_{\lambda}^L, q_{\lambda}^L)}, 1 \} - 1) =: B_{\lambda, X_0, n}^{U,L} \end{cases}$$

In order to deduce the subsequent *detailed* recursive analyses of power divergences, we also employ the obvious relations

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{1}{\lambda(1-\lambda)} - I_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n}) \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \left[-\log(\lambda(1-\lambda)) + \log(H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})) \right] \\ & = \lim_{n \rightarrow \infty} \frac{1}{n} \log(H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})), \quad \text{for } \lambda \in]0, 1[, \end{aligned} \tag{66}$$

as well as

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log(I_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})) = \lim_{n \rightarrow \infty} \frac{1}{n} \left[-\log(\lambda(\lambda-1)) + \log(H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n}) - 1) \right] \\ & = \lim_{n \rightarrow \infty} \frac{1}{n} \left[\log \left(1 - \frac{1}{H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})} \right) + \log(H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})) \right] = \lim_{n \rightarrow \infty} \frac{1}{n} \log(H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n})), \end{aligned} \tag{67}$$

for $\lambda \in \mathbb{R} \setminus [0, 1]$ (provided that $\liminf_{n \rightarrow \infty} H_{\lambda}(P_{\mathcal{A},n}||P_{\mathcal{H},n}) > 1$).

4.2. Detailed Analyses of the Exact Recursive Values of $I_\lambda(\cdot||\cdot)$, i.e., for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP1}) \times (\mathbb{R} \setminus \{0, 1\})$

Corollary 2. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{NI} \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $I_\lambda(P_{A,1}||P_{H,1}) = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda\right) \cdot X_0 \right\}\right) > 0,$
- (b) the sequence $(I_\lambda(P_{A,n}||P_{H,n}))_{n \in \mathbb{N}}$ given by

$$I_\lambda(P_{A,n}||P_{H,n}) = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ a_n^{(q_\lambda^E)} \cdot X_0 \right\}\right) =: V_{\lambda, X_0, n}^I$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} I_\lambda(P_{A,n}||P_{H,n}) = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ x_0^{(q_\lambda^E)} \cdot X_0 \right\}\right) \in]0, \frac{1}{\lambda(1-\lambda)}[,$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{1}{\lambda(1-\lambda)} - I_\lambda(P_{A,n}||P_{H,n}) \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log H_\lambda(P_{A,n}||P_{H,n}) = 0,$
- (e) the map $X_0 \mapsto V_{\lambda, X_0, n}^I$ is strictly increasing.

Corollary 3. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{NI} \times (\mathbb{R} \setminus]0, 1])$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $I_\lambda(P_{A,1}||P_{H,1}) = \frac{1}{\lambda(\lambda-1)} \cdot \left(\exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda\right) \cdot X_0 \right\} - 1\right) > 0,$
- (b) the sequence $(I_\lambda(P_{A,n}||P_{H,n}))_{n \in \mathbb{N}}$ given by

$$I_\lambda(P_{A,n}||P_{H,n}) = \frac{1}{\lambda(\lambda-1)} \cdot \left(\exp \left\{ a_n^{(q_\lambda^E)} \cdot X_0 \right\} - 1\right) =: V_{\lambda, X_0, n}^I$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} I_\lambda(P_{A,n}||P_{H,n}) = \begin{cases} \frac{1}{\lambda(\lambda-1)} \cdot \left(\exp \left\{ x_0^{(q_\lambda^E)} \cdot X_0 \right\} - 1\right) > 0, & \text{if } \lambda \in [\lambda_-, \lambda_+] \setminus]0, 1], \\ \infty, & \text{if } \lambda \in]-\infty, \lambda_-[\cup]\lambda_+, \infty[. \end{cases}$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log I_\lambda(P_{A,n}||P_{H,n}) = \begin{cases} 0, & \text{if } \lambda \in [\lambda_-, \lambda_+] \setminus]0, 1], \\ \infty, & \text{if } \lambda \in]-\infty, \lambda_-[\cup]\lambda_+, \infty[. \end{cases}$
- (e) the map $X_0 \mapsto V_{\lambda, X_0, n}^I$ is strictly increasing.

Corollary 4. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP1} \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $I_\lambda(P_{A,1}||P_{H,1}) = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda\right) \cdot \left(X_0 + \frac{\alpha_A}{\beta_A}\right) \right\}\right) > 0,$
- (b) the sequence $(I_\lambda(P_{A,n}||P_{H,n}))_{n \in \mathbb{N}}$ given by

$$I_\lambda(P_{A,n}||P_{H,n}) = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ a_n^{(q_\lambda^E)} \cdot X_0 + \frac{\alpha_A}{\beta_A} \sum_{k=1}^n a_k^{(q_\lambda^E)} \right\}\right) =: V_{\lambda, X_0, n}^I$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} I_\lambda(P_{A,n}||P_{H,n}) = \frac{1}{\lambda(1-\lambda)},$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{1}{\lambda(1-\lambda)} - I_\lambda(P_{A,n}||P_{H,n}) \right) = \frac{\alpha_A}{\beta_A} \cdot x_0^{(q_\lambda^E)} < 0,$
- (e) the map $X_0 \mapsto V_{\lambda, X_0, n}^I$ is strictly increasing.

Corollary 5. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,1} \times (\mathbb{R} \setminus [0, 1])$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $q_\lambda^E := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $I_\lambda(P_{A,1}||P_{H,1}) = \frac{1}{\lambda(\lambda-1)} \cdot \left(\exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda \right) \cdot \left(X_0 + \frac{\alpha_A}{\beta_A} \right) \right\} - 1 \right) > 0,$
- (b) the sequence $(I_\lambda(P_{A,n}||P_{H,n}))_{n \in \mathbb{N}}$ given by

$$I_\lambda(P_{A,n}||P_{H,n}) = \frac{1}{\lambda(\lambda-1)} \cdot \left(\exp \left\{ a_n^{(q_\lambda^E)} \cdot X_0 + \frac{\alpha_A}{\beta_A} \sum_{k=1}^n a_k^{(q_\lambda^E)} \right\} - 1 \right) =: V_{\lambda, X_0, n}^I$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} I_\lambda(P_{A,n}||P_{H,n}) = \infty,$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log I_\lambda(P_{A,n}||P_{H,n}) = \begin{cases} \frac{\alpha_A}{\beta_A} \cdot x_0^{(q_\lambda^E)} > 0, & \text{if } \lambda \in [\lambda_-, \lambda_+] \setminus [0, 1], \\ \infty, & \text{if } \lambda \in]-\infty, \lambda_-[\cup]\lambda_+, \infty[, \end{cases}$
- (e) the map $X_0 \mapsto V_{\lambda, X_0, n}^I$ is strictly increasing.

In the assertions (a), (b), (d) of the Corollaries 4 and 5 the fraction α_A/β_A can be equivalently replaced by α_H/β_H .

Let us now derive the corresponding detailed results for the bounds of the power divergences for the parameter cases $\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$, where the Hellinger integral, and thus $I_\lambda(P_{A,n}||P_{H,n})$, cannot be determined exactly. The extensive discussion on the Hellinger-integral bounds in the Sections 3.4–3.13, as well as in the Sections 3.16–3.24 can be carried over directly to obtain power-divergence bounds. In the following, we summarize the outcoming key results, referring a detailed discussion on the possible choices of $p_\lambda^A = p^A(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ and $q_\lambda^A = q^A(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ ($A \in \{L, U\}$) to the corresponding above-mentioned subsections.

4.3. Lower Bounds of $I_\lambda(\cdot||\cdot)$ for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$

Corollary 6. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,2} \cup \mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b}) \times]0, 1[$ there exist parameters p_λ^U, q_λ^U which satisfy $p_\lambda^U \in [\alpha_A^\lambda \alpha_H^{1-\lambda}, \alpha_\lambda]$ and $q_\lambda^U \in [\beta_A^\lambda \beta_H^{1-\lambda}, \beta_\lambda]$ [as well as (35) for all $x \in \mathbb{N}_0$, and for all such pairs $(p_\lambda^U, q_\lambda^U)$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds

- (a) $B_{\lambda, X_0, 1}^{I,L} = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ \left(q_\lambda^U - \beta_\lambda \right) \cdot X_0 + p_\lambda^U - \alpha_\lambda \right\} \right) > 0,$
- (b) the sequence $(B_{\lambda, X_0, n}^{I,L})_{n \in \mathbb{N}}$ of lower bounds for $I_\lambda(P_{A,n}||P_{H,n})$ given by

$$B_{\lambda, X_0, n}^{I,L} = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ a_n^{(q_\lambda^U)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^U, q_\lambda^U)} \right\} \right)$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^{I,L} = \lim_{n \rightarrow \infty} I_\lambda(P_{A,n}||P_{H,n}) = \frac{1}{\lambda(1-\lambda)},$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{1}{\lambda(1-\lambda)} - B_{\lambda, X_0, n}^{I,L} \right) = p_\lambda^U \cdot e^{x_0^{(q_\lambda^U)}} - \alpha_\lambda < 0,$
- (e) the map $X_0 \mapsto B_{\lambda, X_0, n}^{I,L}$ is strictly increasing.

Remark 4.

- (a) Notice that in the case $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,2} \times]0, 1[$ —where $\alpha_A^\lambda \alpha_H^{1-\lambda} = \alpha_\lambda = \alpha_A = \alpha_H = \alpha$ —we get the special choice $p_\lambda^U = \alpha$ and $q_\lambda^U = (\alpha + \beta_A)^\lambda (\alpha + \beta_H)^{1-\lambda} - \alpha$ (cf. Section 3.7). For the constellations $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b}) \times]0, 1[$ there exist parameters

$$p_\lambda^U \in [\alpha_A^\lambda \alpha_H^{1-\lambda}, \alpha_\lambda[, q_\lambda^U \in [\beta_A^\lambda \beta_H^{1-\lambda}, \beta_\lambda[\text{ which satisfy (35) for all } x \in \mathbb{N}_0.$$

- (b) For the parameter setups $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,2} \cup \mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b}) \times]0, 1[$ there might exist parameter pairs $(p_\lambda^U, q_\lambda^U)$ satisfying (35) and either $p_\lambda^U = \alpha_\lambda$ or $q_\lambda^U = \beta_\lambda$, for which all assertions of Corollary 6 still hold true.
- (c) Following the discussion in Section 3.10 for all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,3c} \times]0, 1[$ at least part (c) still holds true.

Corollary 7. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4a} \times]0, 1[$ there exist parameters $p_\lambda^U < \alpha_\lambda, 1 > q_\lambda^U > \beta_\lambda = \beta$ such that (35) is satisfied for all $x \in [0, \infty[$ and such that for all initial population sizes $X_0 \in \mathbb{N}$ at least the parts (c) and (d) of Corollary 6 hold true.

As in Section 3.12, for the parameter setup $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4b} \times]0, 1[$ we cannot derive a lower bound for the power divergences which improves the generally valid lower bound $I_\lambda(P_{A,n} || P_{H,n}) \geq 0$ (cf. (10)) by employing our proposed $(p_\lambda^U, q_\lambda^U)$ -method.

4.4. Upper Bounds of $I_\lambda(\cdot || \cdot)$ for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$

Since in this setup the upper bounds of the power divergences can be derived from the lower bounds of the Hellinger integrals, we here appropriately adapt the results of Proposition 6.

Corollary 8. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $p_\lambda^L := \alpha_A^\lambda \alpha_H^{1-\lambda}$ and $q_\lambda^L := \beta_A^\lambda \beta_H^{1-\lambda}$

- (a) $B_{\lambda, X_0, 1}^{I,U} = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ \left(\beta_A^\lambda \beta_H^{1-\lambda} - \beta_\lambda \right) \cdot X_0 + \alpha_A^\lambda \alpha_H^{1-\lambda} - \alpha_\lambda \right\} \right) > 0,$
- (b) the sequence of upper bounds $\left(B_{\lambda, X_0, n}^{I,U} \right)_{n \in \mathbb{N}}$ for $I_\lambda(P_{A,n} || P_{H,n})$ given by

$$B_{\lambda, X_0, n}^{I,U} = \frac{1}{\lambda(1-\lambda)} \cdot \left(1 - \exp \left\{ a_n^{(q_\lambda^L)} \cdot X_0 + \frac{p_\lambda^L}{q_\lambda^L} \sum_{k=1}^n a_k^{(q_\lambda^L)} + n \cdot \left(\frac{p_\lambda^L}{q_\lambda^L} \cdot \beta_\lambda - \alpha_\lambda \right) \right\} \right)$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^{I,U} = \frac{1}{\lambda(1-\lambda)},$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{1}{\lambda(1-\lambda)} - B_{\lambda, X_0, n}^{I,U} \right) = \frac{p_\lambda^L}{q_\lambda^L} \cdot \left(x_0^{(q_\lambda^L)} + \beta_\lambda \right) - \alpha_\lambda = p_\lambda^L \cdot e^{x_0^{(q_\lambda^L)}} - \alpha_\lambda < 0,$
- (e) the map $X_0 \mapsto B_{\lambda, X_0, n}^{I,U}$ is strictly increasing.

4.5. Lower Bounds of $I_\lambda(\cdot || \cdot)$ for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus]0, 1[)$

In order to derive detailed results on lower bounds of the power divergences in the case $\lambda \in \mathbb{R} \setminus]0, 1[$, we have to subsume and adapt the Hellinger-integral concerning lower-bounds investigations from the Sections 3.16–3.23. Recall the λ -sets $\mathcal{I}_{SP,2}, \mathcal{I}_{SP,3a}, \mathcal{I}_{SP,3b}$ (cf. (58), (60), (62)). For the constellations $\mathcal{P}_{SP,2} \times \mathcal{I}_{SP,2}$ we employ the special choice $p_\lambda^L = \alpha_A^\lambda \alpha_H^{1-\lambda} = \alpha_\lambda = \alpha_A = \alpha_H = \alpha$ together with $q_\lambda^L = (\alpha + \beta_A)^\lambda (\alpha + \beta_H)^{1-\lambda} - \alpha > \max\{0, \beta_\lambda\}$ (cf. (58)) which satisfy (35) for all $x \in \mathbb{N}_0$ and (56), whereas for the constellations $(\mathcal{P}_{SP,3a} \times \mathcal{I}_{SP,3a}) \cup (\mathcal{P}_{SP,3b} \times \mathcal{I}_{SP,3b})$ we have proved the existence of parameters p_λ^L, q_λ^L satisfying both (35) for all $x \in \mathbb{N}_0$ and (56) with two strict inequalities. Subsuming this, we obtain

Corollary 9. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,2} \times \mathcal{I}_{SP,2}) \cup (\mathcal{P}_{SP,3a} \times \mathcal{I}_{SP,3a}) \cup (\mathcal{P}_{SP,3b} \times \mathcal{I}_{SP,3b})$ there exist parameters p_λ^L, q_λ^L which satisfy $\max\{0, \alpha_\lambda\} \leq p_\lambda^L \leq \alpha_A^\lambda \alpha_H^{1-\lambda}$, $\max\{0, \beta_\lambda\} < q_\lambda^L \leq \beta_A^\lambda \beta_H^{1-\lambda}$ as well as (35) for all $x \in \mathbb{N}_0$, and for all such pairs $(p_\lambda^L, q_\lambda^L)$ and all initial population sizes $X_0 \in \mathbb{N}$ one gets

(a)
$$B_{\lambda, X_0, 1}^{I, L} = \frac{1}{\lambda(\lambda - 1)} \cdot \left(\exp \left\{ (q_\lambda^L - \beta_\lambda) \cdot X_0 + p_\lambda^L - \alpha_\lambda \right\} - 1 \right) > 0,$$

(b) the sequence $(B_{\lambda, X_0, n}^{I, L})_{n \in \mathbb{N}}$ of lower bounds for $I_\lambda(P_{A, n} || P_{H, n})$ given by

$$B_{\lambda, X_0, n}^{I, L} = \frac{1}{\lambda(\lambda - 1)} \cdot \left(\exp \left\{ a_n^{(q_\lambda^L)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda^L, q_\lambda^L)} \right\} - 1 \right)$$

is strictly increasing,

(c)
$$\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^{I, L} = \lim_{n \rightarrow \infty} I_\lambda(P_{A, n} || P_{H, n}) = \infty,$$

(d)
$$\lim_{n \rightarrow \infty} \frac{1}{n} \log B_{\lambda, X_0, n}^{I, L} = \begin{cases} p_\lambda^L \cdot \exp \left\{ x_0^{(q_\lambda^L)} \right\} - \alpha_\lambda > 0, & \text{if } q_\lambda^L \leq \min \{1; e^{\beta_\lambda - 1}\}, \\ \infty, & \text{if } q_\lambda^L > \min \{1; e^{\beta_\lambda - 1}\}, \end{cases}$$

(e) the map $X_0 \mapsto B_{\lambda, X_0, n}^{I, L}$ is strictly increasing.

Analogously to the discussions in the Sections 3.17–3.20, for the parameter setups $(\mathcal{P}_{SP,2} \times \mathbb{R} \setminus (\mathcal{I}_{SP,2} \cup [0, 1])) \cup (\mathcal{P}_{SP,3a} \times \mathbb{R} \setminus (\mathcal{I}_{SP,3a} \cup [0, 1])) \cup (\mathcal{P}_{SP,3b} \times \mathbb{R} \setminus (\mathcal{I}_{SP,3b} \cup [0, 1])) \cup (\mathcal{P}_{SP,3c} \times \mathbb{R} \setminus [0, 1])$ and for all initial population sizes $X_0 \in \mathbb{N}$ one can still show

$$0 < I_\lambda(P_{A, n} || P_{H, n}), \quad \text{and} \quad \lim_{n \rightarrow \infty} I_\lambda(P_{A, n} || P_{H, n}) = \infty.$$

For the penultimate case we obtain

Corollary 10. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4a} \times (\mathbb{R} \setminus [0, 1])$ there exist parameters $p_\lambda^L > \alpha_\lambda$ (where not necessarily $p_\lambda^L \geq 0$) and $0 < q_\lambda^L < \beta_\lambda = \beta_\bullet < \min\{1, e^{\beta_\bullet - 1}\} = e^{\beta_\bullet - 1}$ such that (35) is satisfied for all $x \in [0, \infty[$ and such that for all initial population sizes $X_0 \in \mathbb{N}$ at least the parts (c) and (d) of Corollary 9 hold true.

Notice that for the last case $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4b} \times \mathbb{R} \setminus [0, 1]$ (where $(\beta_A = \beta_H \geq 1)$ we cannot derive lower bounds of the power divergences which improve the generally valid lower bound $I_\lambda(P_{A, n} || P_{H, n}) \geq 0$ (cf. (11)) by employing our proposed $(p_\lambda^U, q_\lambda^U)$ -method.

4.6. Upper Bounds of $I_\lambda(\cdot || \cdot)$ for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0, 1])$

For these constellations we adapt Proposition 14, which after modulation becomes

Corollary 11. For all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0,1])$ and all initial population sizes $X_0 \in \mathbb{N}$ there holds with $p_{\lambda}^U := \alpha_{\mathcal{A}}^{\lambda} \alpha_{\mathcal{H}}^{1-\lambda}$ and $q_{\lambda}^U := \beta_{\mathcal{A}}^{\lambda} \beta_{\mathcal{H}}^{1-\lambda}$

- (a) $B_{\lambda, X_0, 1}^{I,U} = \frac{1}{\lambda(\lambda-1)} \cdot \left(\exp \left\{ \left(\beta_{\mathcal{A}}^{\lambda} \beta_{\mathcal{H}}^{1-\lambda} - \beta_{\lambda} \right) \cdot X_0 + \alpha_{\mathcal{A}}^{\lambda} \alpha_{\mathcal{H}}^{1-\lambda} - \alpha_{\lambda} \right\} - 1 \right) > 0,$
- (b) the sequence $\left(B_{\lambda, X_0, n}^{I,U} \right)_{n \in \mathbb{N}}$ of upper bounds for $I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ given by

$$B_{\lambda, X_0, n}^{I,U} = \frac{1}{\lambda(\lambda-1)} \cdot \left(\exp \left\{ a_n^{(q_{\lambda}^U)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_{\lambda}^U, q_{\lambda}^U)} \right\} - 1 \right)$$
 is strictly increasing,
- (c) $\lim_{n \rightarrow \infty} B_{\lambda, X_0, n}^{I,U} = \infty,$
- (d) $\lim_{n \rightarrow \infty} \frac{1}{n} \log B_{\lambda, X_0, n}^{I,U} = \begin{cases} p_{\lambda}^U \cdot \exp \left\{ x_0^{(q_{\lambda}^U)} \right\} - \alpha_{\lambda} > 0, & \text{if } \lambda \in [\lambda_-, \lambda_+] \setminus [0,1], \\ \infty, & \text{if } \lambda \in]-\infty, \lambda_-[\cup]\lambda_+, \infty[, \end{cases}$
- (e) the map $X_0 \mapsto B_{\lambda, X_0, n}^{I,U}$ is strictly increasing.

4.7. Applications to Bayesian Decision Making

As explained in Section 2.5, the power divergences fulfill

$$I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = \int_0^1 \Delta \mathcal{BR}_{\mathcal{LO}}(p_{\mathcal{A}}^{\text{prior}}) \cdot (1 - p_{\mathcal{A}}^{\text{prior}})^{\lambda-2} \cdot (p_{\mathcal{A}}^{\text{prior}})^{-1-\lambda} dp_{\mathcal{A}}^{\text{prior}}, \quad \lambda \in \mathbb{R}, \quad (\text{cf. (21)}),$$

and

$$I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = \lim_{\chi \rightarrow p_{\mathcal{A}}^{\text{prior}}} \Delta \mathcal{BR}_{\mathcal{LO}, \lambda, \chi}(p_{\mathcal{A}}^{\text{prior}}), \quad \lambda \in]0, 1[, \quad (\text{cf. (22)}),$$

and thus can be interpreted as (i) *weighted-average* decision risk reduction (weighted-average statistical information measure) about the degree of evidence deg concerning the parameter θ that can be attained by observing the GWI-path \mathcal{X}_n until stage n , and as (ii) *limit* decision risk reduction (limit statistical information measure). Hence, by combining (21) and (22) with the investigations in the previous Sections 4.1–4.6, we obtain exact recursive values respectively recursive bounds of the above-mentioned decision risk reductions. For the sake of brevity, we omit the details here.

5. Kullback-Leibler Information Divergence (Relative Entropy)

5.1. Exact Values Respectively Upper Bounds of $I(\cdot || \cdot)$

From (2), (3) and (6) in Section 2.4, one can immediately see that the Kullback-Leibler information divergence (relative entropy) between two competing Galton-Watson processes without/with immigration can be obtained by the limit

$$I(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = \lim_{\lambda \nearrow 1} I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n}), \quad (68)$$

and the reverse Kullback-Leibler information divergence (reverse relative entropy) by $I(P_{\mathcal{H},n} || P_{\mathcal{A},n}) = \lim_{\lambda \searrow 0} I_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$. Hence, in the following we concentrate only on (68), the reverse case works analogously. Accordingly, we can use (68) in appropriate combination with the $\lambda \in]0, 1[$ -parts of the previous Section 4 (respectively, the corresponding parts of Section 3) in order to obtain detailed analyses for $I(P_{\mathcal{H},n} || P_{\mathcal{A},n})$. Let us start with the following assertions on exact values respectively upper bounds, which will be proved in Appendix A.2:

Theorem 3.

(a) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1})$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ the Kullback-Leibler information divergence (relative entropy) is given by

$$I(P_{A,n}||P_{H,n}) = I_{X_0,n} := \begin{cases} \frac{\beta_A \cdot (\log(\frac{\beta_A}{\beta_H}) - 1) + \beta_H}{1 - \beta_A} \cdot [X_0 - \frac{\alpha_A}{1 - \beta_A}] \cdot (1 - (\beta_A)^n) \\ + \frac{\alpha_A \cdot [\beta_A \cdot (\log(\frac{\beta_A}{\beta_H}) - 1) + \beta_H]}{\beta_A(1 - \beta_A)} \cdot n, & \text{if } \beta_A \neq 1, \\ [\beta_H - \log \beta_H - 1] \cdot [\frac{\alpha_A}{2} \cdot n^2 + (X_0 + \frac{\alpha_A}{2}) \cdot n], & \text{if } \beta_A = 1. \end{cases} \tag{69}$$

(b) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ there holds $I(P_{A,n}||P_{H,n}) \leq E_{X_0,n}^U$, where

$$E_{X_0,n}^U := \begin{cases} \frac{\beta_A \cdot (\log(\frac{\beta_A}{\beta_H}) - 1) + \beta_H}{1 - \beta_A} \cdot [X_0 - \frac{\alpha_A}{1 - \beta_A}] \cdot (1 - (\beta_A)^n) \\ + \left[\frac{\alpha_A \cdot [\beta_A \cdot (\log(\frac{\beta_A}{\beta_H}) - 1) + \beta_H]}{\beta_A(1 - \beta_A)} + \alpha_A \left[\log\left(\frac{\alpha_A \beta_H}{\alpha_H \beta_A}\right) - \frac{\beta_H}{\beta_A} \right] + \alpha_H \right] \cdot n, & \text{if } \beta_A \neq 1, \\ [\beta_H - \log \beta_H - 1] \cdot [\frac{\alpha_A}{2} \cdot n^2 + (X_0 + \frac{\alpha_A}{2}) \cdot n] \\ + \left[\alpha_A \left[\log\left(\frac{\alpha_A \beta_H}{\alpha_H}\right) - \beta_H \right] + \alpha_H \right] \cdot n, & \text{if } \beta_A = 1. \end{cases} \tag{70}$$

Remark 5.

(i) Notice that the exact values respectively upper bounds are in closed form (rather than in recursive form).
 (ii) The n -behaviour of (the bounds of) the Kullback-Leibler information divergence/relative entropy $I(P_{A,n}||P_{H,n})$ in Theorem 3 is influenced by the following facts:

- (a) $\beta_A \cdot (\log(\frac{\beta_A}{\beta_H}) - 1) + \beta_H \geq 0$ with equality iff $\beta_A = \beta_H$.
- (b) In the case $\beta_A \neq 1$ of (70), there holds $\frac{\alpha_A \cdot [\beta_A \cdot (\log(\frac{\beta_A}{\beta_H}) - 1) + \beta_H]}{\beta_A(1 - \beta_A)} + \alpha_A \left[\log\left(\frac{\alpha_A \beta_H}{\alpha_H \beta_A}\right) - \frac{\beta_H}{\beta_A} \right] + \alpha_H \geq 0$, with equality iff $\alpha_A = \alpha_H$ and $\beta_A = \beta_H$.

5.2. Lower Bounds of $I(\cdot||\cdot)$ for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1})$

Again by using (68) in appropriate combination with the “ $\lambda \in]0, 1[$ -parts” of the previous Section 4 (respectively, the corresponding parts of Section 3), we obtain the following (semi-)closed-form lower bounds of $I(P_{H,n}||P_{A,n})$:

Theorem 4. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$

$$I(P_{A,n}||P_{H,n}) \geq E_{X_0,n}^L := \sup_{k \in \mathbb{N}_0, y \in [0, \infty[} \left\{ E_{y, X_0, n}^{L, tan}, E_{k, X_0, n}^{L, sec}, E_{X_0, n}^{L, hor} \right\} \in [0, \infty[, \tag{71}$$

where for all $y \in [0, \infty[$ we define the – possibly negatively valued– finite bound component

$$E_{y, X_0, n}^{L, tan} := \begin{cases} \left[\beta_A \log \left(\frac{\alpha_A + \beta_A y}{\alpha_H + \beta_H y} \right) + \beta_H \left(1 - \frac{\alpha_A + \beta_A y}{\alpha_H + \beta_H y} \right) \right] \cdot \frac{1 - (\beta_A)^n}{1 - \beta_A} \cdot \left[X_0 - \frac{\alpha_A}{1 - \beta_A} \right] \\ + \left[\frac{\alpha_A}{\beta_A(1 - \beta_A)} \left[\beta_A \log \left(\frac{\alpha_A + \beta_A y}{\alpha_H + \beta_H y} \right) + \beta_H \left(1 - \frac{\alpha_A + \beta_A y}{\alpha_H + \beta_H y} \right) \right] \right. \\ \left. + \left(\alpha_H - \alpha_A \frac{\beta_H}{\beta_A} \right) \left(1 - \frac{\alpha_A + \beta_A y}{\alpha_H + \beta_H y} \right) \right] \cdot n, & \text{if } \beta_A \neq 1, \\ \left[\log \left(\frac{\alpha_A + y}{\alpha_H + \beta_H y} \right) + \beta_H \left(1 - \frac{\alpha_A + y}{\alpha_H + \beta_H y} \right) \right] \cdot \left[\frac{\alpha_A}{2} \cdot n^2 + \left(X_0 + \frac{\alpha_A}{2} \right) \cdot n \right] \\ + \left(\alpha_H - \alpha_A \beta_H \right) \left(1 - \frac{\alpha_A + y}{\alpha_H + \beta_H y} \right) \cdot n, & \text{if } \beta_A = 1, \end{cases} \quad (72)$$

and for all $k \in \mathbb{N}_0$ the – possibly negatively valued– finite bound component

$$E_{k, X_0, n}^{L, sec} := \begin{cases} \left[f_A(k+1) \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) - f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) + \beta_H - \beta_A \right] \cdot \frac{1 - (\beta_A)^n}{1 - \beta_A} \cdot \left[X_0 - \frac{\alpha_A}{1 - \beta_A} \right] \\ + \left[\frac{\alpha_A}{\beta_A(1 - \beta_A)} \left(f_A(k+1) \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) - f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) + \beta_H - \beta_A \right) \right. \\ \left. - \left(f_A(k+1) \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) - f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) \right) \right] \cdot \left(k + \frac{\alpha_A}{\beta_A} \right) \\ + f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) - \frac{\alpha_A \beta_H}{\beta_A} + \alpha_H \right] \cdot n, & \text{if } \beta_A \neq 1, \\ \left[f_A(k+1) \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) - f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) + \beta_H - 1 \right] \cdot \left[\frac{\alpha_A}{2} \cdot n^2 + \left(X_0 + \frac{\alpha_A}{2} \right) \cdot n \right] \\ - \left[\left(f_A(k+1) \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) - f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) \right) \right] \cdot \left(k + \alpha_A \right) \\ - f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) + \alpha_A \beta_H - \alpha_H \right] \cdot n, & \text{if } \beta_A = 1. \end{cases} \quad (73)$$

Furthermore, on $\mathcal{P}_{SP, A}$ we set $E_{X_0, n}^{L, hor} := 0$ for all $n \in \mathbb{N}$ whereas on $\mathcal{P}_{SP} \setminus (\mathcal{P}_{SP, 1} \cup \mathcal{P}_{SP, A})$ we define

$$E_{X_0, n}^{L, hor} := \left[(\alpha_A + \beta_A z^*) \cdot \left[\log \left(\frac{\alpha_A + \beta_A z^*}{\alpha_H + \beta_H z^*} \right) - 1 \right] + \alpha_H + \beta_H z^* \right] \cdot n, \quad , n \in \mathbb{N}, \quad (74)$$

with $z^* := \arg \max_{x \in \mathbb{N}_0} \left\{ (\alpha_A + \beta_A x) \left[-\log \left(\frac{\alpha_A + \beta_A x}{\alpha_H + \beta_H x} \right) + 1 \right] - (\alpha_H + \beta_H x) \right\}$.

On $\mathcal{P}_{SP} \setminus (\mathcal{P}_{SP, 1} \cup \mathcal{P}_{SP, 3c})$ one even gets $E_{X_0, n}^{L, hor} > 0$ for all $X_0 \in \mathbb{N}$ and all $n \in \mathbb{N}$.

For the subcase $\mathcal{P}_{SP, 3c}$, one obtains for each fixed $n \in \mathbb{N}$ and each fixed $X_0 \in \mathbb{N}$ the strict positivity $E_{X_0, n}^{L, hor} > 0$ if

$\left(\frac{\partial}{\partial y} E_{y, n}^{L, tan} \right) (y^*) \neq 0$, where $y^* := \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} \in \mathbb{N}$ and hence

$$\begin{aligned} & \left(\frac{\partial}{\partial y} E_{y, X_0, n}^{L, tan} \right) (y^*) & (75) \\ = & \begin{cases} - \frac{(\beta_A - \beta_H)^3}{\alpha_A \beta_H - \alpha_H \beta_A} \cdot \frac{1 - (\beta_A)^n}{1 - \beta_A} \cdot \left[X_0 - \frac{\alpha_A}{1 - \beta_A} \right] - \frac{(\beta_A - \beta_H)^2}{\beta_A} \left(1 + \frac{\alpha_A (\beta_A - \beta_H)}{(1 - \beta_A)(\alpha_A \beta_H - \alpha_H \beta_A)} \right) \cdot n, & \text{if } \beta_A \neq 1, \\ - \frac{(1 - \beta_H)^3}{\alpha_A \beta_H - \alpha_H} \cdot \left[\frac{\alpha_A}{2} \cdot n^2 + \left(X_0 + \frac{\alpha_A}{2} \right) \cdot n \right] - (1 - \beta_H)^2 \cdot n, & \text{if } \beta_A = 1. \end{cases} \end{aligned}$$

A proof of this theorem is given in in Appendix A.2.

Remark 6. Consider the exemplary parameter setup $(\beta_A, \beta_H, \alpha_A, \alpha_H) = (\frac{1}{3}, \frac{2}{3}, 2, 1) \in \mathcal{P}_{SP, 3c}$; within our running-example epidemiological context of Section 2.3, this corresponds to a “semi-mild” infectious-disease-transmission situation (\mathcal{H}) (with subcritical reproduction number $\beta_H = \frac{2}{3}$ and importation mean of $\alpha_H = 1$), whereas (\mathcal{A}) describes a “mild” situation (with “low” subcritical $\beta_A = \frac{1}{3}$ and $\alpha_A = 2$). In the case of $X_0 = 3$ there holds $\left(\frac{\partial}{\partial y} E_{y, X_0, n}^{L, tan} \right) (y^*) = 0$ for all $n \in \mathbb{N}$, whereas for $X_0 \neq 3$ one obtains $\left(\frac{\partial}{\partial y} E_{y, X_0, n}^{L, tan} \right) (y^*) \neq 0$ for all $n \in \mathbb{N}$.

It seems that the optimization problem in (71) admits in general only an implicitly representable solution, and thus we have used the prefix “(semi-)” above. Of course, as a less tight but less involved

explicit lower bound of the Kullback-Leibler information divergence (relative entropy) $I(P_{\mathcal{A},n}||P_{\mathcal{H},n})$ one can use any term of the form $\max \{ E_{y,X_0,n}^{L,tan}, E_{k,X_0,n}^{L,sec}, E_{X_0,n}^{L,hor} \}$ ($y \in [0, \infty[$, $k \in \mathbb{N}_0$), as well as the following

Corollary 12. (a) For all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$

$$I(P_{\mathcal{A},n}||P_{\mathcal{H},n}) \geq E_{X_0,n}^L \geq \tilde{E}_{X_0,n}^L := \max \{ E_{\infty,X_0,n}^{L,tan}, E_{0,X_0,n}^{L,sec}, E_{X_0,n}^{L,hor} \} \in [0, \infty[,$$

with $E_{X_0,n}^{L,hor}$ defined by (74), with – possibly negatively valued– finite bound component $E_{\infty,X_0,n}^{L,tan} := \lim_{y \rightarrow \infty} E_{y,X_0,n}^{L,tan}$, where

$$E_{\infty,X_0,n}^{L,tan} := \begin{cases} \frac{\beta_{\mathcal{A}} \cdot (\log(\frac{\beta_{\mathcal{A}}}{\beta_{\mathcal{H}}}) - 1) + \beta_{\mathcal{H}}}{1 - \beta_{\mathcal{A}}} \cdot [X_0 - \frac{\alpha_{\mathcal{A}}}{1 - \beta_{\mathcal{A}}}] \cdot (1 - (\beta_{\mathcal{A}})^n) \\ + \left[\frac{\alpha_{\mathcal{A}} \cdot [\beta_{\mathcal{A}} \cdot (\log(\frac{\beta_{\mathcal{A}}}{\beta_{\mathcal{H}}}) - 1) + \beta_{\mathcal{H}}]}{\beta_{\mathcal{A}}(1 - \beta_{\mathcal{A}})} + \alpha_{\mathcal{A}} \left(1 - \frac{\beta_{\mathcal{H}}}{\beta_{\mathcal{A}}}\right) + \alpha_{\mathcal{H}} \left(1 - \frac{\beta_{\mathcal{A}}}{\beta_{\mathcal{H}}}\right) \right] \cdot n, & \text{if } \beta_{\mathcal{A}} \neq 1, \\ \left[\beta_{\mathcal{H}} - \log \beta_{\mathcal{H}} - 1 \right] \cdot \left[\frac{\alpha_{\mathcal{A}}}{2} \cdot n^2 + \left(X_0 + \frac{\alpha_{\mathcal{A}}}{2} \right) \cdot n \right] \\ + \left[\alpha_{\mathcal{A}} (1 - \beta_{\mathcal{H}}) + \alpha_{\mathcal{H}} \left(1 - \frac{1}{\beta_{\mathcal{H}}}\right) \right] \cdot n, & \text{if } \beta_{\mathcal{A}} = 1, \end{cases}$$

and –possibly negatively valued–finite bound component

$$E_{0,X_0,n}^{L,sec} = \begin{cases} \left[(\alpha_{\mathcal{A}} + \beta_{\mathcal{A}}) \cdot \log \left(\frac{\alpha_{\mathcal{A}} + \beta_{\mathcal{A}}}{\alpha_{\mathcal{H}} + \beta_{\mathcal{H}}} \right) - \alpha_{\mathcal{A}} \cdot \log \left(\frac{\alpha_{\mathcal{A}}}{\alpha_{\mathcal{H}}} \right) + \beta_{\mathcal{H}} - \beta_{\mathcal{A}} \right] \cdot \frac{1 - (\beta_{\mathcal{A}})^n}{1 - \beta_{\mathcal{A}}} \cdot \left[X_0 - \frac{\alpha_{\mathcal{A}}}{1 - \beta_{\mathcal{A}}} \right] \\ + \left\{ \frac{\alpha_{\mathcal{A}}}{\beta_{\mathcal{A}}(1 - \beta_{\mathcal{A}})} \left((\alpha_{\mathcal{A}} + \beta_{\mathcal{A}}) \cdot \log \left(\frac{\alpha_{\mathcal{A}} + \beta_{\mathcal{A}}}{\alpha_{\mathcal{H}} + \beta_{\mathcal{H}}} \right) - \alpha_{\mathcal{A}} \cdot \log \left(\frac{\alpha_{\mathcal{A}}}{\alpha_{\mathcal{H}}} \right) \right) - \frac{\alpha_{\mathcal{A}}}{1 - \beta_{\mathcal{A}}} (1 - \beta_{\mathcal{H}}) \right. \\ \left. - \alpha_{\mathcal{A}} \left(1 + \frac{\alpha_{\mathcal{A}}}{\beta_{\mathcal{A}}}\right) \cdot \log \left(\frac{\alpha_{\mathcal{H}}(\alpha_{\mathcal{A}} + \beta_{\mathcal{A}})}{\alpha_{\mathcal{A}}(\alpha_{\mathcal{H}} + \beta_{\mathcal{H}})} \right) + \alpha_{\mathcal{H}} \right\} \cdot n, & \text{if } \beta_{\mathcal{A}} \neq 1, \\ \left[(\alpha_{\mathcal{A}} + 1) \cdot \log \left(\frac{\alpha_{\mathcal{A}} + 1}{\alpha_{\mathcal{H}} + \beta_{\mathcal{H}}} \right) - \alpha_{\mathcal{A}} \cdot \log \left(\frac{\alpha_{\mathcal{A}}}{\alpha_{\mathcal{H}}} \right) + \beta_{\mathcal{H}} - 1 \right] \cdot [n \cdot X_0 + \frac{\alpha_{\mathcal{A}}}{2} \cdot n^2] \\ + \left\{ \frac{\alpha_{\mathcal{A}}}{2} \left[(\alpha_{\mathcal{A}} + 1) \cdot \log \left(\frac{\alpha_{\mathcal{A}} + 1}{\alpha_{\mathcal{H}} + \beta_{\mathcal{H}}} \right) - \alpha_{\mathcal{A}} \cdot \log \left(\frac{\alpha_{\mathcal{A}}}{\alpha_{\mathcal{H}}} \right) - \beta_{\mathcal{H}} - 1 \right] \right. \\ \left. - \alpha_{\mathcal{A}} (1 + \alpha_{\mathcal{A}}) \cdot \log \left(\frac{\alpha_{\mathcal{H}}(\alpha_{\mathcal{A}} + 1)}{\alpha_{\mathcal{A}}(\alpha_{\mathcal{H}} + \beta_{\mathcal{H}})} \right) + \alpha_{\mathcal{H}} \right\} \cdot n, & \text{if } \beta_{\mathcal{A}} = 1. \end{cases}$$

For the cases $\mathcal{P}_{SP,2} \cup \mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b}$ one gets even $\tilde{E}_{X_0,n}^L > 0$ for all $X_0 \in \mathbb{N}$ and all $n \in \mathbb{N}$.

5.3. Applications to Bayesian Decision Making

As explained in Section 2.5, the Kullback-Leibler information divergence fulfills

$$I(P_{\mathcal{A},n}||P_{\mathcal{H},n}) = \int_0^1 \Delta \mathcal{BR}_{\widehat{CO}}(p_{\mathcal{A}}^{\text{prior}}) \cdot (1 - p_{\mathcal{A}}^{\text{prior}})^{-1} \cdot (p_{\mathcal{A}}^{\text{prior}})^{-2} dp_{\mathcal{A}}^{\text{prior}}, \quad (\text{cf. (21) with } \lambda = 1),$$

and thus can be interpreted as *weighted-average* decision risk reduction (weighted-average statistical information measure) about the degree of evidence deg concerning the parameter θ that can be attained by observing the GWI-path \mathcal{X}_n until stage n . Hence, by combining (21) with the investigations in the previous Sections 5.1 and 5.2, we obtain exact values respectively bounds of the above-mentioned decision risk reductions. For the sake of brevity, we omit the details here.

6. Explicit Closed-Form Bounds of Hellinger Integrals

6.1. Principal Approach

Depending on the parameter constellation $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P} \times (\mathbb{R} \setminus \{0, 1\})$, for the Hellinger integrals $H_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ we have derived in Section 3 corresponding lower/upper bounds respectively exact values—of recursive nature— which can be obtained by choosing appropriate $p = p_{\lambda}^A = p^A(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda)$, $q = q_{\lambda}^A = q^A(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda)$ ($A \in \{E, L, U\}$) and by using those together with the recursion $(a_n^{(q)})_{n \in \mathbb{N}}$ defined by (36) as well as the sequence $(b_n^{(p,q)})_{n \in \mathbb{N}}$ obtained from $(a_n^{(q)})_{n \in \mathbb{N}}$ by the linear transformation (38). Both sequences are “stepwise fully evaluable” but generally seem not to admit a closed-form representation in the observation horizons n ; consequently, the time-evolution $n \mapsto H_{\lambda}(P_{\mathcal{A},n} || P_{\mathcal{H},n})$ —respectively the time-evolution of the corresponding recursive bounds— can generally *not be seen explicitly*. On order to avoid this *intransparency* (at the expense of losing some precision) one can approximate (36) by a recursion that allows for a closed-form representation; by the way, this will also turn out to be useful for investigations concerning diffusion limits (cf. the next Section 7).

To explain the basic underlying principle, let us first assume some *general* $q \in]0, \beta_{\lambda}[$ and $\lambda \in]0, 1[$. With Properties 1 (P1) we see that the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ is strictly negative, strictly decreasing and converges to $x_0^{(q)} \in]-\beta_{\lambda}, q - \beta_{\lambda}[$. Recall that this sequence is obtained by the recursive application of the function $\zeta_{\lambda}^{(q)}(x) := q \cdot e^x - \beta_{\lambda}$, through $a_1^{(q)} = \zeta_{\lambda}^{(q)}(0) = q - \beta_{\lambda} < 0$, $a_n^{(q)} = \zeta_{\lambda}^{(q)}(a_{n-1}^{(q)}) = qe^{a_{n-1}^{(q)}} - \beta_{\lambda}$ (cf. (36)). As a first step, we want to approximate $\zeta_{\lambda}^{(q)}(\cdot)$ by a linear function on the interval $[x_0^{(q)}, 0]$. Due to convexity (P9), this is done by using the tangent line of $\zeta_{\lambda}^{(q)}(\cdot)$ at $x_0^{(q)}$

$$\zeta_{\lambda}^{(q),T}(x) := c^{(q),T} + d^{(q),T} \cdot x := x_0^{(q)} \left(1 - q \cdot e^{x_0^{(q)}} \right) + q \cdot e^{x_0^{(q)}} \cdot x, \tag{76}$$

as a linear lower bound, and the secant line of $\zeta_{\lambda}^{(q)}(\cdot)$ across its arguments 0 and $x_0^{(q)}$

$$\zeta_{\lambda}^{(q),S}(x) := c^{(q),S} + d^{(q),S} \cdot x := q - \beta_{\lambda} + \frac{x_0^{(q)} - (q - \beta_{\lambda})}{x_0^{(q)}} \cdot x, \tag{77}$$

as a linear upper bound. With the help of these functions, we can define the *linear* recursions

$$a_0^{(q),T} := 0, \quad a_n^{(q),T} := \zeta_{\lambda}^{(q),T}(a_{n-1}^{(q),T}), \quad n \in \mathbb{N}, \tag{78}$$

as well as
$$a_0^{(q),S} := 0, \quad a_n^{(q),S} := \zeta_{\lambda}^{(q),S}(a_{n-1}^{(q),S}), \quad n \in \mathbb{N}. \tag{79}$$

In the following, we will refer to these sequences as the *rudimentary closed-form sequence-bounds*. Clearly, both sequences are strictly negative (on \mathbb{N}), strictly decreasing, and one gets the sandwiching

$$a_n^{(q),T} < a_n^{(q)} \leq a_n^{(q),S} \tag{80}$$

for all $n \in \mathbb{N}$, with equality on the right side iff $n = 1$ (where $a_1^{(q)} = q - \beta_{\lambda} < 0$); moreover,

$$\lim_{n \rightarrow \infty} a_n^{(q),T} = \lim_{n \rightarrow \infty} a_n^{(q),S} = \lim_{n \rightarrow \infty} a_n^{(q)} = x_0^{(q)}. \tag{81}$$

Furthermore, such linear recursions allow for a closed-form representation, namely

$$a_n^{(q),*} = \frac{c^{(q),*}}{1 - d^{(q),*}} \cdot \left(1 - \left(d^{(q),*}\right)^n\right) = x_0^{(q)} \cdot \left(1 - \left(d^{(q),*}\right)^n\right), \tag{82}$$

where the “*” stands for either *S* or *T*. Notice that this representation is valid due to $d^{(q),T}, d^{(q),S} \in]0, 1[$. So far, we have considered the case $q \in]0, \beta_\lambda[$. If $q = \beta_\lambda$, then one can see from Properties 1 (P2) that $a_n^{(q)} \equiv 0$, which is also an explicitly given (though trivial) sequence. For the remaining case, where $q > \beta_\lambda$ and thus $\zeta_\lambda^{(q)}(0) = a_1^{(q)} = q - \beta_\lambda > 0$, we want to exclude $q \geq \min\{1, e^{\beta_\lambda - 1}\}$ for the following reasons. Firstly, if $q > \min\{1, e^{\beta_\lambda - 1}\}$, then from (P3) we see that the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ is strictly increasing and divergent to ∞ , at a rate faster than exponentially (P3b); but a linear recursion is too weak to approximate such a growth pattern. Secondly, if $q = \min\{1, e^{\beta_\lambda - 1}\}$, then one necessarily gets $q = e^{\beta_\lambda - 1} < 1$ (since we have required $q > \beta_\lambda$, and otherwise one obtains the contradiction $\beta_\lambda < q = 1 \leq e^{\beta_\lambda - 1}$). This means that the function $\zeta_\lambda^{(q)}(\cdot)$ now touches the straight line $id(\cdot)$ in the point $-\log(q)$, i.e., $\zeta_\lambda^{(q)}(-\log(q)) = -\log(q)$. Our above-proposed method, namely to use the tangent line of $\zeta_\lambda^{(q)}(\cdot)$ at $x = x_0^{(q)} = -\log(q)$ as a linear lower bound for $\zeta_\lambda^{(q)}(\cdot)$, leads then to the recursion $a_n^{(q),T} \equiv 0$ (cf. (78)). This is due to the fact that the tangent line $\zeta_\lambda^{(q),T}(\cdot)$ is in the current case equivalent with the straight line $id(\cdot)$. Consequently, (81) would not be satisfied.

Notice that in the case $\beta_\lambda < q < \min\{1, e^{\beta_\lambda - 1}\}$, the above-introduced functions $\zeta_\lambda^{(q),T}(\cdot), \zeta_\lambda^{(q),S}(\cdot)$ constitute again linear lower and upper bounds for $\zeta_\lambda^{(q)}(\cdot)$, however, this time on the interval $[0, x_0^{(q)}]$. The sequences defined in (78) and (79) still fulfill the assertions (80) and (81), and additionally allow for the closed-form representation (82). Furthermore, let us mention that these rudimentary closed-form sequence-bounds can be defined analogously for $\lambda \in \mathbb{R} \setminus]0, 1]$ and either $0 < q < \beta_\lambda$, or $q = \beta_\lambda$, or $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}$.

In a second step, we want to *improve* the above-mentioned linear (lower and upper) approximations of the sequence $a_n^{(q)}$ by reducing the faced error within each iteration. To do so, in both cases of lower and upper approximates we shall employ context-adapted linear *inhomogeneous difference equations* of the form

$$\tilde{a}_0 := 0 \quad ; \quad \tilde{a}_n := \tilde{\zeta}(\tilde{a}_{n-1}) + \rho_{n-1}, \quad n \in \mathbb{N}, \tag{83}$$

with

$$\tilde{\zeta}(x) := c + d \cdot x, \quad x \in \mathbb{R}, \tag{84}$$

$$\rho_{n-1} := K_1 \cdot \varkappa^{n-1} + K_2 \cdot \nu^{n-1}, \quad n \in \mathbb{N}, \tag{85}$$

for some constants $c \in \mathbb{R}, d \in]0, 1[, K_1, K_2, \varkappa, \nu \in \mathbb{R}$ with $0 \leq \nu < \varkappa \leq d$. This will be applied to $c := c^{(q),S}, c := c^{(q),T}, d := d^{(q),S}$ and $d := d^{(q),T}$ later on. Meanwhile, let us first present some facts and expressions which are insightful for further formulations and analyses.

Lemma 2. Consider the sequence $(\tilde{a}_n)_{n \in \mathbb{N}_0}$ defined in (83) to (85). If $0 \leq \nu < \varkappa < d$, then one gets the closed-form representation

$$\tilde{a}_n = \tilde{a}_n^{hom} + \tilde{c}_n \quad \text{with} \quad \tilde{a}_n^{hom} = c \cdot \frac{1 - d^n}{1 - d} \quad \text{and} \quad \tilde{c}_n = K_1 \cdot \frac{d^n - \varkappa^n}{d - \varkappa} + K_2 \cdot \frac{d^n - \nu^n}{d - \nu}, \tag{86}$$

which leads for all $n \in \mathbb{N}$ to

$$\sum_{k=1}^n \tilde{a}_k = \left(\frac{K_1}{d - \varkappa} + \frac{K_2}{d - \nu} - \frac{c}{1 - d} \right) \cdot \frac{d \cdot (1 - d^n)}{1 - d} - \frac{K_1 \cdot \varkappa \cdot (1 - \varkappa^n)}{(d - \varkappa)(1 - \varkappa)} - \frac{K_2 \cdot \nu \cdot (1 - \nu^n)}{(d - \nu)(1 - \nu)} + \frac{c}{1 - d} \cdot n. \tag{87}$$

If $0 \leq \nu < \varkappa = d$, then one gets the closed-form representation

$$\tilde{a}_n = \tilde{a}_n^{hom} + \tilde{c}_n \quad \text{with} \quad \tilde{a}_n^{hom} = c \cdot \frac{1 - d^n}{1 - d} \quad \text{and} \quad \tilde{c}_n = K_1 \cdot n \cdot d^{n-1} + K_2 \cdot \frac{d^n - \nu^n}{d - \nu}, \quad (88)$$

which leads for all $n \in \mathbb{N}$ to

$$\sum_{k=1}^n \tilde{a}_k = \left(\frac{K_1}{d(1-d)} + \frac{K_2}{d-\nu} - \frac{c}{1-d} \right) \cdot \frac{d \cdot (1-d^n)}{1-d} - \frac{K_2 \cdot \nu \cdot (1-\nu^n)}{(d-\nu)(1-\nu)} + \left(\frac{c}{1-d} - \frac{K_1 \cdot d^n}{1-d} \right) \cdot n. \quad (89)$$

Lemma 2 will be proved in Appendix A.3. Notice that (88) is consistent with taking the limit $\varkappa \nearrow d$ in (86). Furthermore, for the special case $K_2 = -K_1 > 0$ one has from (85) for all integers $n \geq 2$ the relation $\rho_{n-1} < 0$ and thus $\tilde{a}_n - \tilde{a}_n^{hom} < 0$, leading to

$$\tilde{c}_n < 0 \quad \text{and} \quad \sum_{k=1}^n \tilde{c}_k < 0. \quad (90)$$

Lemma 2 gives explicit expressions for a linear inhomogeneous recursion of the form (83) possessing the extra term given by (85). Therefrom we derive lower and upper bounds for the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ by employing $a_n^{(q),T}$ resp. $a_n^{(q),S}$ as the homogeneous solution of (83), i.e., by setting $\tilde{a}_n^{hom} := a_n^{(q),T}$ resp. $\tilde{a}_n^{hom} := a_n^{(q),S}$. Moreover, our concrete approximation-error-reducing “correction terms” ρ_n will have different form, depending on whether $0 < q < \beta_\lambda$ or $q > \max\{0, \beta_\lambda\}$. In both cases, we express ρ_n by means of the slopes $d^{(q),T} = qe^{x_0^{(q)}}$ resp. $d^{(q),S} = \frac{x_0^{(q)} - (q - \beta_\lambda)}{x_0^{(q)}}$ of the tangent line $\xi_\lambda^{(q),T}(\cdot)$ (cf. (76)) resp. the secant line $\xi_\lambda^{(q),S}(\cdot)$ (cf. (77)), as well as in terms of the parameters

$$\Gamma_{<}^{(q)} := \frac{1}{2} \cdot (x_0^{(q)})^2 \cdot q \cdot e^{x_0^{(q)}}, \quad \text{for } 0 < q < \beta_\lambda, \quad \text{and} \quad \Gamma_{>}^{(q)} := \frac{q}{2} \cdot (x_0^{(q)})^2, \quad \text{for } q > \max\{0, \beta_\lambda\}. \quad (91)$$

In detail, let us first define the lower approximate by

$$\underline{a}_0^{(q)} := 0, \quad \underline{a}_n^{(q)} := \xi_\lambda^{(q),T}(\underline{a}_{n-1}^{(q)}) + \underline{\rho}_{n-1}^{(q)}, \quad n \in \mathbb{N}, \quad (92)$$

where

$$\underline{\rho}_{n-1}^{(q)} := \begin{cases} \Gamma_{<}^{(q)} \cdot (d^{(q),T})^{2(n-1)}, & \text{if } 0 < q < \beta_\lambda, \\ \Gamma_{>}^{(q)} \cdot (d^{(q),S})^{2(n-1)}, & \text{if } \max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}. \end{cases} \quad (93)$$

The upper approximate is defined by

$$\bar{a}_0^{(q)} := 0, \quad \bar{a}_n^{(q)} := \xi_\lambda^{(q),S}(\bar{a}_{n-1}^{(q)}) + \bar{\rho}_{n-1}^{(q)}, \quad n \in \mathbb{N}, \quad (94)$$

where

$$\bar{\rho}_{n-1}^{(q)} := \begin{cases} -\Gamma_{<}^{(q)} \cdot (d^{(q),T})^{n-1} \cdot [1 - (d^{(q),S})^{n-1}], & \text{if } 0 < q < \beta_\lambda, \\ -\Gamma_{>}^{(q)} \cdot (d^{(q),S})^{n-1} \cdot [1 - (d^{(q),T})^{n-1}], & \text{if } \max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}. \end{cases} \quad (95)$$

In terms of (85), we use for $\underline{\rho}_n^{(q)}$ the constants $K_2 = \nu = 0$ as well as $K_1 = \Gamma_{<}^{(q)}$, $\varkappa = (d^{(q),T})^2$ for $0 < q < \beta_\lambda$ respectively $K_1 = \Gamma_{>}^{(q)}$, $\varkappa = (d^{(q),S})^2$ for $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}$. For $\bar{\rho}_n^{(q)}$ we

shall employ the constants $-K_1 = K_2 = \Gamma_{<}^{(q)}$, $\varkappa = d^{(q),T}$, $\nu = d^{(q),S}d^{(q),T}$ for $0 < q < \beta_\lambda$, and $-K_1 = K_2 = \Gamma_{>}^{(q)}$, $\varkappa = d^{(q),S}$, $\nu = d^{(q),S}d^{(q),T}$ for $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}$. Recall from (76) the constants $c^{(q),T} := x_0^{(q)}(1 - qe^{x_0^{(q)}})$, $d^{(q),T} := qe^{x_0^{(q)}}$ and from (77) $c^{(q),S} := q - \beta_\lambda$, $d^{(q),S} := \frac{x_0^{(q)} - (q - \beta_\lambda)}{x_0^{(q)}}$.

In the following, we will refer to the sequences $\underline{a}_n^{(q)}$ resp. $\bar{a}_n^{(q)}$ as the *improved closed-form sequence-bounds*. Putting all ingredients together, we arrive at the

Lemma 3. For all $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}$ there holds with $d^{(q),T} = qe^{x_0^{(q)}}$ and $d^{(q),S} = \frac{x_0^{(q)} - (q - \beta_\lambda)}{x_0^{(q)}}$

(a) in the case $0 < q < \beta_\lambda$:

(i)

$$\underline{a}_n^{(q)} < a_n^{(q)} \leq \bar{a}_n^{(q)} \quad \text{for all } n \in \mathbb{N},$$

with equality on the right-hand side iff $n = 1$, where

$$\underline{a}_n^{(q)} = x_0^{(q)} \cdot \left(1 - \left(d^{(q),T}\right)^n\right) + \Gamma_{<}^{(q)} \cdot \frac{\left(d^{(q),T}\right)^{n-1}}{1 - d^{(q),T}} \cdot \left(1 - \left(d^{(q),T}\right)^n\right) > a_n^{(q),T}, \quad \text{and}$$

$$\bar{a}_n^{(q)} = x_0^{(q)} \cdot \left(1 - \left(d^{(q),S}\right)^n\right) - \Gamma_{<}^{(q)} \cdot \left[\frac{\left(d^{(q),S}\right)^n - \left(d^{(q),T}\right)^n}{d^{(q),S} - d^{(q),T}} - \left(d^{(q),S}\right)^{n-1} \frac{1 - \left(d^{(q),T}\right)^n}{1 - d^{(q),T}} \right] \leq a_n^{(q),S},$$

with $a_n^{(q),T}$ and $a_n^{(q),S}$ defined by (78) and (79).

(ii) Both sequences $\left(\underline{a}_n^{(q)}\right)_{n \in \mathbb{N}}$ and $\left(\bar{a}_n^{(q)}\right)_{n \in \mathbb{N}}$ are strictly decreasing.

(iii)

$$\lim_{n \rightarrow \infty} \underline{a}_n^{(q)} = \lim_{n \rightarrow \infty} \bar{a}_n^{(q)} = \lim_{n \rightarrow \infty} a_n^{(q)} = x_0^{(q)} \in] - \beta_\lambda, q - \beta_\lambda[.$$

(b) in the case $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}$:

(i)

$$\underline{a}_n^{(q)} < a_n^{(q)} \leq \bar{a}_n^{(q)}, \quad \text{for all } n \in \mathbb{N},$$

with equality on the right-hand side iff $n = 1$, where

$$\underline{a}_n^{(q)} = x_0^{(q)} \cdot \left(1 - \left(d^{(q),T}\right)^n\right) + \Gamma_{>}^{(q)} \cdot \frac{\left(d^{(q),T}\right)^n - \left(d^{(q),S}\right)^{2n}}{d^{(q),T} - \left(d^{(q),S}\right)^2} > a_n^{(q),T} \quad \text{and}$$

$$\bar{a}_n^{(q)} = x_0^{(q)} \cdot \left(1 - \left(d^{(q),S}\right)^n\right) - \Gamma_{>}^{(q)} \cdot \left(d^{(q),S}\right)^{n-1} \left[n - \frac{1 - \left(d^{(q),T}\right)^n}{1 - d^{(q),T}} \right] \leq a_n^{(q),S},$$

with $a_n^{(q),T}$ and $a_n^{(q),S}$ defined by (78) and (79).

(ii) Both sequences $\left(\underline{a}_n^{(q)}\right)_{n \in \mathbb{N}}$ and $\left(\bar{a}_n^{(q)}\right)_{n \in \mathbb{N}}$ are strictly increasing.

(iii)

$$\lim_{n \rightarrow \infty} \underline{a}_n^{(q)} = \lim_{n \rightarrow \infty} \bar{a}_n^{(q)} = \lim_{n \rightarrow \infty} a_n^{(q)} = x_0^{(q)} \in]q - \beta_\lambda, -\log(q)[.$$

A detailed proof of Lemma 3 is provided in Appendix A.3. In the following, we employ the above-mentioned investigations in order to derive the desired closed-form bounds of the Hellinger integrals $H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n})$.

6.2. Explicit Closed-Form Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP1}) \times (\mathbb{R} \setminus \{0, 1\})$

Recall that in this setup, we have obtained the recursive, non-explicit exact values $V_{\lambda, X_0, n} = H_{\lambda}(P_{A,n} || P_{H,n})$ given in (39) of Theorem 1, where we used $q = q_{\lambda}^E = q^E(\beta_A, \beta_H, \lambda) = \beta_A^{\lambda} \beta_H^{1-\lambda} \in]0, \beta_{\lambda}[$ in the case $\lambda \in]0, 1[$ respectively $q = q_{\lambda}^E = \beta_A^{\lambda} \beta_H^{1-\lambda} > \max\{0, \beta_{\lambda}\}$ in the case $\lambda \in \mathbb{R} \setminus]0, 1[$. For the latter, Lemma 1 implies that $q_{\lambda}^E < \min\{1, e^{\beta_{\lambda}-1}\}$ iff $\lambda \in]\lambda_-, \lambda_+[\setminus]0, 1[$. This—together with (39) from Theorem 1, Lemma 2 and with the quantities $d^{(q),T}, d^{(q),S}, \Gamma_{<}^{(q)}$ and $\Gamma_{>}^{(q)}$ as defined in (76) and (77) resp. (91)—leads to

Theorem 5. Let $p_{\lambda}^E := \alpha_A^{\lambda} \alpha_H^{1-\lambda}$ and $q_{\lambda}^E := \beta_A^{\lambda} \beta_H^{1-\lambda}$. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP1}) \times (]\lambda_-, \lambda_+[\setminus \{0, 1\})$, all initial population sizes $X_0 \in \mathbb{N}$ and for all observation horizons $n \in \mathbb{N}$ the following assertions hold true:

(a) the Hellinger integral can be bounded by the closed-form lower and upper bounds

$$C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), T} \leq C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), L} \leq V_{\lambda, X_0, n} = H_{\lambda}(P_{A,n} || P_{H,n}) \leq C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), U} \leq C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), S},$$

(b)

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log(V_{\lambda, X_0, n}) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log(C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), L}) = \lim_{n \rightarrow \infty} \frac{1}{n} \log(C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), U}) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log(C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), T}) = \lim_{n \rightarrow \infty} \frac{1}{n} \log(C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), S}) = \frac{\alpha_A}{\beta_A} \cdot x_0^{(q_{\lambda}^E)}, \end{aligned}$$

where the involved closed-form lower bounds are defined by

$$\begin{aligned} C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), L} &:= C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), T} \cdot \exp\left\{\zeta_n^{(q_{\lambda}^E)} \cdot X_0 + \frac{\alpha_A}{\beta_A} \cdot \underline{\vartheta}_n^{(q_{\lambda}^E)}\right\}, \quad \text{with} \tag{96} \\ C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), T} &:= \exp\left\{x_0^{(q_{\lambda}^E)} \cdot \left[X_0 - \frac{\alpha_A}{\beta_A} \cdot \frac{d^{(q_{\lambda}^E), T}}{1 - d^{(q_{\lambda}^E), T}}\right] \cdot \left(1 - (d^{(q_{\lambda}^E), T})^n\right) + \frac{\alpha_A}{\beta_A} x_0^{(q_{\lambda}^E)} \cdot n\right\}, \end{aligned}$$

and the closed-form upper bounds are defined by

$$\begin{aligned} C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), U} &:= C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), S} \cdot \exp\left\{-\bar{\zeta}_n^{(q_{\lambda}^E)} \cdot X_0 - \frac{\alpha_A}{\beta_A} \cdot \bar{\vartheta}_n^{(q_{\lambda}^E)}\right\}, \quad \text{with} \tag{97} \\ C_{\lambda, X_0, n}^{(p_{\lambda}^E, q_{\lambda}^E), S} &:= \exp\left\{x_0^{(q_{\lambda}^E)} \cdot \left[X_0 - \frac{\alpha_A}{\beta_A} \cdot \frac{d^{(q_{\lambda}^E), S}}{1 - d^{(q_{\lambda}^E), S}}\right] \cdot \left(1 - (d^{(q_{\lambda}^E), S})^n\right) + \frac{\alpha_A}{\beta_A} x_0^{(q_{\lambda}^E)} \cdot n\right\}, \end{aligned}$$

where in the case $\lambda \in]0, 1[$

$$\zeta_n^{(q_{\lambda}^E)} := \Gamma_{<}^{(q_{\lambda}^E)} \cdot \frac{(d^{(q_{\lambda}^E), T})^{n-1}}{1 - d^{(q_{\lambda}^E), T}} \cdot \left(1 - (d^{(q_{\lambda}^E), T})^n\right) > 0, \tag{98}$$

$$\underline{\vartheta}_n^{(q_{\lambda}^E)} := \Gamma_{<}^{(q_{\lambda}^E)} \cdot \frac{1 - (d^{(q_{\lambda}^E), T})^n}{(1 - d^{(q_{\lambda}^E), T})^2} \cdot \left[1 - \frac{d^{(q_{\lambda}^E), T} (1 + (d^{(q_{\lambda}^E), T})^n)}{1 + d^{(q_{\lambda}^E), T}}\right] > 0, \tag{99}$$

$$\bar{\zeta}_n^{(q_{\lambda}^E)} := \Gamma_{<}^{(q_{\lambda}^E)} \cdot \left[\frac{(d^{(q_{\lambda}^E), S})^n - (d^{(q_{\lambda}^E), T})^n}{d^{(q_{\lambda}^E), S} - d^{(q_{\lambda}^E), T}} - (d^{(q_{\lambda}^E), S})^{n-1} \cdot \frac{1 - (d^{(q_{\lambda}^E), T})^n}{1 - d^{(q_{\lambda}^E), T}}\right] > 0, \tag{100}$$

$$\bar{\vartheta}_n^{(q_{\lambda}^E)} := \Gamma_{<}^{(q_{\lambda}^E)} \cdot \frac{d^{(q_{\lambda}^E), T}}{1 - d^{(q_{\lambda}^E), T}} \cdot \left[\frac{1 - (d^{(q_{\lambda}^E), S} d^{(q_{\lambda}^E), T})^n}{1 - d^{(q_{\lambda}^E), S} d^{(q_{\lambda}^E), T}} - \frac{(d^{(q_{\lambda}^E), S})^n - (d^{(q_{\lambda}^E), T})^n}{d^{(q_{\lambda}^E), S} - d^{(q_{\lambda}^E), T}}\right] > 0, \tag{101}$$

and where in the case $\lambda \in]\lambda_-, \lambda_+[\setminus]0, 1[$

$$\underline{\zeta}_n^{(q_\lambda^E)} := \Gamma_{>}^{(q_\lambda^E)} \cdot \frac{\left(d^{(q_\lambda^E),T}\right)^n - \left(d^{(q_\lambda^E),S}\right)^{2n}}{d^{(q_\lambda^E),T} - \left(d^{(q_\lambda^E),S}\right)^2} > 0, \tag{102}$$

$$\underline{\vartheta}_n^{(q_\lambda^E)} := \frac{\Gamma_{>}^{(q_\lambda^E)}}{d^{(q_\lambda^E),T} - \left(d^{(q_\lambda^E),S}\right)^2} \left[\frac{d^{(q_\lambda^E),T} \left(1 - \left(d^{(q_\lambda^E),T}\right)^n\right)}{1 - d^{(q_\lambda^E),T}} - \frac{\left(d^{(q_\lambda^E),S}\right)^2 \left(1 - \left(d^{(q_\lambda^E),S}\right)^{2n}\right)}{1 - \left(d^{(q_\lambda^E),S}\right)^2} \right] > 0, \tag{103}$$

$$\bar{\zeta}_n^{(q_\lambda^E)} := \Gamma_{>}^{(q_\lambda^E)} \cdot \left(d^{(q_\lambda^E),S}\right)^{n-1} \cdot \left[n - \frac{1 - \left(d^{(q_\lambda^E),T}\right)^n}{1 - d^{(q_\lambda^E),T}} \right] > 0, \tag{104}$$

$$\bar{\vartheta}_n^{(q_\lambda^E)} := \Gamma_{>}^{(q_\lambda^E)} \cdot \left[\frac{d^{(q_\lambda^E),S} - d^{(q_\lambda^E),T}}{\left(1 - d^{(q_\lambda^E),S}\right)^2 \left(1 - d^{(q_\lambda^E),T}\right)} \cdot \left(1 - \left(d^{(q_\lambda^E),S}\right)^n\right) + \frac{d^{(q_\lambda^E),T} \left(1 - \left(d^{(q_\lambda^E),S} d^{(q_\lambda^E),T}\right)^n\right)}{\left(1 - d^{(q_\lambda^E),T}\right) \left(1 - d^{(q_\lambda^E),S} d^{(q_\lambda^E),T}\right)} - \frac{\left(d^{(q_\lambda^E),S}\right)^n}{1 - d^{(q_\lambda^E),S}} \cdot n \right] > 0. \tag{105}$$

Notice that $\frac{\alpha_A}{\beta_A}$ can be equivalently be replaced by $\frac{\alpha_H}{\beta_H}$ in (96) and in (97).

A proof of Theorem 5 is given in Appendix A.3.

6.3. Explicit Closed-Form Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$

To derive (explicit) closed-form lower bounds of the (nonexplicit) recursive lower bounds $B_{\lambda, X_0, n}^L$ for the Hellinger integral $H_\lambda(P_{A,n} || P_{H,n})$ respectively closed-form upper bounds of the recursive upper bounds $B_{\lambda, X_0, n}^U$ for all parameters cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus \{0, 1\})$, we combine part (b) of Theorem 1, Lemma 2, Lemma 3 together with appropriate parameters $p_\lambda^L = p^L(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$, $p_\lambda^U = p^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \geq 0$ and $q_\lambda^L = q^L(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$, $q_\lambda^U = q^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) > 0$ satisfying (35). Notice that the representations of the lower and upper closed-form sequence-bounds depend on whether $0 < q_\lambda^A < \beta_\lambda$, $0 < q_\lambda^A = \beta_\lambda$ or $\max\{0, \beta_\lambda\} < q_\lambda^A < \min\{1, e^{\beta_\lambda - 1}\}$ ($A \in \{L, U\}$).

Let us start with closed-form lower bounds for the case $\lambda \in]0, 1[$; recall that the choice $p_\lambda^L = \alpha_A^\lambda \alpha_H^{1-\lambda}$, $q_\lambda^L = \beta_A^\lambda \beta_H^{1-\lambda}$ led to the optimal recursive lower bounds $B_{\lambda, X_0, n}^L$ of the Hellinger integral (cf. Theorem 1(b) and Section 3.5). Correspondingly, we can derive

Theorem 6. Let $p_\lambda^L = \alpha_A^\lambda \alpha_H^{1-\lambda}$ and $q_\lambda^L = \beta_A^\lambda \beta_H^{1-\lambda}$. Then, the following assertions hold true:

(a) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,2} \cup \mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b} \cup \mathcal{P}_{SP,3c}) \times]0, 1[$ (for which particularly $0 < q_\lambda^L < \beta_\lambda, \beta_A \neq \beta_H$), all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ there holds

$$C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} \leq C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), L} \leq B_{\lambda, X_0, n}^L < 1,$$

where $C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), L} := C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} \cdot \exp \left\{ \underline{\zeta}_n^{(q_\lambda^L)} \cdot X_0 + \frac{p_\lambda^L}{q_\lambda^L} \cdot \underline{\vartheta}_n^{(q_\lambda^L)} \right\}$ (106)

with $C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} := \exp \left\{ x_0^{(q_\lambda^L)} \cdot \left[X_0 - \frac{p_\lambda^L}{q_\lambda^L} \cdot \frac{d^{(q_\lambda^L), T}}{1 - d^{(q_\lambda^L), T}} \right] \cdot \left(1 - \left(d^{(q_\lambda^L), T} \right)^n \right) + \left(\frac{p_\lambda^L}{q_\lambda^L} \cdot \left(\beta_\lambda + x_0^{(q_\lambda^L)} \right) - \alpha_\lambda \right) \cdot n \right\},$

and with $\underline{\zeta}_n^{(q_\lambda^L)} := \Gamma_{<}^{(q_\lambda^L)} \cdot \frac{\left(d^{(q_\lambda^L), T} \right)^{n-1}}{1 - d^{(q_\lambda^L), T}} \cdot \left(1 - \left(d^{(q_\lambda^L), T} \right)^n \right) > 0,$ (107)

$$\underline{\vartheta}_n^{(q_\lambda^L)} := \Gamma_{<}^{(q_\lambda^L)} \cdot \frac{1 - \left(d^{(q_\lambda^L), T} \right)^n}{\left(1 - d^{(q_\lambda^L), T} \right)^2} \cdot \left[1 - \frac{d^{(q_\lambda^L), T} \left(1 + \left(d^{(q_\lambda^L), T} \right)^n \right)}{1 + d^{(q_\lambda^L), T}} \right] > 0. \quad (108)$$

(b) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,4a} \cup \mathcal{P}_{SP,4b}) \times]0, 1[$ (for which particularly $0 < q_\lambda^L = \beta_\lambda, \beta_A = \beta_H$), all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ there holds

$$C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), L} := C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} := B_{\lambda, X_0, n}^L = \exp \left\{ \left(p_\lambda^L - \alpha_\lambda \right) \cdot n \right\} < 1.$$

(c) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$ and all initial population sizes $X_0 \in \mathbb{N}$ one gets

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} \right) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), L} \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(B_{\lambda, X_0, n}^L \right) \\ &= \frac{p_\lambda^L}{q_\lambda^L} \cdot \left(\beta_\lambda + x_0^{(q_\lambda^L)} \right) - \alpha_\lambda < 0, \end{aligned}$$

where in the case $\beta_A = \beta_H$ there holds $q_\lambda^L = \beta_\lambda$ and $x_0^{(q_\lambda^L)} = 0$.

The proof will be provided in Appendix A.3.

In order to deduce closed-form upper bounds for the case $\lambda \in]0, 1[$, we first recall from the Sections 3.6–3.13, that we have to employ suitable parameters $p_\lambda^U = p^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$, $q_\lambda^U = q^U(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda)$ satisfying (35). Notice that we automatically obtain $p_\lambda^U \geq p_\lambda^L = \alpha_A^\lambda \alpha_H^{1-\lambda} > 0$. Correspondingly, we obtain

Theorem 7. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times]0, 1[$, all coefficients p_λ^U, q_λ^U which satisfy (35) for all $x \in \mathbb{N}_0$ and additionally either $0 < q_\lambda^U \leq \beta_\lambda$ or $\beta_\lambda < q_\lambda^U < \min\{1, e^{\beta_\lambda - 1}\}$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ the following assertions hold true:

$$C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} \geq C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), U} \geq \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} \geq B_{\lambda, X_0, n}^U, \quad \text{where} \quad (109)$$

(a) in the case $0 < q_\lambda^U < \beta_\lambda$ one has

$$C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), U} := C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} \cdot \exp \left\{ -\bar{\zeta}_n^{(q_\lambda^U)} \cdot X_0 - \frac{p_\lambda^U}{q_\lambda^U} \cdot \bar{\vartheta}_n^{(q_\lambda^U)} \right\} \tag{110}$$

with
$$C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} := \exp \left\{ x_0^{(q_\lambda^U)} \cdot \left[X_0 - \frac{p_\lambda^U}{q_\lambda^U} \cdot \frac{d^{(q_\lambda^U), S}}{1 - d^{(q_\lambda^U), S}} \right] \cdot \left(1 - \left(d^{(q_\lambda^U), S} \right)^n \right) + \left(\frac{p_\lambda^U}{q_\lambda^U} \cdot \left(\beta_\lambda + x_0^{(q_\lambda^U)} \right) - \alpha_\lambda \right) \cdot n \right\},$$

$$\bar{\zeta}_n^{(q_\lambda^U)} := \Gamma_{<}^{(q_\lambda^U)} \cdot \left[\frac{\left(d^{(q_\lambda^U), S} \right)^n - \left(d^{(q_\lambda^U), T} \right)^n}{d^{(q_\lambda^U), S} - d^{(q_\lambda^U), T}} - \left(d^{(q_\lambda^U), S} \right)^{n-1} \cdot \frac{1 - \left(d^{(q_\lambda^U), T} \right)^n}{1 - d^{(q_\lambda^U), T}} \right] > 0, \tag{111}$$

$$\bar{\vartheta}_n^{(q_\lambda^U)} := \Gamma_{<}^{(q_\lambda^U)} \cdot \frac{d^{(q_\lambda^U), T}}{1 - d^{(q_\lambda^U), T}} \cdot \left[\frac{1 - \left(d^{(q_\lambda^U), S} d^{(q_\lambda^U), T} \right)^n}{1 - d^{(q_\lambda^U), S} d^{(q_\lambda^U), T}} - \frac{\left(d^{(q_\lambda^U), S} \right)^n - \left(d^{(q_\lambda^U), T} \right)^n}{d^{(q_\lambda^U), S} - d^{(q_\lambda^U), T}} \right] > 0; \tag{112}$$

furthermore, whenever p_λ^U, q_λ^U satisfy additionally (47) (such parameters exist particularly in the setups $\mathcal{P}_{SP,2} \cup \mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b}$, cf. Sections 3.7–3.9), then

$$1 > C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} \quad \text{and} \quad \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} = B_{\lambda, X_0, n}^U \quad \forall n \in \mathbb{N};$$

(b) in the case $0 < q_\lambda^U = \beta_\lambda$ one has

$$C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), U} := C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} := \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} = \exp \left\{ \left(p_\lambda^U - \alpha_\lambda \right) \cdot n \right\};$$

(c) in the case $\beta_\lambda < q_\lambda^U < \min \{ 1, e^{\beta_\lambda - 1} \}$ the formulas (109) and (110) remain valid, but with

$$\bar{\zeta}_n^{(q_\lambda^U)} := \Gamma_{>}^{(q_\lambda^U)} \cdot \left(d^{(q_\lambda^U), S} \right)^{n-1} \cdot \left[n - \frac{1 - \left(d^{(q_\lambda^U), T} \right)^n}{1 - d^{(q_\lambda^U), T}} \right] > 0, \tag{113}$$

$$\bar{\vartheta}_n^{(q_\lambda^U)} := \Gamma_{>}^{(q_\lambda^U)} \cdot \left[\frac{d^{(q_\lambda^U), S} - d^{(q_\lambda^U), T}}{\left(1 - d^{(q_\lambda^U), S} \right)^2 \left(1 - d^{(q_\lambda^U), T} \right)} \cdot \left(1 - \left(d^{(q_\lambda^U), S} \right)^n \right) + \frac{d^{(q_\lambda^U), T} \left(1 - \left(d^{(q_\lambda^U), S} d^{(q_\lambda^U), T} \right)^n \right)}{\left(1 - d^{(q_\lambda^U), T} \right) \left(1 - d^{(q_\lambda^U), S} d^{(q_\lambda^U), T} \right)} - \frac{\left(d^{(q_\lambda^U), S} \right)^n}{1 - d^{(q_\lambda^U), S}} \cdot n \right] > 0; \tag{114}$$

(d) for all cases (a) to (c) one gets

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} \right) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), U} \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U)} \right) \\ &= \frac{p_\lambda^U}{q_\lambda^U} \cdot \left(\beta_\lambda + x_0^{(q_\lambda^U)} \right) - \alpha_\lambda, \end{aligned}$$

where in the case $q_\lambda^U = \beta_\lambda$ there holds $x_0^{(q_\lambda^U)} = 0$.

This Theorem 7 will be proved in Appendix A.3. Notice that for an inadequate choice of p_λ^U, q_λ^U it may hold that $\frac{p_\lambda^U}{q_\lambda^U}(\beta_\lambda + x_0^{(q_\lambda^U)}) - \alpha_\lambda > 0$ in part (d) of Theorem 7.

6.4. Explicit Closed-Form Bounds for the Cases $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0, 1])$

For $\lambda \in \mathbb{R} \setminus [0, 1]$, let us now construct closed-form lower bounds of the recursive lower bound components $\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)}$, for suitable parameters $p_\lambda^L \geq 0$ and either $0 < q_\lambda^L \leq \beta_\lambda$ or $\max\{0, \beta_\lambda\} < q_\lambda^L < \min\{1, e^{\beta_\lambda - 1}\}$ satisfying (35).

Theorem 8. For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\mathbb{R} \setminus [0, 1])$, all coefficients $p_\lambda^L \geq 0, q_\lambda^L > 0$ which satisfy (35) for all $x \in \mathbb{N}_0$ and either $0 < q_\lambda^L \leq \beta_\lambda$ or $\max\{0, \beta_\lambda\} < q_\lambda^L < \min\{1, e^{\beta_\lambda - 1}\}$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ the following assertions hold true:

$$C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} \leq C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), L} \leq \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} \leq B_{\lambda, X_0, n}^L, \quad \text{where} \tag{115}$$

(a) in the case $0 < q_\lambda^L < \beta_\lambda$ one has

$$C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), L} := C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} \cdot \exp \left\{ \zeta_n^{(q_\lambda^L)} \cdot X_0 + \frac{p_\lambda^L}{q_\lambda^L} \cdot \vartheta_n^{(q_\lambda^L)} \right\}, \tag{116}$$

$$\begin{aligned} \text{with } C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} := & \exp \left\{ x_0^{(q_\lambda^L)} \cdot \left[X_0 - \frac{p_\lambda^L}{q_\lambda^L} \cdot \frac{d^{(q_\lambda^L), T}}{1 - d^{(q_\lambda^L), T}} \right] \cdot \left(1 - \left(d^{(q_\lambda^L), T} \right)^n \right) \right. \\ & \left. + \left(\frac{p_\lambda^L}{q_\lambda^L} \cdot \left(\beta_\lambda + x_0^{(q_\lambda^L)} \right) - \alpha_\lambda \right) \cdot n \right\} \end{aligned}$$

$$\zeta_n^{(q_\lambda^L)} := \Gamma_{<}^{(q_\lambda^L)} \cdot \frac{\left(d^{(q_\lambda^L), T} \right)^{n-1}}{1 - d^{(q_\lambda^L), T}} \cdot \left(1 - \left(d^{(q_\lambda^L), T} \right)^n \right) > 0, \tag{117}$$

$$\vartheta_n^{(q_\lambda^L)} := \Gamma_{<}^{(q_\lambda^L)} \cdot \frac{1 - \left(d^{(q_\lambda^L), T} \right)^n}{\left(1 - d^{(q_\lambda^L), T} \right)^2} \cdot \left[1 - \frac{d^{(q_\lambda^L), T} \left(1 + \left(d^{(q_\lambda^L), T} \right)^n \right)}{1 + d^{(q_\lambda^L), T}} \right] > 0; \tag{118}$$

furthermore, whenever p_λ^L, q_λ^L satisfy additionally (56) (such parameters exist particularly in the setups $\mathcal{P}_{SP,2} \cup \mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b}$, cf. Sections 3.17–3.19), then

$$1 < C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} \quad \text{and} \quad \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} = B_{\lambda, X_0, n}^L \quad \forall n \in \mathbb{N};$$

(b) in the case $0 < q_\lambda^L = \beta_\lambda$ one has

$$C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), L} := C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} = \tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} = \exp \left\{ \left(p_\lambda^L - \alpha_\lambda \right) \cdot n \right\};$$

(c) in the case $\max\{0, \beta_\lambda\} < q_\lambda^L < \min\{1, e^{\beta_\lambda - 1}\}$ the formulas (115) and (116) remain valid, but with

$$\zeta_n^{(q_\lambda^L)} := \Gamma_{>}^{(q_\lambda^L)} \cdot \frac{\left(d^{(q_\lambda^L), T} \right)^n - \left(d^{(q_\lambda^L), S} \right)^{2n}}{d^{(q_\lambda^L), T} - \left(d^{(q_\lambda^L), S} \right)^2} > 0, \tag{119}$$

$$\vartheta_n^{(q_\lambda^L)} := \frac{\Gamma_{>}^{(q_\lambda^L)}}{d^{(q_\lambda^L), T} - \left(d^{(q_\lambda^L), S} \right)^2} \cdot \left[\frac{d^{(q_\lambda^L), T} \cdot \left(1 - \left(d^{(q_\lambda^L), T} \right)^n \right)}{1 - d^{(q_\lambda^L), T}} - \frac{\left(d^{(q_\lambda^L), S} \right)^2 \cdot \left(1 - \left(d^{(q_\lambda^L), S} \right)^{2n} \right)}{1 - \left(d^{(q_\lambda^L), S} \right)^2} \right] > 0; \tag{120}$$

(d) for all cases (a) to (c) one gets

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), T} \right) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L), L} \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\tilde{B}_{\lambda, X_0, n}^{(p_\lambda^L, q_\lambda^L)} \right) \\ &= \frac{p_\lambda^L}{q_\lambda^L} \cdot \left(\beta_\lambda + x_0^{(q_\lambda^L)} \right) - \alpha_\lambda, \end{aligned}$$

where in the case $q_\lambda^L = \beta_\lambda$ there holds $x_0^{(q_\lambda^L)} = 0$.

For the proof of Theorem 8, see Appendix A.3. Notice that for an inadequate choice of p_λ^L, q_λ^L it may hold that $\frac{p_\lambda^L}{q_\lambda^L} (\beta_\lambda + x_0^{(q_\lambda^L)}) - \alpha_\lambda < 0$ in the last assertion of Theorem 8.

To derive closed-form upper bounds of the recursive upper bounds $B_{\lambda, X_0, n}^U$ of the Hellinger integral in the case $\lambda \in \mathbb{R} \setminus [0, 1]$, let us first recall from Section 3.24 that we have to use the parameters $p_\lambda^U = \alpha_A^\lambda \alpha_H^{1-\lambda} > 0$ and $q_\lambda^U = \beta_A^\lambda \beta_H^{1-\lambda} > 0$. Furthermore, in the case $\beta_A \neq \beta_H$ we obtain from Lemma 1 (setting $q_\lambda = q_\lambda^U$) the assertion that $\max\{0, \beta_\lambda\} < q_\lambda^U < \min\{1, e^{\beta_\lambda - 1}\}$ iff $\lambda \in]\lambda_-, \lambda_+[\setminus [0, 1]$ (implying that the sequence $(a_n^{(q_\lambda^U)})_{n \in \mathbb{N}}$ converges). In the case $\beta_A = \beta_H$ one gets $q_\lambda^U = \beta_A^\lambda \beta_H^{1-\lambda} = \beta_A = \beta_H = \beta_\lambda$ and therefore (cf. (P2)) $a_n^{(q_\lambda^U)} = 0$ for all $n \in \mathbb{N}$ and for all $\lambda \in \mathbb{R} \setminus [0, 1]$. Correspondingly, we deduce

Theorem 9. Let $p_\lambda^U = \alpha_A^\lambda \alpha_H^{1-\lambda}$ and $q_\lambda^U = \beta_A^\lambda \beta_H^{1-\lambda}$. Then, the following assertions hold true:

(a) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,2} \cup \mathcal{P}_{SP,3a} \cup \mathcal{P}_{SP,3b} \cup \mathcal{P}_{SP,3c}) \times (]\lambda_-, \lambda_+[\setminus [0, 1])$ (in particular for $\beta_A \neq \beta_H$), all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ there holds

$$\infty > C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} \geq C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), U} \geq B_{\lambda, X_0, n}^U > 1,$$

where
$$C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), U} := C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} \cdot \exp \left\{ -\bar{\zeta}_n^{(q_\lambda^U)} \cdot X_0 - \frac{p_\lambda^U}{q_\lambda^U} \cdot \bar{\vartheta}_n^{(q_\lambda^U)} \right\} \tag{121}$$

with
$$C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} := \exp \left\{ x_0^{(q_\lambda^U)} \cdot \left[X_0 - \frac{p_\lambda^U}{q_\lambda^U} \cdot \frac{d^{(q_\lambda^U), T}}{1 - d^{(q_\lambda^U), T}} \right] \cdot \left(1 - \left(d^{(q_\lambda^U), T} \right)^n \right) + \left(\frac{p_\lambda^U}{q_\lambda^U} \cdot \left(\beta_\lambda + x_0^{(q_\lambda^U)} \right) - \alpha_\lambda \right) \cdot n \right\},$$

$$\bar{\zeta}_n^{(q_\lambda^U)} := \Gamma_{>}^{(q_\lambda^U)} \cdot \left(d^{(q_\lambda^U), S} \right)^{n-1} \cdot \left[n - \frac{1 - \left(d^{(q_\lambda^U), T} \right)^n}{1 - d^{(q_\lambda^U), T}} \right] > 0, \tag{122}$$

$$\begin{aligned} \bar{\vartheta}_n^{(q_\lambda^U)} := \Gamma_{>}^{(q_\lambda^U)} \cdot \left[\frac{d^{(q_\lambda^U), S} - d^{(q_\lambda^U), T}}{\left(1 - d^{(q_\lambda^U), S} \right)^2 \left(1 - d^{(q_\lambda^U), T} \right)} \cdot \left(1 - \left(d^{(q_\lambda^U), S} \right)^n \right) + \frac{d^{(q_\lambda^U), T} \left(1 - \left(d^{(q_\lambda^U), S} d^{(q_\lambda^U), T} \right)^n \right)}{\left(1 - d^{(q_\lambda^U), T} \right) \left(1 - d^{(q_\lambda^U), S} d^{(q_\lambda^U), T} \right)} - \frac{\left(d^{(q_\lambda^U), S} \right)^n}{1 - d^{(q_\lambda^U), S}} \cdot n \right] > 0. \end{aligned} \tag{123}$$

(b) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP,4a} \cup \mathcal{P}_{SP,4b}) \times (\mathbb{R} \setminus [0, 1])$ (for which particularly $0 < q_\lambda^U = \beta_\lambda, \beta_A = \beta_H$), all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ there holds

$$C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), U} := C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} := B_{\lambda, X_0, n}^U = \exp \left\{ \left(p_\lambda^U - \alpha_\lambda \right) \cdot n \right\} > 1.$$

(c) For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in (\mathcal{P}_{SP} \setminus \mathcal{P}_{SP,1}) \times (\lambda_-, \lambda_+ \setminus [0, 1])$ and all initial population sizes $X_0 \in \mathbb{N}$ one gets

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), S} \right) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^U, q_\lambda^U), U} \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(B_{\lambda, X_0, n}^U \right) \\ &= \frac{p_\lambda^U}{q_\lambda^U} \cdot \left(\beta_\lambda + x_0^{(q_\lambda^U)} \right) - \alpha_\lambda > 0, \end{aligned}$$

where in the case $\beta_A = \beta_H$ there holds $q_\lambda^U = \beta_\lambda$ and $x_0^{(q_\lambda^U)} = 0$.

A proof of Theorem 9 is provided in Appendix A.3.

Remark 7. Substituting $a_n^{(q)}$ by $a_n^{(q),T}$ resp. $a_n^{(q),S}$ (cf. (78) resp. (79)) in $\tilde{B}_{\lambda, X_0, n}^{(p,q)}$ from (42) leads to the “rudimentary” closed-form bounds $C_{\lambda, X_0, n}^{(p,q),T}$ resp. $C_{\lambda, X_0, n}^{(p,q),S}$, whereas substituting $a_n^{(q)}$ by $\underline{a}_n^{(q)}$ resp. $\bar{a}_n^{(q)}$ (cf. (92) resp. (94)) in $\tilde{B}_{\lambda, X_0, n}^{(p,q)}$ from (42) leads to the “improved” closed-form bounds $C_{\lambda, X_0, n}^{(p,q),L}$ resp. $C_{\lambda, X_0, n}^{(p,q),U}$ in all the Theorems 5–9.

6.5. Totally Explicit Closed-Form Bounds

The above-mentioned results give closed-form lower bounds $C_{\lambda, X_0, n}^{(p,q),L}$, $C_{\lambda, X_0, n}^{(p,q),T}$ resp. closed-form upper bounds $C_{\lambda, X_0, n}^{(p,q),U}$, $C_{\lambda, X_0, n}^{(p,q),S}$ of the Hellinger integrals $H_\lambda(P_{A,n} || P_{H,n})$ for case-dependent choices of p, q . However, these bounds still involve the fixed point $x_0^{(q)}$ which in general has to be calculated implicitly. In order to get “totally” explicit but “slightly” less tight closed-form bounds of $H_\lambda(P_{A,n} || P_{H,n})$, one can proceed as follows:

1. in all the closed-form lower bound formulas of the Theorems 5, 6 and 8—including the definitions (76), (77) and (91)—replace the implicit $x_0^{(q)}$ by a close explicitly known point $\underline{x}_0^{(q)} < x_0^{(q)}$;
2. in all closed-form upper bound formulas of the Theorems 5, 7 and 9—including (76), (77) and (91)—replace $x_0^{(q)}$ by a close explicitly known point $\bar{x}_0^{(q)} > x_0^{(q)}$.

For instance, one can use the following choices which will be also employed as an auxiliary tool for the diffusion-limit-concerning proof of Lemma A6 in Appendix A.4:

$$\underline{x}_0^{(q)} := \begin{cases} q^{-1} \cdot e^{-\underline{x}_0^{(q)}} \cdot \left[(1-q) - \sqrt{(1-q)^2 - 2 \cdot q \cdot e^{\underline{x}_0^{(q)}} \cdot (q - \beta_\lambda)} \right], & \text{if } q \in]0, \beta_\lambda[, \\ q^{-1} \cdot \left[(1-q) - \sqrt{(1-q)^2 - 2 \cdot q \cdot (q - \beta_\lambda)} \right], & \text{if } \max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}, \end{cases} \tag{124}$$

$$\text{where } \underline{x}_0^{(q)} := \begin{cases} \max\left\{-\beta_\lambda, \frac{q - \beta_\lambda}{1 - q}\right\}, & \text{if } q \in]0, 1[, \\ -\beta_\lambda, & \text{if } q \geq 1, \end{cases} \tag{125}$$

$$\bar{x}_0^{(q)} := \begin{cases} q^{-1} \cdot \left[(1-q) - \sqrt{(1-q)^2 - 2 \cdot q \cdot (q - \beta_\lambda)} \right], & \text{if } q \in]0, \beta_\lambda[, \\ (1-q) - \sqrt{(1-q)^2 - 2 \cdot q \cdot (q - \beta_\lambda)}, & \text{if } \max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\} \\ & \text{and } (1-q)^2 - 2 \cdot q \cdot (q - \beta_\lambda) \geq 0, \\ \bar{x}_0^{(q)} := -\log(q) & \text{if } \max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\} \\ & \text{and } (1-q)^2 - 2 \cdot q \cdot (q - \beta_\lambda) < 0. \end{cases} \tag{126}$$

Behind this choice “lies” the idea that—in contrast to the solution $x_0^{(q)}$ of $\xi_\lambda^{(q)}(x) := qe^x - \beta_\lambda = x$ —the point $\underline{x}_0^{(q)}$ is a solution of (the obviously explicitly solvable) $\underline{Q}_\lambda^{(q)}(x) := \underline{a}_\lambda^{(q)}x^2 + \underline{b}_\lambda^{(q)}x + \underline{c}_\lambda^{(q)} = x$ in both cases $0 < q < \beta_\lambda$ and $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}$, whereas the point $\bar{x}_0^{(q)}$ is a solution of $\bar{Q}_\lambda^{(q)}(x) := \bar{a}_\lambda^{(q)}x^2 + \bar{b}_\lambda^{(q)}x + \bar{c}_\lambda^{(q)} = x$ in the case $0 < q < \beta_\lambda$ and in the case $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}$ together with $(1 - q)^2 - 2 \cdot q \cdot (q - \beta_\lambda) \geq 0$. Thereby, $\underline{Q}_\lambda^{(q)}(\cdot)$ and $\bar{Q}_\lambda^{(q)}(\cdot)$ are the lower resp. upper quadratic approximates of $\xi_\lambda^{(q)}(\cdot)$ satisfying the following constraints:

- for $q \in]0, \beta_\lambda[$ (mostly but not only for $\lambda \in]0, 1[$) (lower bound):

$$\underline{Q}_\lambda^{(q)}(0) = \xi_\lambda^{(q)}(0) = q - \beta_\lambda, \quad \underline{Q}_\lambda^{(q)'}(0) = \xi_\lambda^{(q)'}(0) = q, \quad \underline{Q}_\lambda^{(q)''}(x) = \xi_\lambda^{(q)''}(y) = qe^y, \quad x \in \mathbb{R},$$

for some explicitly known approximate $y < x_0^{(q)}$ (leading to the (tighter) explicit lower approximate $\underline{x}_0^{(q)} \in]y, x_0^{(q)}[$); here, we choose

$$y := \underline{x}_0^{(q)} := \begin{cases} \max\{-\beta_\lambda, \frac{q-\beta_\lambda}{1-q}\}, & \text{if } q < 1, \\ -\beta_\lambda, & \text{if } q \geq 1; \end{cases}$$

- for $q \in]0, \beta_\lambda[$ (mostly but not only for $\lambda \in]0, 1[$) (upper bound):

$$\bar{Q}_\lambda^{(q)}(0) = \xi_\lambda^{(q)}(0) = q - \beta_\lambda, \quad \bar{Q}_\lambda^{(q)'}(0) = \xi_\lambda^{(q)'}(0) = q, \quad \bar{Q}_\lambda^{(q)''}(x) = \xi_\lambda^{(q)''}(0) = q, \quad x \in \mathbb{R};$$

- for $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}$ (mostly but not only for $\lambda \in \mathbb{R} \setminus [0, 1]$) (lower bound):

$$\underline{Q}_\lambda^{(q)}(0) = \xi_\lambda^{(q)}(0) = q - \beta_\lambda, \quad \underline{Q}_\lambda^{(q)'}(0) = \xi_\lambda^{(q)'}(0) = q, \quad \underline{Q}_\lambda^{(q)''}(x) = \xi_\lambda^{(q)''}(0) = q, \quad x \in \mathbb{R};$$

- for $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}$ in combination with $(1 - q)^2 - 2 \cdot q \cdot (q - \beta_\lambda) \geq 0$ (mostly but not only for $\lambda \in \mathbb{R} \setminus [0, 1]$) (upper bound):

$$\bar{Q}_\lambda^{(q)}(0) = \xi_\lambda^{(q)}(0) = q - \beta_\lambda, \quad \bar{Q}_\lambda^{(q)'}(0) = \xi_\lambda^{(q)'}(0) = q, \quad \bar{Q}_\lambda^{(q)''}(x) = \xi_\lambda^{(q)''}(-\log(q)) = 1, \quad x \in \mathbb{R}.$$

If $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}$ and $(1 - q)^2 - 2 \cdot q \cdot (q - \beta_\lambda) < 0$, then a real-valued solution $\bar{Q}_\lambda^{(q)}(x) = x$ does not exist and we set $\bar{x}_0^{(q)} := \bar{x}_0^{(q)} := -\log(q)$, with $\xi_\lambda^{(q)'}(\bar{x}_0^{(q)}) = 1$. The above considerations lead to corresponding unique choices of constants $\underline{a}_\lambda^{(q)}, \underline{b}_\lambda^{(q)}, \underline{c}_\lambda^{(q)}, \bar{a}_\lambda^{(q)}, \bar{b}_\lambda^{(q)}, \bar{c}_\lambda^{(q)}$ culminating in

$$\underline{Q}_\lambda^{(q)}(x) := \begin{cases} \frac{q}{2} \cdot e^{\underline{x}_0^{(q)}} \cdot x^2 + q \cdot x + q - \beta_\lambda, & \text{if } 0 < q < \beta_\lambda, \\ \frac{q}{2} \cdot x^2 + q \cdot x + q - \beta_\lambda, & \text{if } \max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}, \end{cases} \tag{127}$$

$$\bar{Q}_\lambda^{(q)}(x) := \begin{cases} \frac{q}{2} \cdot x^2 + q \cdot x + q - \beta_\lambda, & \text{if } 0 < q < \beta_\lambda, \\ \frac{1}{2} \cdot x^2 + q \cdot x + q - \beta_\lambda, & \text{if } \max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda-1}\}. \end{cases} \tag{128}$$

6.6. Closed-Form Bounds for Power Divergences of Non-Kullback-Leibler-Information-Divergence Type

Analogously to Section 4 (see especially Section 4.1), for orders $\lambda \in \mathbb{R} \setminus \{0, 1\}$ all the results of the previous Sections 6.1–6.5 carry correspondingly over from closed-form bounds of the Hellinger

integrals $H_\lambda(\cdot||\cdot)$ to closed-form bounds of the total variation distance $V(\cdot||\cdot)$, by virtue of the relation (cf. (12))

$$2 \left(1 - H_{\frac{1}{2}}(P_{\mathcal{A},n}||P_{\mathcal{H},n})\right) \leq V(P_{\mathcal{A},n}||P_{\mathcal{H},n}) \leq 2 \sqrt{1 - \left(H_{\frac{1}{2}}(P_{\mathcal{A},n}||P_{\mathcal{H},n})\right)^2},$$

to closed-form bounds of the Renyi divergences $R_\lambda(\cdot||\cdot)$, by virtue of the relation (cf. (7))

$$0 \leq R_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) = \frac{1}{\lambda(\lambda-1)} \log H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}), \quad \text{with } \log 0 := -\infty,$$

as well as to closed-form bounds of the power divergences $I_\lambda(\cdot||\cdot)$, by virtue of the relation (cf. (2))

$$I_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) = \frac{1 - H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n})}{\lambda \cdot (1 - \lambda)}, \quad n \in \mathbb{N}.$$

For the sake of brevity, the—merely repetitive—exact details are omitted.

6.7. Applications to Decision Making

The above-mentioned investigations of the Sections 6.1 to 6.6 can be applied to the context of Section 2.5 on *dichotomous* decision making on the space of all possible path scenarios (path space) of Poissonian Galton-Watson processes without (with) immigration GW(I) (e.g., in combination with our running-example epidemiological context of Section 2.3). More detailed, for the minimal mean decision loss (Bayes risk) \mathcal{R}_n defined by (18) we can derive explicit closed-form upper (respectively lower) bounds by using (19) respectively (20) together with the results of the Sections 6.1–6.5 concerning Hellinger integrals of order $\lambda \in]0, 1[$; we can proceed analogously in the Neyman-Pearson context in order to deduce closed-form bounds of type II error probabilities, by means of (23) and (24). Moreover, in an analogous way we can employ the investigations of Section 6.6 on power divergences in order to obtain closed-form bounds of (i) the corresponding (cf. (21)) *weighted-average* decision risk reduction (weighted-average statistical information measure) about the degree of evidence deg concerning the parameter θ that can be attained by observing the GW(I)-path \mathcal{X}_n until stage n , as well as (ii) the corresponding (cf. (22)) *limit* decision risk reduction (limit statistical information measure). For the sake of brevity, the—merely repetitive—exact details are omitted.

7. Hellinger Integrals and Power Divergences of Galton-Watson Type Diffusion Approximations

7.1. Branching-Type Diffusion Approximations

One can show that a properly rescaled Galton-Watson process without (respectively with) immigration GW(I) converges weakly to a diffusion process $\tilde{X} := \{\tilde{X}_s, s \in [0, \infty[\}$ which is the unique, strong, nonnegative – and in case of $\frac{\eta}{\sigma^2} \geq \frac{1}{2}$ strictly positive– solution of the stochastic differential equation (SDE) of the form

$$d\tilde{X}_s = \left(\eta - \kappa \tilde{X}_s\right) ds + \sigma \sqrt{\tilde{X}_s} dW_s, \quad s \in [0, \infty[, \quad \tilde{X}_0 \in]0, \infty[\text{ given}, \quad (129)$$

where $\eta \in [0, \infty[$, $\kappa \in [0, \infty[$, $\sigma \in]0, \infty[$ are constants and $\{W_s, s \in [0, \infty[\}$ denotes a standard Brownian motion with respect to the underlying probability measure P ; see e.g., Feller [130], Jirina [131], Lamperti [132,133], Lindvall [134,135], Grimvall [136], Jagers [56], Borovkov [137], Ethier & Kurtz [138], Durrett [139] for the non-immigration case corresponding to $\eta = 0, \kappa \geq 0$, Kawazu & Watanabe [140], Wei & Winnicki [141], Winnicki [64] for the immigration case corresponding to $\eta \neq 0, \kappa = 0$, as well as Sriram [142] for the general case $\eta \in [0, \infty[, \kappa \in \mathbb{R}$. Feller-type branching processes of the form (129), which are special cases of continuous state branching processes with immigration (see e.g., Kawazu & Watanabe [140], Li [143], as well as Dawson & Li [144] for imbeddings to affine processes) play

for instance an important role in the modelling of the term structure of interest rates, cf. the seminal Cox-Ingersoll-Ross CIR model [145] and the vast follow-up literature thereof. Furthermore, (129) is also prominently used as (a special case of) Cox & Ross’s [146] constant elasticity of variance CEV asset price process, as (part of) Heston’s [147] stochastic asset-volatility framework, as a model of neuron activity (see e.g., Lansky & Lanska [148], Giorno et al. [149], Lanska et al. [150], Lansky et al [151], Ditlevsen & Lansky [152], Höpfner [153], Lansky & Ditlevsen [154]), as a time-dynamic description of the nitrous oxide emission rate from the soil surface (see e.g., Pedersen [155]), as well as a model for the individual hazard rate in a survival analysis context (see e.g., Aalen & Gjessing [156]).

Along these lines of branching-type diffusion limits, it makes sense to consider the solutions of two SDEs (129) with different fixed parameter sets $(\eta, \kappa_{\mathcal{A}}, \sigma)$ and $(\eta, \kappa_{\mathcal{H}}, \sigma)$, determine for each of them a corresponding approximating GW(I), investigate the Hellinger integral between the laws of these two GW(I), and finally calculate the limit of the Hellinger integral (bounds) as the GW(I) approach their SDE solutions. Notice that for technicality reasons (which will be explained below), the constants η and σ ought to be independent of \mathcal{A}, \mathcal{H} in our current context.

In order to make the above-mentioned limit procedure rigorous, it is reasonable to work with appropriate approximations such that in each convergence step m one faces the setup $\mathcal{P}_{\text{NI}} \cup \mathcal{P}_{\text{SP},1}$ (i.e., the non-immigration or the equal-fraction case), where the corresponding Hellinger integral can be calculated exactly in a recursive way, as stated in Theorem 1. Let us explain the details in the following.

Consider a sequence of GW(I) $(X^{(m)})_{m \in \mathbb{N}}$ with probability laws $P_{\bullet}^{(m)}$ on a measurable space (Ω, \mathcal{F}) , where as above the subscript \bullet stands for either the hypothesis \mathcal{H} or the alternative \mathcal{A} . Analogously to (1), we use for each fixed step $m \in \mathbb{N}$ the representation $X^{(m)} := \{X_{\ell}^{(m)}, \ell \in \mathbb{N}\}$ with

$$X_{\ell}^{(m)} := \sum_{j=1}^{X_{\ell-1}^{(m)}} Y_{\ell-1,j}^{(m)} + \tilde{Y}_{\ell}^{(m)}, \quad \ell \in \mathbb{N}, \quad X_0^{(m)} \in \mathbb{N} \text{ given}, \tag{130}$$

where under the law $P_{\bullet}^{(m)}$

- the collection $Y^{(m)} := \{Y_{i,j}^{(m)}, i \in \mathbb{N}_0, j \in \mathbb{N}\}$ consists of i.i.d. random variables which are Poisson distributed with parameter $\beta_{\bullet}^{(m)} > 0$,
- the collection $\tilde{Y}^{(m)} := \{\tilde{Y}_i^{(m)}, i \in \mathbb{N}\}$ consists of i.i.d. random variables which are Poisson distributed with parameter $\alpha_{\bullet}^{(m)} \geq 0$,
- $Y^{(m)}$ and $\tilde{Y}^{(m)}$ are independent.

From arbitrary drift-parameters $\eta \in [0, \infty[$, $\kappa_{\bullet} \in [0, \infty[$, and diffusion-term-parameter $\sigma > 0$, we construct the offspring-distribution-parameter and the immigration-distribution parameter of the sequence $(X_{\ell}^{(m)})_{\ell \in \mathbb{N}}$ by

$$\beta_{\bullet}^{(m)} := 1 - \frac{\kappa_{\bullet}}{\sigma^2 m} \quad \text{and} \quad \alpha_{\bullet}^{(m)} := \beta_{\bullet}^{(m)} \cdot \frac{\eta}{\sigma^2}. \tag{131}$$

Here and henceforth, we always assume that the approximation step m is large enough to ensure that $\beta_{\bullet}^{(m)} \in]0, 1]$ and at least one of $\beta_{\mathcal{A}}^{(m)}, \beta_{\mathcal{H}}^{(m)}$ is strictly less than 1; this will be abbreviated by $m \in \bar{\mathbb{N}}$. Let us point out that – as mentioned above – our choice entails the best-to-handle setup $\mathcal{P}_{\text{NI}} \cup \mathcal{P}_{\text{SP},1}$ (which does not happen if instead of η one uses η_{\bullet} with $\eta_{\mathcal{A}} \neq \eta_{\mathcal{H}}$). Based on the GW(I) $X^{(m)}$, let us construct the *continuous-time* branching process $\tilde{X}^{(m)} := \{\tilde{X}_s^{(m)}, s \in [0, \infty[\}$ by

$$\tilde{X}_s^{(m)} := \frac{1}{m} X_{\lfloor \sigma^2 m s \rfloor}^{(m)}, \tag{132}$$

living on the state space $E^{(m)} := \frac{1}{m}\mathbb{N}_0$. Notice that $\tilde{X}^{(m)}$ is constant on each time-interval $\left[\frac{k}{\sigma^2 m}, \frac{k+1}{\sigma^2 m}\right]$ and takes at $s = \frac{k}{\sigma^2 m}$ the value $\frac{1}{m}X_k^{(m)}$ of the k -th GW(I) generation size, divided by m , i.e., it “jumps” with the jump-size $\frac{1}{m} \left(X_k^{(m)} - X_{k-1}^{(m)}\right)$ which is equal to the $\frac{1}{m}$ -fold difference to the previous generation size. From (132) one can immediately see the necessity of having σ to be independent of \mathcal{A} , \mathcal{H} because for the required law-equivalence in (the corresponding version of) (13) both models at stake have to “live” on the same time-scale $\tau_s^{(m)} := \lfloor \sigma^2 ms \rfloor$. For this setup, one obtains the following convergenc result:

Theorem 10. *Let $\eta \in [0, \infty[$, $\kappa_\bullet \in [0, \infty[$, $\sigma \in]0, \infty[$ and $\tilde{X}^{(m)}$ be as defined in (130) to (132). Furthermore, let us suppose that $\lim_{m \rightarrow \infty} \frac{1}{m} X_0^{(m)} = \tilde{X}_0 > 0$ and denote by $D([0, \infty[, [0, \infty[)$ the space of right-continuous functions $f : [0, \infty[\mapsto [0, \infty[$ with left limits. Then the sequence of processes $\left(\tilde{X}^{(m)}\right)_{m \in \bar{\mathbb{N}}}$ convergences in distribution in $D([0, \infty[, [0, \infty[)$ to a diffusion process \tilde{X} which is the unique strong, nonnegative—and in case of $\frac{\eta}{\sigma^2} \geq \frac{1}{2}$ strictly positive—solution of the SDE*

$$d\tilde{X}_s = (\eta - \kappa_\bullet \tilde{X}_s) ds + \sigma \sqrt{\tilde{X}_s} dW_s^\bullet, \quad s \in [0, \infty[, \quad \tilde{X}_0 \in]0, \infty[\text{ given}, \quad (133)$$

where $\{W_s^\bullet, s \in [0, \infty[\}$ denotes a standard Brownian motion with respect to the limit probability measure \tilde{P}_\bullet .

Remark 8. *Notice that the condition $\frac{\eta}{\sigma^2} \geq \frac{1}{2}$ can be interpreted in our approximation setup (131) as $\alpha_\bullet^{(m)} \geq \beta_\bullet^{(m)}/2$, which quantifies the intuitively reasonable indication that if the probability $P_\bullet[\tilde{Y}_\ell^{(m)} = 0] = e^{-\alpha_\bullet^{(m)}}$ of having no immigration is small enough relative to the probability $P_\bullet[Y_{\ell-1,k}^{(m)} = 0] = e^{-\beta_\bullet^{(m)}}$ of having no offspring ($m \in \bar{\mathbb{N}}$), then the limiting diffusion \tilde{X} never hits zero almost surely.*

The corresponding proof of Theorem 10—which is outlined in Appendix A.4—is an adaption of the proof of Theorem 9.1.3 in Ethier & Kurtz [138] which deals with drift-parameters $\eta = 0$, $\kappa_\bullet = 0$ in the SDE (133) whose solution is approached on a σ -independent time scale by a sequence of (critical) Galton-Watson processes without immigration but with general offspring distribution with mean 1 and variance σ . Notice that due to (131) the latter is inconsistent with our Poissonian setup, but this is compensated by our chosen σ -dependent time scale. Other limit investigations for (133) involving offspring/immigration distributions and parametrizations which are also incompatible to ours, are e.g., treated in Sriram [142].

As illustration of our proposed approach, let us give the following

Example 3. *Consider the parameter setup $(\eta, \kappa_\bullet, \sigma) = (5, 2, 0.4)$ and initial generation size $\tilde{X}_0 = 3$. Figure 4 shows the diffusion-approximation $\tilde{X}_s^{(m)}$ (blue) of the corresponding solution \tilde{X}_s of the SDE (133) up to the time horizon $T = 10$, for the approximation steps $m \in \{13, 50, 200, 1000\}$. Notice that in this setup there holds $\bar{\mathbb{N}} = \{k \in \mathbb{N} : k \geq 13\}$ (recall that $\bar{\mathbb{N}}$ is the subset of the positive integers such that $\beta_\bullet^{(m)} = 1 - \frac{\kappa_\bullet}{\sigma^2 \cdot m} > 0$). The “long-term mean” of the limit process \tilde{X}_s is $\frac{\eta}{\kappa_\bullet} = 2.5$ and is indicated as red line. The “long-term mean” of the approximations $\tilde{X}_s^{(m)}$ is equal to $\frac{\alpha_\bullet^{(m)}}{1 - \beta_\bullet^{(m)}} = \frac{\eta}{\kappa_\bullet} - \frac{\eta}{\sigma^2 \cdot m} = 2.5 - 31.25/m$ and is displayed as green line.*

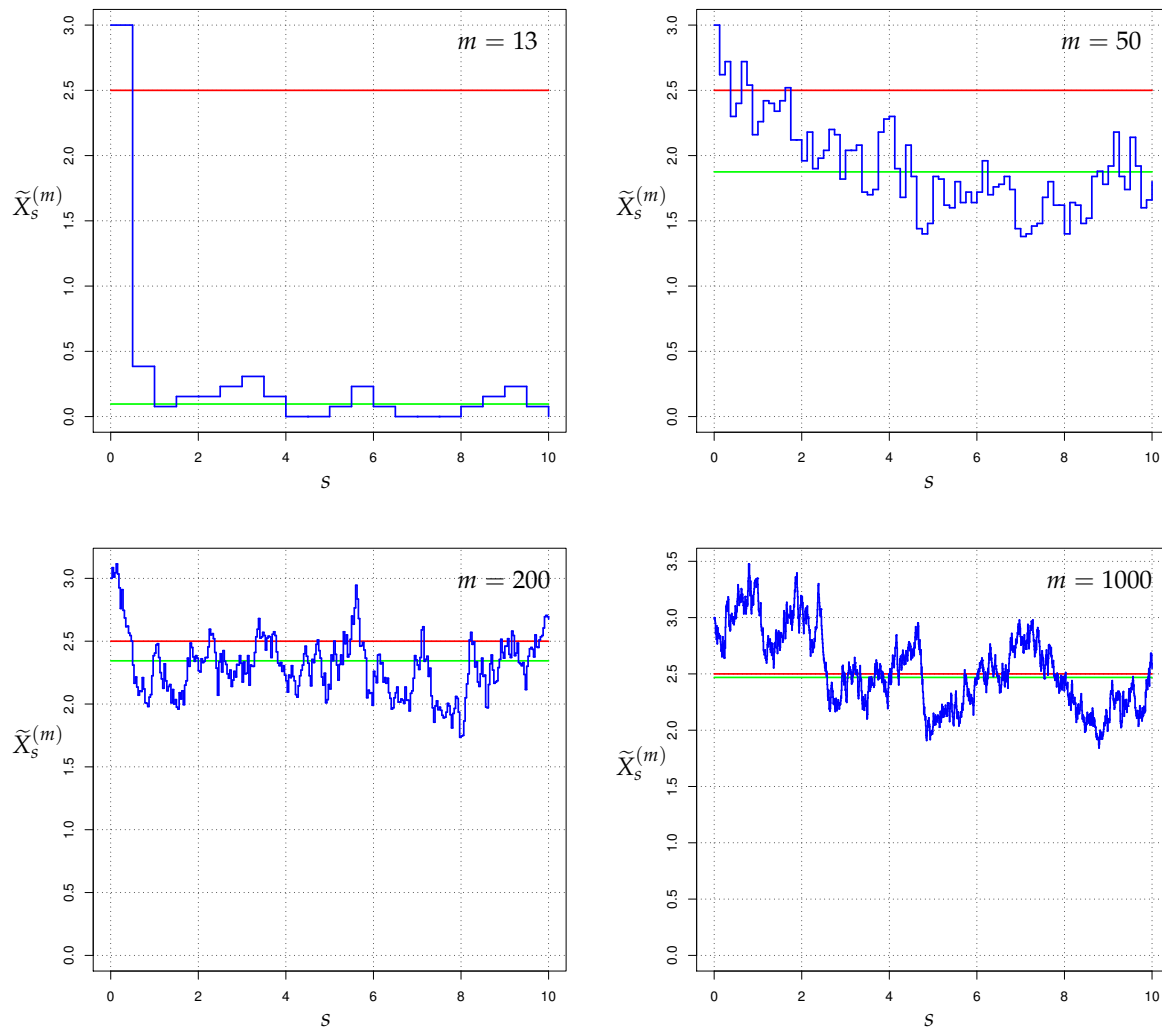


Figure 4. Simulation of the process $\tilde{X}_s^{(m)}$ for the approximation steps $m \in \{13, 50, 200, 1000\}$ in the parameter setup $(\eta, \kappa_\bullet, \sigma) = (5, 2, 0.4)$ and with initial starting value $\tilde{X}_0 = 3$.

7.2. Bounds of Hellinger Integrals for Diffusion Approximations

For each approximation step m and each observation horizon $t \in [0, \infty[$, let us now investigate the behaviour of the Hellinger integrals $H_\lambda \left(P_{\mathcal{A},t}^{(m),CDA} \middle| \middle| P_{\mathcal{H},t}^{(m),CDA} \right)$, where $P_{\bullet,t}^{(m),CDA}$ denotes the canonical law (under \mathcal{H} resp. \mathcal{A}) of the *continuous-time diffusion approximation* $\tilde{X}^{(m)}$ (cf. (132)), restricted to $[0, t]$. It is easy to see that $H_\lambda \left(P_{\mathcal{A},t}^{(m),CDA} \middle| \middle| P_{\mathcal{H},t}^{(m),CDA} \right)$ coincides with $H_\lambda \left(P_{\mathcal{A},[\sigma^2 mt]}^{(m)} \middle| \middle| P_{\mathcal{H},[\sigma^2 mt]}^{(m)} \right)$ of the law restrictions of the GW(I) generations sizes $\left(X_\ell^{(m)} \right)_{\ell \in \{0, \dots, [\sigma^2 mt]\}}$, where $\frac{[\sigma^2 mt]}{\sigma^2 m}$ can be interpreted as the last “jump-time” of $\tilde{X}^{(m)}$ before t . These Hellinger integrals obey the results of

- the Propositions 2 and 3 (for $\eta = 0$) respectively the Propositions 4 and 5 (for $\eta \in]0, \infty[$), as far as recursively computable exact values are concerned,
- Theorem 5 as far as closed-form bounds are concerned; recall that the current setup is of type $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$, and thus we can use the simplifications proposed in the Remark 7(a).

In order to obtain the desired Hellinger integral limits $\lim_{m \rightarrow \infty} H_\lambda \left(P_{\mathcal{A},[\sigma^2 mt]}^{(m)} \middle| \middle| P_{\mathcal{H},[\sigma^2 mt]}^{(m)} \right)$, one faces several technical problems which will be described in the following. To begin with, for fixed

$m \in \overline{\mathbb{N}}$ we apply the Propositions 2(b), 3(b), 4(b), 5(b) to the current setup $(\beta_{\mathcal{A}}^{(m)}, \beta_{\mathcal{H}}^{(m)}, \alpha_{\mathcal{A}}^{(m)}, \alpha_{\mathcal{H}}^{(m)}) \in \mathcal{P}_{\text{NI}} \cup \mathcal{P}_{\text{SP},1}$ with

$$\beta_{\bullet}^{(m)} := \beta_{\bullet}(m, \kappa_{\bullet}, \sigma^2) := 1 - \frac{\kappa_{\bullet}}{\sigma^2 m} \quad \text{and} \quad \alpha_{\bullet}^{(m)} := \alpha_{\bullet}(m, \kappa_{\bullet}, \sigma^2, \eta) := \beta_{\bullet}^{(m)} \cdot \frac{\eta}{\sigma^2} \quad (\text{cf. (131)}).$$

Notice that $\eta = 0$ corresponds to the no-immigration (NI) case and that $\frac{\alpha_{\bullet}^{(m)}}{\beta_{\bullet}^{(m)}} = \frac{\eta}{\sigma^2}$. Accordingly, we set $\alpha_{\lambda}^{(m)} := \lambda \cdot \alpha_{\mathcal{A}}^{(m)} + (1 - \lambda) \cdot \alpha_{\mathcal{H}}^{(m)}$, $\beta_{\lambda}^{(m)} := \lambda \cdot \beta_{\mathcal{A}}^{(m)} + (1 - \lambda) \cdot \beta_{\mathcal{H}}^{(m)}$. By using

$$q_{\lambda}^{(m)} := q(m, \kappa_{\bullet}, \sigma^2, \lambda) := \left(\beta_{\mathcal{A}}^{(m)}\right)^{\lambda} \left(\beta_{\mathcal{H}}^{(m)}\right)^{1-\lambda}, \quad \lambda \in \mathbb{R} \setminus \{0, 1\}, \tag{134}$$

as well as the connected sequence $\left(a_n^{(m)}\right)_{n \in \mathbb{N}} := \left(a_n^{(q_{\lambda}^{(m)})}\right)_{n \in \mathbb{N}}$ we arrive at the

Corollary 13. For all $(\beta_{\mathcal{A}}^{(m)}, \beta_{\mathcal{H}}^{(m)}, \alpha_{\mathcal{A}}^{(m)}, \alpha_{\mathcal{H}}^{(m)}, \lambda) \in (\mathcal{P}_{\text{NI}} \cup \mathcal{P}_{\text{SP},1}) \times (\mathbb{R} \setminus \{0, 1\})$ and all population sizes $X_0^{(m)} \in \mathbb{N}$ there holds

$$h_{\lambda}^{(m)} := H_{\lambda} \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) = \exp \left\{ a_{[\sigma^2 mt]}^{(q_{\lambda}^{(m)})} \cdot X_0^{(m)} + \frac{\eta}{\sigma^2} \sum_{k=1}^{[\sigma^2 mt]} a_k^{(q_{\lambda}^{(m)})} \right\} \tag{135}$$

with $\eta = 0$ in the NI case.

In the following, we employ the SDE-parameter constellations (which are consistent with (131) in combination with our requirement to work here only on $(\mathcal{P}_{\text{NI}} \cup \mathcal{P}_{\text{SP},1})$)

$$\tilde{\mathcal{P}}_{\text{NI}} := \left\{ (\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta), \eta = 0, \kappa_{\mathcal{A}} \in [0, \infty[, \kappa_{\mathcal{H}} \in [0, \infty[, \kappa_{\mathcal{A}} \neq \kappa_{\mathcal{H}} \right\}, \tag{136}$$

$$\tilde{\mathcal{P}}_{\text{SP},1} := \left\{ (\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta), \eta > 0, \kappa_{\mathcal{A}} \in [0, \infty[, \kappa_{\mathcal{H}} \in [0, \infty[, \kappa_{\mathcal{A}} \neq \kappa_{\mathcal{H}} \right\}. \tag{137}$$

Due to the—not in closed-form representable—recursive nature of the sequences $\left(a_n^{(q)}\right)_{n \in \mathbb{N}}$ defined by (36), the calculation of $\lim_{m \rightarrow \infty} h_{\lambda}^{(m)}$ in (135) seems to be not (straightforwardly) tractable; after all, one “has to move along” a sequence of recursions (roughly speaking) since $[\sigma^2 mt] \rightarrow \infty$ as m tends to infinity. One way to “circumvent” such technical problems is to compute instead of the limit $\lim_{m \rightarrow \infty} h_{\lambda}^{(m)}$ of the (exact values of the) Hellinger integrals $h_{\lambda}^{(m)}$, the limits of the corresponding (explicit) closed-form lower resp. upper bounds adapted from Theorem 5. In order to achieve this, one first needs a preparatory step, due to the fact that the sequence $\left(a_{[\sigma^2 mt]}^{(q_{\lambda}^{(m)})}\right)_{m \in \overline{\mathbb{N}}}$ (and hence its bounds leading to closed-form expressions) does not necessarily converge for all $\lambda \in \mathbb{R} \setminus [0, 1]$; roughly, this can be conjectured from the Propositions 3(c) and 5(c) in combination with $[\sigma^2 mt] \rightarrow \infty$. Correspondingly, for our “sequence-of-recursions” context equipped with the diffusion-limit’s drift-parameter constellations $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta)$ we have to derive a “convergence interval” $[\tilde{\lambda}_-, \tilde{\lambda}_+] \setminus [0, 1]$ which replaces the single-recursion-concerning $[\lambda_-, \lambda_+] \setminus [0, 1]$ (cf. Lemma 1). This amounts to

Proposition 15. For all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta) \in \tilde{\mathcal{P}}_{\text{NI}} \cup \tilde{\mathcal{P}}_{\text{SP},1}$ define

$$0 > \tilde{\lambda}_- := \begin{cases} -\infty, & \text{if } \kappa_{\mathcal{A}} < \kappa_{\mathcal{H}}, \\ -\frac{\kappa_{\mathcal{H}}^2}{\kappa_{\mathcal{A}}^2 - \kappa_{\mathcal{H}}^2}, & \text{if } \kappa_{\mathcal{A}} > \kappa_{\mathcal{H}}, \end{cases} \quad \text{and} \quad 1 < \tilde{\lambda}_+ := \begin{cases} \frac{\kappa_{\mathcal{H}}^2}{\kappa_{\mathcal{H}}^2 - \kappa_{\mathcal{A}}^2}, & \text{if } \kappa_{\mathcal{A}} < \kappa_{\mathcal{H}}, \\ \infty, & \text{if } \kappa_{\mathcal{A}} > \kappa_{\mathcal{H}}. \end{cases} \tag{138}$$

Then, for all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta, \lambda) \in (\tilde{\mathcal{P}}_{NI} \cup \tilde{\mathcal{P}}_{SP,1}) \times]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus [0, 1]$ there holds for all sufficiently large $m \in \bar{\mathbb{N}}$

$$q_{\lambda}^{(m)} := \left(1 - \frac{\kappa_{\mathcal{A}}}{\sigma^2 m}\right)^{\lambda} \left(1 - \frac{\kappa_{\mathcal{H}}}{\sigma^2 m}\right)^{1-\lambda} < \min \left\{1, e^{\beta_{\lambda}^{(m)} - 1}\right\}, \tag{139}$$

and thus the sequence $\left(a_n^{(q_{\lambda}^{(m)})}\right)_{n \in \mathbb{N}}$ converges to the fixed point $x_0^{(m)} \in]0, -\log(q_{\lambda}^{(m)})[$.

This will be proved in Appendix A.4.

We are now in the position to determine bounds of the Hellinger integral limits $\lim_{m \rightarrow \infty} H_{\lambda} \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \middle\| \middle\| P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right)$ in form of m -limits of appropriate versions of closed-form bounds

from Section 6. For the sake of brevity, let us henceforth use the abbreviations $x_0^{(m)} := x_0^{(q_{\lambda}^{(m)})}$, $\Gamma_{<}^{(m)} := \Gamma_{<}^{(q_{\lambda}^{(m)})} = \frac{q_{\lambda}^{(m)}}{2} \cdot e^{x_0^{(m)}} \cdot (x_0^{(m)})^2$, $\Gamma_{>}^{(m)} := \Gamma_{>}^{(q_{\lambda}^{(m)})} = \frac{q_{\lambda}^{(m)}}{2} \cdot (x_0^{(m)})^2$, $d^{(m),S} := d^{(q_{\lambda}^{(m)})},S = \frac{x_0^{(m)} - (q_{\lambda}^{(m)} - \beta_{\lambda}^{(m)})}{x_0^{(m)}}$

and $d^{(m),T} := d^{(q_{\lambda}^{(m)})},T = q_{\lambda}^{(m)} \cdot e^{x_0^{(m)}}$. By the above considerations, the Theorem 5 (together with Remark 7(a)) adapts to the current setup as follows:

Corollary 14. (a) For all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta, \lambda) \in (\tilde{\mathcal{P}}_{NI} \cup \tilde{\mathcal{P}}_{SP,1}) \times]0, 1[$, all $t \in [0, \infty[$, all approximation steps $m \in \bar{\mathbb{N}}$ and all initial population sizes $X_0^{(m)} \in \mathbb{N}$ the Hellinger integral can be bounded by

$$\begin{aligned} C_{\lambda, X_0^{(m)}, t}^{(m),L} &:= \exp \left\{ x_0^{(m)} \cdot \left[X_0^{(m)} - \frac{\eta}{\sigma^2} \frac{d^{(m),T}}{1 - d^{(m),T}} \right] \left(1 - (d^{(m),T})^{[\sigma^2 mt]} \right) + x_0^{(m)} \frac{\eta}{\sigma^2} \cdot [\sigma^2 mt] \right. \\ &\quad \left. + \underline{\zeta}_{[\sigma^2 mt]}^{(m)} \cdot X_0^{(m)} + \frac{\eta}{\sigma^2} \cdot \underline{\vartheta}_{[\sigma^2 mt]}^{(m)} \right\} \end{aligned} \tag{140}$$

$$\begin{aligned} &\leq H_{\lambda} \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \middle\| \middle\| P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) \\ &\leq \exp \left\{ x_0^{(m)} \cdot \left[X_0^{(m)} - \frac{\eta}{\sigma^2} \frac{d^{(m),S}}{1 - d^{(m),S}} \right] \left(1 - (d^{(m),S})^{[\sigma^2 mt]} \right) + x_0^{(m)} \frac{\eta}{\sigma^2} \cdot [\sigma^2 mt] \right. \\ &\quad \left. - \bar{\zeta}_{[\sigma^2 mt]}^{(m)} \cdot X_0^{(m)} - \frac{\eta}{\sigma^2} \cdot \bar{\vartheta}_{[\sigma^2 mt]}^{(m)} \right\} =: C_{\lambda, X_0^{(m)}, t}^{(m),U}, \end{aligned} \tag{141}$$

where we define analogously to (98) to (101)

$$\underline{\zeta}_n^{(m)} := \Gamma_{<}^{(m)} \cdot \frac{(d^{(m),T})^{n-1}}{1 - d^{(m),T}} \cdot \left(1 - (d^{(m),T})^n \right) > 0, \tag{142}$$

$$\underline{\vartheta}_n^{(m)} := \Gamma_{<}^{(m)} \cdot \frac{1 - (d^{(m),T})^n}{(1 - d^{(m),T})^2} \cdot \left[1 - \frac{d^{(m),T} (1 + (d^{(m),T})^n)}{1 + d^{(m),T}} \right] > 0, \tag{143}$$

$$\bar{\zeta}_n^{(m)} := \Gamma_{<}^{(m)} \cdot \left[\frac{(d^{(m),S})^n - (d^{(m),T})^n}{d^{(m),S} - d^{(m),T}} - (d^{(m),S})^{n-1} \cdot \frac{1 - (d^{(m),T})^n}{1 - d^{(m),T}} \right] > 0, \tag{144}$$

$$\bar{\vartheta}_n^{(m)} := \Gamma_{<}^{(m)} \cdot \frac{d^{(m),T}}{1 - d^{(m),T}} \cdot \left[\frac{1 - (d^{(m),S} d^{(m),T})^n}{1 - d^{(m),S} d^{(m),T}} - \frac{(d^{(m),S})^n - (d^{(m),T})^n}{d^{(m),S} - d^{(m),T}} \right] > 0. \tag{145}$$

Notice that (140) and (141) simplify significantly for $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta, \lambda) \in \tilde{\mathcal{P}}_{NI} \times]0, 1[$ for which $\eta = 0$ holds.

(b) For all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta, \lambda) \in (\tilde{\mathcal{P}}_{NI} \cup \tilde{\mathcal{P}}_{SP,1}) \times]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus [0, 1]$ and all initial population sizes $X_0^{(m)} \in \mathbb{N}$ the

Hellinger integral bounds (140) and (141) are valid for all sufficiently large $m \in \mathbb{N}$, where the expressions (142) to (145) have to be replaced by

$$\underline{\zeta}_n^{(m)} := \Gamma_{>}^{(m)} \cdot \frac{(d^{(m),T})^n - (d^{(m),S})^{2n}}{d^{(m),T} - (d^{(m),S})^2} > 0, \tag{146}$$

$$\underline{\vartheta}_n^{(m)} := \frac{\Gamma_{>}^{(m)}}{d^{(m),T} - (d^{(m),S})^2} \cdot \left[\frac{d^{(m),T} \cdot (1 - (d^{(m),T})^n)}{1 - d^{(m),T}} - \frac{(d^{(m),S})^2 \cdot (1 - (d^{(m),S})^{2n})}{1 - (d^{(m),S})^2} \right] > 0,$$

$$\underline{\zeta}_n^{(m)} := \Gamma_{>}^{(m)} \cdot (d^{(m),S})^{n-1} \cdot \left[n - \frac{1 - (d^{(m),T})^n}{1 - d^{(m),T}} \right] > 0, \tag{147}$$

$$\overline{\vartheta}_n^{(m)} := \Gamma_{>}^{(m)} \cdot \left[\frac{d^{(m),S} - d^{(m),T}}{(1 - d^{(m),S})^2 (1 - d^{(m),T})} \cdot (1 - (d^{(m),S})^n) \right. \tag{148}$$

$$\left. + \frac{d^{(m),T} (1 - (d^{(m),S} d^{(m),T})^n)}{(1 - d^{(m),T}) (1 - d^{(m),S} d^{(m),T})} - \frac{(d^{(m),S})^n}{1 - d^{(m),S}} \cdot n \right]. \tag{149}$$

Let us finally present the desired assertions on the limits of the bounds given in Corollary 14 as the approximation step m tends to infinity, by employing for $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+ [\supsetneq]0, 1[$ the quantities

$$\kappa_\lambda := \lambda \kappa_{\mathcal{A}} + (1 - \lambda) \kappa_{\mathcal{H}} \quad \text{as well as} \quad \Lambda_\lambda := \sqrt{\lambda \kappa_{\mathcal{A}}^2 + (1 - \lambda) \kappa_{\mathcal{H}}^2}, \tag{150}$$

for which the following relations hold:

$$\Lambda_\lambda > \kappa_\lambda > 0, \quad \text{for } \lambda \in]0, 1[, \tag{151}$$

$$0 < \Lambda_\lambda < \kappa_\lambda, \quad \text{for } \lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+ [\setminus]0, 1[. \tag{152}$$

Theorem 11. *Let the initial SDE-value $\tilde{X}_0 \in]0, \infty[$ be arbitrary but fixed, and suppose that $\lim_{m \rightarrow \infty} \frac{1}{m} X_0^{(m)} = \tilde{X}_0$. Then, for all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta, \lambda) \in (\tilde{\mathcal{P}}_{NI} \cup \tilde{\mathcal{P}}_{SP,1}) \times]\tilde{\lambda}_-, \tilde{\lambda}_+ [\setminus \{0, 1\}$ and all $t \in [0, \infty[$ the Hellinger integral limit can be bounded by*

$$D_{\lambda, \tilde{X}_0, t}^L := \exp \left\{ - \frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \left[\tilde{X}_0 - \frac{\eta}{\Lambda_\lambda} \right] \left(1 - e^{-\Lambda_\lambda \cdot t} \right) - \frac{\eta}{\sigma^2} (\Lambda_\lambda - \kappa_\lambda) \cdot t \right. \\ \left. + L_\lambda^{(1)}(t) \cdot \tilde{X}_0 + \frac{\eta}{\sigma^2} \cdot L_\lambda^{(2)}(t) \right\} \tag{153}$$

$$\leq \lim_{m \rightarrow \infty} H_\lambda \left(P_{\mathcal{A}, [\sigma^2 m t]}^{(m)} \middle| \middle| P_{\mathcal{H}, [\sigma^2 m t]}^{(m)} \right) \\ \leq \exp \left\{ - \frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \left[\tilde{X}_0 - \frac{\eta}{\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)} \right] \left(1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \right) - \frac{\eta}{\sigma^2} (\Lambda_\lambda - \kappa_\lambda) \cdot t \right. \\ \left. - U_\lambda^{(1)}(t) \cdot \tilde{X}_0 - \frac{\eta}{\sigma^2} \cdot U_\lambda^{(2)}(t) \right\} =: D_{\lambda, \tilde{X}_0, t}^U, \tag{154}$$

where for the (sub)case of all $\lambda \in]0, 1[$ and all $t \geq 0$

$$L_\lambda^{(1)}(t) := \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2\sigma^2 \cdot \Lambda_\lambda} \cdot e^{-\Lambda_\lambda \cdot t} \cdot (1 - e^{-\Lambda_\lambda \cdot t}), \tag{155}$$

$$L_\lambda^{(2)}(t) := \frac{1}{4} \cdot \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda}\right)^2 \cdot (1 - e^{-\Lambda_\lambda \cdot t})^2, \tag{156}$$

$$U_\lambda^{(1)}(t) := \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\sigma^2} \cdot \left[\frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda - \kappa_\lambda} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} (1 - e^{-\Lambda_\lambda \cdot t})}{2 \cdot \Lambda_\lambda} \right], \tag{157}$$

$$U_\lambda^{(2)}(t) := \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\Lambda_\lambda} \cdot \left[\frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda) \cdot t}}{3\Lambda_\lambda + \kappa_\lambda} + \frac{e^{-\Lambda_\lambda \cdot t} - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda - \kappa_\lambda} \right], \tag{158}$$

and for the remaining (sub)case of all $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+ [\setminus]0, 1[$ and all $t \geq 0$

$$L_\lambda^{(1)}(t) := \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2\sigma^2 \cdot \kappa_\lambda} \cdot e^{-\Lambda_\lambda \cdot t} \cdot (1 - e^{-\kappa_\lambda \cdot t}), \tag{159}$$

$$L_\lambda^{(2)}(t) := \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2 \cdot \kappa_\lambda} \cdot \left[\frac{1 - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda} - \frac{1 - e^{-(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda + \kappa_\lambda} \right], \tag{160}$$

$$U_\lambda^{(1)}(t) := \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2 \cdot \sigma^2} \cdot e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \cdot \left[t - \frac{1 - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda} \right], \tag{161}$$

$$U_\lambda^{(2)}(t) := (\Lambda_\lambda - \kappa_\lambda)^2 \cdot \left[\frac{(\Lambda_\lambda - \kappa_\lambda) (1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t})}{\Lambda_\lambda \cdot (\Lambda_\lambda + \kappa_\lambda)^2} + \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda \cdot (3\Lambda_\lambda + \kappa_\lambda)} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda + \kappa_\lambda} \cdot t \right]. \tag{162}$$

Notice that the components $L_\lambda^{(i)}(t)$ and $U_\lambda^{(i)}(t)$ (for $i = 1, 2$ and in both cases $\lambda \in]0, 1[$ and $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+ [\setminus]0, 1[$) are strictly positive for $t > 0$ and do not depend on the parameter η . Furthermore, the bounds $D_{\lambda, \tilde{X}_0, t}^L$ and $D_{\lambda, \tilde{X}_0, t}^U$ simplify significantly in the case $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta) \in \tilde{\mathcal{P}}_{NI}$, for which $\eta = 0$ holds.

This will be proved in Appendix A.4. For the time-asymptotics, we obtain the

Corollary 15. Let the initial SDE-value $\tilde{X}_0 \in]0, \infty[$ be arbitrary but fixed, and suppose that $\lim_{m \rightarrow \infty} \frac{1}{m} X_0^{(m)} = \tilde{X}_0$. Then:

(a) For all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta, \lambda) \in \tilde{\mathcal{P}}_{NI} \times]\tilde{\lambda}_-, \tilde{\lambda}_+ [\setminus \{0, 1\}$ the Hellinger integral limit converges to

$$\lim_{t \rightarrow \infty} \lim_{m \rightarrow \infty} \log \left(H_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \middle| \middle| P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) \right) = - \frac{\tilde{X}_0}{\sigma^2} \cdot (\Lambda_\lambda - \kappa_\lambda) \begin{cases} < 0, & \text{for } \lambda \in]0, 1[, \\ > 0, & \text{for } \lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+ [\setminus]0, 1[. \end{cases}$$

(b) For all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta, \lambda) \in \tilde{\mathcal{P}}_{SP,1} \times]\tilde{\lambda}_-, \tilde{\lambda}_+ [\setminus \{0, 1\}$ the Hellinger integral limit possesses the asymptotical behaviour

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \left(\lim_{m \rightarrow \infty} H_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \middle| \middle| P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) \right) = - \frac{\eta}{\sigma^2} \cdot (\Lambda_\lambda - \kappa_\lambda) \begin{cases} < 0, & \text{for } \lambda \in]0, 1[, \\ > 0, & \text{for } \lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+ [\setminus]0, 1[. \end{cases}$$

The assertions of Corollary 15 follow immediately by inspecting the expressions in the exponential of (153) and (154) in combination with (155) to (162).

7.3. Bounds of Power Divergences for Diffusion Approximations

Analogously to Section 4 (see especially Section 4.1), for orders $\lambda \in \mathbb{R} \setminus \{0, 1\}$ all the results of the previous Section 7.2 carry correspondingly over from (limits of) bounds of the Hellinger integrals $H_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right)$ to (limits of) bounds of the total variation distance $V \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right)$ (by virtue of (12)), to (limits of) bounds of the Renyi divergences $R_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right)$ (by virtue of (7)) as well as to (limits of) bounds of the power divergences $I_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right)$ (by virtue of (2)). For the sake of brevity, the—merely repetitive—exact details are omitted. Moreover, by combining the outcoming results on the above-mentioned power divergences with parts of the Bayesian-decision-making context of Section 2.5, we obtain corresponding assertions on (i) the (cf. (21)) *weighted-average* decision risk reduction (weighted-average statistical information measure) about the degree of evidence deg concerning the parameter θ that can be attained by observing the GWI-path \mathcal{X}_n until stage n , as well as (ii) the (cf. (22)) *limit* decision risk reduction (limit statistical information measure).

In the following, let us concentrate on the derivation of the Kullback-Leibler information divergence KL (relative entropy) within the current diffusion-limit framework. Notice that altogether we face two limit procedures simultaneously: by the first limit $\lim_{\lambda \uparrow 1} I_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right)$ we obtain the KL $I \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right)$ for every fixed approximation step $m \in \overline{\mathbb{N}}$; on the other hand, for each fixed $\lambda \in]0, 1[$, the second limit $\lim_{m \rightarrow \infty} I_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right)$ describes the limit of the power divergence – as the sequence of rescaled and continuously interpolated GW(I)’s $\left((\tilde{X}_s^{(m)})_{s \in [0, \infty[} \right)_{m \in \overline{\mathbb{N}}}$ (equipped with probability law $P_{\mathcal{A}, [\sigma^2 mt]}^{(m)}$ resp. $P_{\mathcal{H}, [\sigma^2 mt]}^{(m)}$ up to time $[\sigma^2 mt]$) converges weakly to the continuous-time CIR-type diffusion process $(\tilde{X}_s)_{s \in [0, \infty[}$ (with probability law $\tilde{P}_{\mathcal{A}, t}$ resp. $\tilde{P}_{\mathcal{H}, t}$ up to time t). In Appendix A.4 we shall prove that these two limits can be interchanged:

Theorem 12. *Let the initial SDE-value $\tilde{X}_0 \in]0, \infty[$ be arbitrary but fixed, and suppose that $\lim_{m \rightarrow \infty} \frac{1}{m} X_0^{(m)} = \tilde{X}_0$. Then, for all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta) \in \tilde{\mathcal{P}}_{NI} \cup \tilde{\mathcal{P}}_{SP, 1}$ and all $t \in [0, \infty[$, one gets the Kullback-Leibler information divergence (relative entropy) convergences*

$$\begin{aligned} \lim_{m \rightarrow \infty} I \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) &= \lim_{m \rightarrow \infty} \lim_{\lambda \nearrow 1} I_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) \\ &= \begin{cases} \frac{(\kappa_{\mathcal{A}} - \kappa_{\mathcal{H}})^2}{2\sigma^2 \cdot \kappa_{\mathcal{A}}} \cdot \left[\left(\tilde{X}_0 - \frac{\eta}{\kappa_{\mathcal{A}}} \right) \cdot (1 - e^{-\kappa_{\mathcal{A}} t}) + \eta \cdot t \right], & \text{if } \kappa_{\mathcal{A}} > 0, \\ \frac{\kappa_{\mathcal{H}}^2}{2\sigma^2} \cdot \left[\frac{\eta}{2} \cdot t^2 + \tilde{X}_0 \cdot t \right], & \text{if } \kappa_{\mathcal{A}} = 0, \end{cases} \\ &= \lim_{\lambda \nearrow 1} \lim_{m \rightarrow \infty} I_\lambda \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right). \end{aligned} \tag{163}$$

This immediately leads to the following

Corollary 16. Let the initial SDE-value $\tilde{X}_0 \in]0, \infty[$ be arbitrary but fixed, and suppose that $\lim_{m \rightarrow \infty} \frac{1}{m} X_0^{(m)} = \tilde{X}_0$. Then, the KL limit (163) possesses the following time-asymptotical behaviour:

(a) For all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta) \in \tilde{\mathcal{P}}_{NI}$ (i.e., $\eta = 0$) one gets

$$(i) \quad \text{in the case } \kappa_{\mathcal{A}} > 0 \quad \lim_{t \rightarrow \infty} \lim_{m \rightarrow \infty} I \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) = \frac{\tilde{X}_0 \cdot (\kappa_{\mathcal{A}} - \kappa_{\mathcal{H}})^2}{2\sigma^2 \cdot \kappa_{\mathcal{A}}},$$

$$(ii) \quad \text{in the case } \kappa_{\mathcal{A}} = 0 \quad \lim_{t \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{1}{t} \cdot I \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) = \frac{\tilde{X}_0 \cdot \kappa_{\mathcal{H}}^2}{4\sigma^2}.$$

(b) For all $(\kappa_{\mathcal{A}}, \kappa_{\mathcal{H}}, \eta) \in \tilde{\mathcal{P}}_{SP,1}$ (i.e., $\eta > 0$) one gets

$$(i) \quad \text{in the case } \kappa_{\mathcal{A}} > 0 \quad \lim_{t \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{1}{t} \cdot I \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) = \frac{\eta \cdot (\kappa_{\mathcal{A}} - \kappa_{\mathcal{H}})^2}{2\sigma^2 \cdot \kappa_{\mathcal{A}}},$$

$$(ii) \quad \text{in the case } \kappa_{\mathcal{A}} = 0 \quad \lim_{t \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{1}{t^2} \cdot I \left(P_{\mathcal{A}, [\sigma^2 mt]}^{(m)} \parallel P_{\mathcal{H}, [\sigma^2 mt]}^{(m)} \right) = \frac{\eta \cdot \kappa_{\mathcal{H}}^2}{4\sigma^2}.$$

Remark 9. In Appendix A.4 we shall see that the proof of the last (limit-interchange concerning) equality in (163) relies heavily on the use of the extra terms $L_{\lambda}^{(1)}(t)$, $L_{\lambda}^{(2)}(t)$, $U_{\lambda}^{(1)}(t)$, $U_{\lambda}^{(2)}(t)$ in (153) and (154). Recall that these terms ultimately stem from (manipulations of) the corresponding parts of the “improved closed-form bounds” in Theorem 5, which were derived by using the linear inhomogeneous difference equations $\underline{a}_n^{(q)}$ resp. $\bar{a}_n^{(q)}$ (cf. (92) resp. (94)) instead of the linear homogeneous difference equations $a_n^{(q),T}$ resp. $a_n^{(q),S}$ (cf. (78) resp. (79)) as explicit approximates of the sequence $a_n^{(q)}$. Not only this fact shows the importance of this more tedious approach.

Interesting comparisons of the above-mentioned results in Sections 7.2 and 7.3 with corresponding information measures of the solutions of the SDE (129) themselves (rather their branching approximations), can be found in Kammerer [157].

7.4. Applications to Decision Making

Analogously to Section 6.7, the above-mentioned investigations of the Sections 7.1–7.3 can be applied to the context of Section 2.5 on dichotomous decision making about GW(I)-type diffusion approximations of solutions of the stochastic differential Equation (129). For the sake of brevity, the—merely repetitive—exact details are omitted.

Author Contributions: Conceptualization, N.B.K. and W.S.; Formal analysis, N.B.K. and W.S.; Methodology, N.B.K. and W.S.; Visualization, N.B.K.; Writing, N.B.K. and W.S. All authors have read and agreed to the published version of the manuscript.

Funding: Niels B. Kammerer received a scholarship of the “Studienstiftung des Deutschen Volkes” for his PhD Thesis.

Acknowledgments: We are very grateful to the referees for their patience to review this long manuscript, and for their helpful suggestions. Moreover, we would like to thank Andreas Greven for some useful remarks.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proofs and Auxiliary Lemmas

Appendix A.1. Proofs and Auxiliary Lemmas for Section 3

Lemma A1. For all real numbers $x, y, z > 0$ and all $\lambda \in \mathbb{R}$ one has

$$x^\lambda y^{1-\lambda} - (\lambda x z^{\lambda-1} + (1-\lambda) y z^\lambda) \begin{cases} \leq 0, & \text{for } \lambda \in]0, 1[, \\ = 0, & \text{for } \lambda \in \{0, 1\} , \\ \geq 0, & \text{for } \lambda \in \mathbb{R} \setminus]0, 1[, \end{cases}$$

with equality in the cases $\lambda \in \mathbb{R} \setminus \{0, 1\}$ iff $\frac{x}{y} = z$.

Proof of Lemma A1. For fixed $\tilde{x} := xz^{\lambda-1} > 0, \tilde{y} := yz^\lambda > 0$ with $\tilde{x} \neq \tilde{y}$ we inspect the function g on \mathbb{R} defined by $g(\lambda) := \tilde{x}^\lambda \tilde{y}^{1-\lambda} - (\lambda \tilde{x} + (1-\lambda)\tilde{y})$ which satisfies $g(0) = g(1) = 0, g'(0) = \tilde{y} \log(\tilde{x}/\tilde{y}) - (\tilde{x} - \tilde{y}) < \tilde{y}((\tilde{x}/\tilde{y}) - 1) - (\tilde{x} - \tilde{y}) = 0$ and which is strictly convex. Thus, the assertion follows immediately by taking into account the obvious case $\tilde{x} = \tilde{y}$. \square

Proof of Properties 1. Property (P9) is trivially valid. To show (P1) we assume $0 < q < \beta_\lambda$, which implies $a_1^{(q)} = \zeta_\lambda^{(q)}(0) = q - \beta_\lambda < 0$. By induction, $(a_n)_{n \in \mathbb{N}}$ is strictly negative and strictly decreasing. As stated in (P9), the function $\zeta_\lambda^{(q)}$ is strictly increasing, strictly convex and converges to $-\beta_\lambda$ for $x \rightarrow -\infty$. Thus, it hits the straight line $id(x) = x$ once and only once on the negative real line at $x_0^{(q)} \in]-\beta_\lambda, 0[$ (cf. (44)). This implies that the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ converges to $x_0^{(q)} \in]-\beta_\lambda, q - \beta_\lambda[$. Property (P2) follows immediately. In order to prove (P3), let us fix $q > \max\{0, \beta_\lambda\}$, implying $a_1^{(q)} = \zeta_\lambda^{(q)}(0) = q - \beta_\lambda > 0$; notice that in this setup, the special choice $q = 1$ implies $\min\{1, e^{\beta_\lambda-1}\} = e^{\beta_\lambda-1} < q$. By induction, $(a_n^{(q)})_{n \in \mathbb{N}}$ is strictly positive and strictly increasing. Since $\lim_{x \rightarrow \infty} \zeta_\lambda^{(q)}(x) = \infty$, the function $\zeta_\lambda^{(q)}$ does not necessarily hit the straight line $id(x) = x$ on the positive real line. In fact, due to strict convexity (cf. (P9)), this is excluded if $\zeta_\lambda^{(q)'}(0) = q \geq 1$. Suppose that $q < 1$. To prove that there exists a positive solution of the equation $\zeta_\lambda^{(q)}(x) = x$ it is sufficient to show that the unique global minimum of the strict convex function $h_\lambda^{(q)}(x) := \zeta_\lambda^{(q)}(x) - x$ is taken at some point $x_0 \in]0, \infty[$ and that $h_\lambda^{(q)}(x_0) \leq 0$. It holds $h_\lambda^{(q)'}(x) = q \cdot e^x - 1$, and therefore $h_\lambda^{(q)'}(x) = 0$ iff $x = x_0 = -\log q$. We have $h_\lambda^{(q)}(-\log q) = 1 - \beta_\lambda + \log q$, which is less or equal to zero iff $q \leq e^{\beta_\lambda-1}$. It remains to show that for $q > \beta_\lambda$ and $q > \min\{1, e^{\beta_\lambda-1}\}$ the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ grows faster than exponentially, i.e., there do not exist constants $c_1, c_2 \in \mathbb{R}$ such that $a_n^{(q)} \leq e^{c_1+c_2n}$ for all $n \in \mathbb{N}$. We already know that (in the current case) $a_n^{(q)} \xrightarrow{n \rightarrow \infty} \infty$. Notice that it is sufficient to verify $\limsup_{n \rightarrow \infty} (\log(a_{n+1}^{(q)}) - \log(a_n^{(q)})) = \infty$. For the case $\beta_\lambda \geq 0$ the latter is obtained by

$$\begin{aligned} \log(a_{n+1}^{(q)}) - \log(a_n^{(q)}) &= \log\left((q - \beta_\lambda)e^{a_n^{(q)}} + \beta_\lambda(e^{a_n^{(q)}} - 1)\right) - \log\left(qe^{a_{n-1}^{(q)}} - \beta_\lambda\right) \\ &\geq (\log(q - \beta_\lambda) - \log(q)) + \left(qe^{a_{n-1}^{(q)}} - \beta_\lambda - a_{n-1}^{(q)}\right) \xrightarrow{a_{n-1}^{(q)} \rightarrow \infty} \infty. \end{aligned}$$

An analogous consideration works out for the case $\beta_\lambda < 0$. Property (P4) is trivial, and (P5) to (P8) are direct implications of the already proven properties (P1) to (P4). \square

Proof of Lemma 1. (a) Let $\beta_A > 0, \beta_H > 0$ with $\beta_A \neq \beta_H, \lambda \in \mathbb{R} \setminus]0, 1[$, $\beta_\lambda := \lambda\beta_A + (1 - \lambda)\beta_H$ and $q_\lambda := \beta_A^\lambda \beta_H^{1-\lambda} > \max\{0, \beta_\lambda\}$ (cf. Lemma A1). Below, we follow the lines of Linkov & Lunyova [53], appropriately adapted to our context. We have to find those $\lambda \in \mathbb{R} \setminus]0, 1[$ for which the following two conditions hold:

- (i) $q_\lambda \leq 1$, i.e., $\xi_\lambda^{(q_\lambda)'}(0) \leq 1$,
- (ii) $q_\lambda \leq e^{\beta_\lambda - 1}$ (cf. (P3a)), which is equivalent with the existence of a positive, if (i) is satisfied, solution of the equation $\xi_\lambda^{(q_\lambda)}(x) = x$.

Notice that the case $q_\lambda = 1, \lambda \in \mathbb{R} \setminus [0, 1]$, cannot appear in (i), provided that (ii) holds (since due to Lemma A1 $e^{\beta_\lambda - 1} < e^{q_\lambda - 1} = 1$). For (i), it is easy to check that we have to require

$$\lambda \begin{cases} < \frac{\log(\beta_H)}{\log(\beta_H/\beta_A)}, & \text{if } \beta_A > \beta_H, \\ > \frac{\log(\beta_H)}{\log(\beta_H/\beta_A)}, & \text{if } \beta_A < \beta_H. \end{cases} \tag{A1}$$

To proceed, straightforward analysis leads to $-\log(q_\lambda) = \arg \min_{x \in \mathbb{R}} \{\xi_\lambda^{(q_\lambda)}(x) - x\}$. To check (ii), we first notice that $q_\lambda \leq e^{\beta_\lambda - 1}$ iff $\xi_\lambda^{(q_\lambda)}(x) - x \leq 0$ for some $x \in \mathbb{R}$. Hence, we calculate

$$\begin{aligned} \xi_\lambda^{(q_\lambda)}(-\log(q_\lambda)) + \log(q_\lambda) \leq 0 &\iff 1 - \lambda(\beta_A - \beta_H) - \beta_H + \lambda \log\left(\frac{\beta_A}{\beta_H}\right) + \log(\beta_H) \leq 0 \\ \iff \lambda \cdot \left[\beta_H \left(1 - \frac{\beta_A}{\beta_H}\right) + \log\left(\frac{\beta_A}{\beta_H}\right) \right] &\leq \beta_H - 1 - \log(\beta_H). \end{aligned} \tag{A2}$$

In order to isolate λ in (A2), one has to find out for which (β_A, β_H) the term in the square bracket is positive resp. zero resp. negative. To achieve this, we aim for the substitutions $x := \beta_A/\beta_H, \beta = \beta_H$ and thus study first the auxiliary function $h_\beta(x) := \log(x) - \beta(x - 1), x > 0$, with fixed parameters $\beta > 0$. Straightforwardly, we obtain $h'_\beta(x) = x^{-1} - \beta$ and $h''_\beta(x) = -x^{-2}$. Thus, the function $h_\beta(\cdot)$ is strictly concave and attains a maximum at $x = \beta^{-1}$. Since additionally $h_\beta(1) = 0$ and $h'_\beta(1) = 1 - \beta$, there exists a second solution $z(\beta) \neq 1$ of the equation $h_\beta(x) = 0$ iff $\beta \neq 1$. Thus, one gets

- for $\beta = 1$: for all $x > 0$ there holds $h_\beta(x) \leq 0$, with equality iff $x = \beta^{-1}$,
- for $\beta < 1$: $h_\beta(x) \geq 0$ iff $x \in [1, z(\beta)]$, with equality iff $x \in \{1, z(\beta)\}$ (notice that $z(\beta) > 1$),
- for $\beta > 1$: $h_\beta(x) \geq 0$ iff $x \in [z(\beta), 1]$, with equality iff $x \in \{z(\beta), 1\}$ (notice that $z(\beta) < 1$).

Suppose that $\lambda < 0$.

Case 1: If $\beta_H = 1$, then condition (ii) is not satisfied whenever $\beta_A \neq \beta_H$, since the right side of (A2) is equal to zero and the left side is strictly greater than zero. Hence, $\lambda_- = 0$.

Case 2: Let $\beta_H > 1$. If $\beta_A < \beta_H$, then condition (i) is not satisfied and hence $\lambda_- = 0$. If $\beta_A > \beta_H$, then condition (i) is satisfied iff $\lambda < \check{\lambda} := \check{\lambda}(\beta_A, \beta_H) := \frac{\log(\beta_H)}{\log(\beta_H/\beta_A)} < 0$. On the other hand, incorporating the discussion of the function $h_\beta(\cdot)$, we see that $h_{\beta_H}\left(\frac{\beta_A}{\beta_H}\right) < 0$. Thus, (A2) implies that condition (ii) is satisfied when $\lambda \geq \check{\lambda} := \check{\lambda}(\beta_A, \beta_H) := \frac{\beta_H - 1 - \log(\beta_H)}{\beta_H - \beta_A + \log\left(\frac{\beta_A}{\beta_H}\right)}$. We claim that $\check{\lambda} < \check{\lambda}$ and conclude that the conditions (i) and (ii) are not fulfilled jointly, which leads to $\lambda_- = 0$. To see this, we notice that due to $1 < \beta_H < \beta_A$ we get $\log(\beta_A)/(\beta_A - 1) < \log(\beta_H)/(\beta_H - 1)$ and thus

$$\begin{aligned}
 & \log(\beta_A)(\beta_H - 1) < \log(\beta_H)(\beta_A - 1) \\
 \iff & \beta_H \log(\beta_H) - \beta_A \log(\beta_H) < \beta_H \log(\beta_H) - \beta_H \log(\beta_A) - \log(\beta_H) + \log(\beta_A) \\
 \iff & \log(\beta_H)(\beta_H - \beta_A) + \log(\beta_H) \log\left(\frac{\beta_A}{\beta_H}\right) < \log\left(\frac{\beta_H}{\beta_A}\right)(\beta_H - 1) + \log(\beta_H) \log\left(\frac{\beta_A}{\beta_H}\right) \\
 \iff & \frac{\log(\beta_H)}{\log\left(\frac{\beta_H}{\beta_A}\right)} < \frac{\beta_H - 1 - \log(\beta_H)}{\beta_H - \beta_A + \log\left(\frac{\beta_A}{\beta_H}\right)} \iff \check{\lambda} < \check{\lambda}. \tag{A3}
 \end{aligned}$$

Case 3: Let $\beta_H < 1$. For this, one gets $h_{\beta_H}\left(\frac{\beta_A}{\beta_H}\right) \geq 0$ for $\beta_A \in]\beta_H, \beta_H z(\beta_H)]$. Hence, condition (ii) is satisfied if either $\beta_A \in]\beta_H, \beta_H z(\beta_H)]$, or $\beta_A \notin]\beta_H, \beta_H z(\beta_H)]$ and $\lambda \geq \check{\lambda}$. If $\beta_A > \beta_H z(\beta_H)$, then condition (i) is trivially satisfied for all $\lambda < 0$. In the case $\beta_A < \beta_H$, condition (i) is satisfied whenever $\lambda > \check{\lambda}$. Notice that since $0 < \beta_A < \beta_H < 1$, an analogous consideration as in (A3) leads to $\check{\lambda} < \check{\lambda}$. This implies that $\lambda_- = \check{\lambda}$. The last case $\beta_A \in]\beta_H, \beta_H z(\beta_H)]$ is easy to handle: since $\frac{\log(\beta_H)}{\log(\beta_H/\beta_A)} > 0$ as well as $z_{\beta_H}\left(\frac{\beta_A}{\beta_H}\right) > 0$, both conditions (i) and (ii) hold trivially.

The representation of λ_+ follows straightforwardly from the λ_- -result and the skew symmetry (8), by employing $1 - \check{\lambda}(\beta_H, \beta_A) = \check{\lambda}(\beta_A, \beta_H)$. Alternatively, one can proceed analogously to the λ_- -case.

Part (b) is much easier to prove: if $\beta_\bullet := \beta_A = \beta_H > 0$, then for all $\lambda \in \mathbb{R} \setminus [0, 1]$ one gets $q_\lambda = \beta_A^\lambda \beta_H^{1-\lambda} = \beta_\bullet$ as well as $\beta_\lambda = \beta_\bullet$. Hence, Properties 1 (P2) implies that $a_n^{(q_\lambda)} \equiv 0$ and thus it is convergent, independently of the choice $\lambda \in \mathbb{R} \setminus [0, 1]$. \square

Proof of Formula (51). For the parameter constellation in Section 3.10, we employ as upper bound for $\phi_\lambda(x)$ ($x \in \mathbb{N}_0$) the function

$$\overline{\phi}_\lambda(x) := \begin{cases} \phi_\lambda(0), & \text{if } x = 0, \\ 0, & \text{if } x > 0. \end{cases}$$

Notice that this method is rather crude, and gives in the other cases treated in the Sections 3.7–3.9 worse bounds than those derived there. Since $\lambda \in]0, 1[$ and $\alpha_A \neq \alpha_H$, one has $\phi_\lambda(0) < 0$. In order to derive an upper bound of the Hellinger integral, we first set $\bar{\epsilon} := 1 - e^{\phi_\lambda(0)} \in]0, 1[$. Hence, for all $n \in \mathbb{N} \setminus \{1\}$ we obtain the auxiliary expression

$$\begin{aligned}
 \sum_{x_{n-1}=0}^{\infty} \frac{[\varphi_\lambda(x_{n-2})]^{x_{n-1}}}{x_{n-1}!} \cdot \exp\{\phi_\lambda(x_{n-1})\} & \leq \sum_{x_{n-1}=0}^{\infty} \frac{[\varphi_\lambda(x_{n-2})]^{x_{n-1}}}{x_{n-1}!} \cdot \exp\{\overline{\phi}_\lambda(x_{n-1})\} \\
 & = \exp\{\varphi_\lambda(x_{n-2})\} - \bar{\epsilon} = \exp\{\varphi_\lambda(x_{n-2})\} \cdot [1 - \bar{\epsilon} \cdot \exp\{-\varphi_\lambda(x_{n-2})\}].
 \end{aligned}$$

Moreover, since $\beta_A \neq \beta_H$, one gets $\lim_{x \rightarrow \infty} \phi_\lambda(x) = -\infty$ (cf. Properties 3 (P20) and Lemma A1). This—together with the nonnegativity of $\varphi_\lambda(\cdot)$ —implies

$$\sup_{x \in \mathbb{N}_0} \left\{ \exp\{\phi_\lambda(x)\} \cdot [1 - \bar{\epsilon} \cdot \exp\{-\varphi_\lambda(x)\}] \right\} =: \bar{\delta} \in]0, 1[.$$

Incorporating these considerations as well as the formulas (27) to (32), we get for $n = 1$ the relation $H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) = \exp\{\phi_\lambda(x_0)\} \leq 1$ (with equality iff $x_0 = x^* = \frac{\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}}}{\beta_{\mathcal{H}} - \beta_{\mathcal{A}}}$), and—as a continuation of formula (29)— for all $n \in \mathbb{N} \setminus \{1\}$ (recall that $\vec{x} := (x_0, x_1, \dots) \in \Omega$)

$$\begin{aligned}
 H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) &= \sum_{x_1=0}^\infty \cdots \sum_{x_n=0}^\infty \prod_{k=1}^n Z_{n,k}^{(\lambda)}(\vec{x}) \\
 &= \sum_{x_1=0}^\infty \cdots \sum_{x_{n-1}=0}^\infty \prod_{k=1}^{n-1} Z_{n,k}^{(\lambda)}(\vec{x}) \\
 &\quad \cdot \exp\left\{ (f_{\mathcal{A}}(x_{n-1}))^\lambda (f_{\mathcal{H}}(x_{n-1}))^{(1-\lambda)} - (\lambda f_{\mathcal{A}}(x_{n-1}) + (1-\lambda)f_{\mathcal{H}}(x_{n-1})) \right\} \\
 &= \sum_{x_1=0}^\infty \cdots \sum_{x_{n-2}=0}^\infty \prod_{k=1}^{n-2} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot \exp\{-f_\lambda(x_{n-2})\} \sum_{x_{n-1}=0}^\infty \frac{[\varphi_\lambda(x_{n-2})]^{x_{n-1}}}{x_{n-1}!} \cdot \exp\{\phi_\lambda(x_{n-1})\} \\
 &\leq \sum_{x_1=0}^\infty \cdots \sum_{x_{n-2}=0}^\infty \prod_{k=1}^{n-2} Z_{n,k}^{(\lambda)}(\vec{x}) \cdot \exp\{\phi_\lambda(x_{n-2})\} \cdot [1 - \bar{\epsilon} \cdot \exp\{-\phi_\lambda(x_{n-2})\}] \\
 &\leq \bar{\delta} \cdot \sum_{x_1=0}^\infty \cdots \sum_{x_{n-2}=0}^\infty \prod_{k=1}^{n-2} Z_{n,k}^{(\lambda)}(\vec{x}) \leq \dots \leq \bar{\delta}^{\lfloor n/2 \rfloor}. \tag{A4}
 \end{aligned}$$

Hence, $H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) < 1$ for (at least) all $n \in \mathbb{N} \setminus \{1\}$, and $\lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) = 0$. \square

Notice that the above proof method of formula (51) does not work for the parameter setup in Section 3.11, because there one gets $\bar{\delta} = \sup_{x \in \mathbb{N}_0} \left\{ \exp\{\phi_\lambda(x)\} \cdot [1 - \bar{\epsilon} \cdot \exp\{-\phi_\lambda(x)\}] \right\} = 1$.

Proof of Proposition 9. In the setup $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{\text{SP},4a} \times]0, 1[$ we require $\beta_\bullet := \beta_{\mathcal{A}} = \beta_{\mathcal{H}} < 1$. As a linear upper bound for $\phi_\lambda(\cdot)$, we employ the tangent line at $y \geq 0$ (cf. (52))

$$\phi_{\lambda,y}^{\text{tan}}(x) := (p_y - \alpha_\lambda) + (q_y - \beta_\bullet) \cdot x := (p_{\lambda,y}^{\text{tan}} - \alpha_\lambda) + (q_{\lambda,y}^{\text{tan}} - \beta_\lambda) \cdot x := (\phi_\lambda(y) - y \cdot \phi'_\lambda(y)) + \phi'_\lambda(y) \cdot x. \tag{A5}$$

Since in the current setup $\mathcal{P}_{\text{SP},4a}$ the function $\phi_\lambda(\cdot)$ is strictly increasing, the slope $\phi'_\lambda(y)$ of the tangent line at y is positive. Thus we have $q_y > \beta_\lambda$ and Properties 1 (P3) implies that the sequence $(a_n^{(q_y)})_{n \in \mathbb{N}}$ is strictly increasing and converges to $x_0^{(q_y)} \in]0, -\log(q_y)]$ iff $q_y \leq \min\{1, e^{\beta_\bullet - 1}\} = e^{\beta_\bullet - 1} < 1$ (cf. (P3a)), where $x_0^{(q_y)}$ is the smallest solution of the equation $\zeta_\lambda^{(q_y)}(x) = q_y \cdot e^x - \beta_\bullet = x$. Since $q_y \searrow \beta_\bullet$ for $y \rightarrow \infty$ (cf. Properties 3 (P18)) and additionally $e^{\beta_\bullet - 1} > \beta_\bullet$, there exists a large enough $y \geq 0$ such that the sequence $(a_n^{(q_y)})_{n \in \mathbb{N}}$ converges. If this y is also large enough to additionally guarantee $h(y) < 0$ for

$$h(y) := \lim_{n \rightarrow \infty} \frac{1}{n} \log(\tilde{B}_{\lambda, X_0, n}^{(p_y, q_y)}) = p_y \cdot e^{x_0^{(q_y)}} - \alpha_\lambda,$$

then one can conclude that $\lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) = 0$. As a first step, for verifying $h(y) < 0$ we look for an upper bound $\bar{x}_0^{(q_y)}$ for the fixed point $x_0^{(q_y)}$ where the latter exists for $y \geq y_1$ (say). Notice that

$$\bar{Q}_\lambda^{(q_y)}(x) := \frac{1}{2}x^2 + q_y x + q_y - \beta_\bullet \geq q_y \cdot e^x - \beta_\bullet = \zeta_\lambda^{(q_y)}(x), \tag{A6}$$

since $\bar{Q}_\lambda^{(q_y)}(0) = \zeta_\lambda^{(q_y)}(0)$, $\bar{Q}_\lambda^{(q_y)'}(0) = \zeta_\lambda^{(q_y)'}(0)$ and $\bar{Q}_\lambda^{(q_y)''}(x) \geq \zeta_\lambda^{(q_y)''}(x)$ for $x \in [0, -\log(q_y)]$. For sufficiently large $y \geq y_2 \geq y_1$ (say), we easily obtain the smaller solution of $\bar{Q}_\lambda^{(q_y)}(x) = x$ as

$$\bar{x}_0^{(q_y)} = (1 - q_y) - \sqrt{(1 - q_y)^2 - 2(q_y - \beta_\bullet)} = (1 - \phi'_\lambda(y) - \beta_\bullet) - \sqrt{(1 - \phi'_\lambda(y) - \beta_\bullet)^2 - 2\phi'_\lambda(y)} \geq x_0^{(q_y)} \tag{A7}$$

where the expression in the root is positive since $q_y \searrow \beta_\bullet$ for $y \rightarrow \infty$. We now have

$$h(y) = p_y \cdot e^{\bar{x}_0^{(q_y)}} - \alpha_\lambda \leq p_y \cdot e^{\bar{x}_0^{(q_y)}} - \alpha_\lambda =: \bar{h}(y), \quad \forall y \geq y_2. \tag{A8}$$

Hence, it suffices to show that $\bar{h}(y) < 0$ for some $y \geq y_2$. We recall from Properties 3 (P15), (P17) and (P19) that

$$\begin{aligned} \phi_\lambda(y) &= (\alpha_{\mathcal{A}} + \beta_\bullet \cdot y)^\lambda (\alpha_{\mathcal{H}} + \beta_\bullet \cdot y)^{1-\lambda} - \lambda (\alpha_{\mathcal{A}} + \beta_\bullet \cdot y) - (1-\lambda) (\alpha_{\mathcal{H}} + \beta_\bullet \cdot y) < 0, \\ \phi'_\lambda(y) &= \lambda \cdot \beta_\bullet \cdot \left(\frac{\alpha_{\mathcal{A}} + \beta_\bullet \cdot y}{\alpha_{\mathcal{H}} + \beta_\bullet \cdot y}\right)^{\lambda-1} + (1-\lambda) \cdot \beta_\bullet \cdot \left(\frac{\alpha_{\mathcal{A}} + \beta_\bullet \cdot y}{\alpha_{\mathcal{H}} + \beta_\bullet \cdot y}\right)^\lambda - \beta_\bullet > 0 \quad \text{and that} \\ \phi''_\lambda(y) &= - \left(\frac{\alpha_{\mathcal{A}} + \beta_\bullet \cdot y}{\alpha_{\mathcal{H}} + \beta_\bullet \cdot y}\right)^\lambda \cdot \frac{\lambda(1-\lambda) \cdot \beta_\bullet^2 \cdot (\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}})^2}{(\alpha_{\mathcal{A}} + \beta_\bullet \cdot y)^2 (\alpha_{\mathcal{H}} + \beta_\bullet \cdot y)} < 0, \end{aligned} \tag{A9}$$

which immediately implies $\lim_{y \rightarrow \infty} \phi_\lambda(y) = \lim_{y \rightarrow \infty} \phi'_\lambda(y) = \lim_{y \rightarrow \infty} \phi''_\lambda(y) = 0$ and with l'Hospital's rule

$$\begin{aligned} \lim_{y \rightarrow \infty} y \cdot \phi_\lambda(y) &= \lim_{y \rightarrow \infty} -y^2 \cdot \phi'_\lambda(y) = \lim_{y \rightarrow \infty} \frac{y^3}{2} \cdot \phi''_\lambda(y) \\ &= -\frac{1}{2} \lim_{y \rightarrow \infty} \left(\frac{\alpha_{\mathcal{A}} + \beta_\bullet \cdot y}{\alpha_{\mathcal{H}} + \beta_\bullet \cdot y}\right)^\lambda \cdot \frac{\lambda(1-\lambda) \cdot \beta_\bullet^2 \cdot (\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}})^2}{(\alpha_{\mathcal{A}}/y + \beta_\bullet)^2 (\alpha_{\mathcal{H}}/y + \beta_\bullet)} = -\frac{1}{2} \lambda(1-\lambda) \cdot \frac{(\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}})^2}{\beta_\bullet}. \end{aligned} \tag{A10}$$

The formulas (A5), (A7) and (A9) imply the limits $\lim_{y \rightarrow \infty} p_y = \alpha_\lambda$, $\lim_{y \rightarrow \infty} q_y = \beta_\bullet$, $\lim_{y \rightarrow \infty} \bar{x}_0^{(q_y)} = 0$. Notice that $p_y < \alpha_\lambda$ holds trivially for all $y \geq 0$ since the intercept $(p_y - \alpha_\lambda)$ of the tangent line $\phi_{\lambda,y}^{\tan}(\cdot)$ is negative. Incorporating (A8) we therefore obtain $\lim_{y \rightarrow \infty} h(y) \leq \lim_{y \rightarrow \infty} \bar{h}(y) = 0$. As mentioned before, for the proof it is sufficient to show that $\bar{h}(y) < 0$ for some $y \geq y_2$. This holds true if $\lim_{y \rightarrow \infty} y \cdot \bar{h}(y) < 0$. To verify this, notice first that from (A5), (A7) and (A8) we get

$$\bar{h}'(y) = -p_y \cdot e^{\bar{x}_0^{(q_y)}} \cdot \phi''_\lambda(y) \cdot \left[1 - \frac{2 - \phi'_\lambda(y) - \beta_\bullet}{\sqrt{(1 - q_y)^2 - 2(q_y - \beta_\bullet)}}\right] - y \cdot \phi''_\lambda(y) \cdot e^{\bar{x}_0^{(q_y)}} \xrightarrow{y \rightarrow \infty} 0. \tag{A11}$$

Finally we obtain with (A10)

$$\begin{aligned} \lim_{y \rightarrow \infty} y \cdot \bar{h}(y) &= - \lim_{y \rightarrow \infty} y^2 \cdot \bar{h}'(y) \\ &= \lim_{y \rightarrow \infty} p_y \cdot e^{\bar{x}_0^{(q_y)}} \cdot y^2 \cdot \phi''_\lambda(y) \cdot \left[1 - \frac{2 - \phi'_\lambda(y) - \beta_\bullet}{\sqrt{(1 - q_y)^2 - 2(q_y - \beta_\bullet)}}\right] + y^3 \cdot \phi''_\lambda(y) \cdot e^{\bar{x}_0^{(q_y)}} \\ &= 0 - \lambda(1-\lambda) \cdot \frac{(\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}})^2}{\beta_\bullet} < 0. \quad \square \end{aligned}$$

Proof of Corollary 1. Part (a) follows directly from Proposition 1 (a),(b) and the limit $\lim_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n}) = 0$ in the respective part (c) of the Propositions 7, 8, 9 as well as from (51). To prove part (b), according to (26) we have to verify $\liminf_{\lambda \nearrow 1} \{\liminf_{n \rightarrow \infty} H_\lambda(P_{\mathcal{A},n} || P_{\mathcal{H},n})\} = 1$. From part (c) of Proposition 2 we see that this is satisfied iff $\lim_{\lambda \uparrow 1} x_0^{(q_\lambda^E)} = 0$. Recall that for fixed $\lambda \in]0, 1[$ we have $\beta_\lambda = \lambda \beta_{\mathcal{A}} + (1-\lambda) \beta_{\mathcal{H}} > 0$, $q_\lambda^E = \beta_{\mathcal{A}}^\lambda \beta_{\mathcal{H}}^{1-\lambda} < \beta_\lambda$ (cf. Lemma A1) and from Properties 1 (P1) the unique negative solution $x_0^{(q_\lambda^E)} \in]-\beta_\lambda, q_\lambda^E - \beta_\lambda[$ of $\xi_\lambda^{(q_\lambda^E)}(x) = q_\lambda^E e^x - \beta_\lambda = x$ (cf. (44)). Due to the continuity and boundedness of the map $\lambda \mapsto x_0^{(q_\lambda^E)}$ (for $\lambda \in [0, 1]$) one gets that $\lim_{\lambda \nearrow 1} x_0^{(q_\lambda^E)}$ exists and is the smallest nonpositive solution of $\beta_{\mathcal{A}} e^x - \beta_{\mathcal{A}} = x$. From this, the part (b) as well as the non-contiguity in part (c) follow immediately. The other part of (c) is a direct consequence of

Proposition 1 (a),(b) and Proposition 2 (c). □

Proof of Formula (59) . One can proceed similarly to the proof of formula (51) above. Recall $H_\lambda(P_{\mathcal{A},1}||P_{\mathcal{H},1}) = \exp\{\phi_\lambda(X_0)\} > 1$ for $X_0 \in \mathbb{N}$ (cf. (28), Lemma A1 and $f_{\mathcal{A}}(X_0) \neq f_{\mathcal{H}}(X_0)$ for all $X_0 \in \mathbb{N}$). For $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}, \lambda) \in \mathcal{P}_{SP,2} \times (\mathbb{R} \setminus [0, 1])$ one gets $\phi_\lambda(0) = 0, \phi_\lambda(1) > 0$, and we define for $x \geq 0$

$$\underline{\phi}_\lambda(x) := \begin{cases} \phi_\lambda(1), & \text{if } x = 1, \\ 0, & \text{if } x \neq 1. \end{cases}$$

By means of the choice $\underline{\epsilon} := \phi_\lambda(1) \cdot (e^{\phi_\lambda(1)} - 1) > 0$, we obtain for all $n \in \mathbb{N} \setminus \{1\}$

$$\begin{aligned} \sum_{x_{n-1}=0}^\infty \frac{[\phi_\lambda(x_{n-2})]^{x_{n-1}}}{x_{n-1}!} \cdot \exp\{\phi_\lambda(x_{n-1})\} &\geq \sum_{x_{n-1}=0}^\infty \frac{[\phi_\lambda(x_{n-2})]^{x_{n-1}}}{x_{n-1}!} \cdot \exp\{\underline{\phi}_\lambda(x_{n-1})\} \\ &= \exp\{\phi_\lambda(x_{n-2})\} + \underline{\epsilon} = \exp\{\phi_\lambda(x_{n-2})\} \cdot [1 + \underline{\epsilon} \cdot \exp\{-\phi_\lambda(x_{n-2})\}]. \end{aligned}$$

Incorporating

$$\inf_{x \in \mathbb{N}_0} \left\{ \exp\{\phi_\lambda(x)\} \cdot [1 + \underline{\epsilon} \cdot \exp\{-\phi_\lambda(x)\}] \right\} =: \underline{\delta} > 1,$$

one can show analogously to (A4) that

$$H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) \geq \dots \geq \underline{\delta}^{\lfloor n/2 \rfloor} \xrightarrow{n \rightarrow \infty} \infty. \quad \square$$

Proof of the Formulas (61), (63) and (64). In the following, we slightly adapt the above-mentioned proof of formula (59). Let us define

$$\underline{\phi}_\lambda(x) := \begin{cases} \phi_\lambda(0), & \text{if } x = 0, \\ 0, & \text{if } x > 0. \end{cases}$$

In all respective subcases one clearly has $\underline{\phi}_\lambda(0) = \phi_\lambda(0) > 0$. With $\underline{\epsilon} := e^{\phi_\lambda(0)} - 1 > 0$ we obtain for all $n \in \mathbb{N} \setminus \{1\}$

$$\begin{aligned} \sum_{x_{n-1}=0}^\infty \frac{[\phi_\lambda(x_{n-2})]^{x_{n-1}}}{x_{n-1}!} \cdot \exp\{\phi_\lambda(x_{n-1})\} &\geq \sum_{x_{n-1}=0}^\infty \frac{[\phi_\lambda(x_{n-2})]^{x_{n-1}}}{x_{n-1}!} \cdot \exp\{\underline{\phi}_\lambda(x_{n-1})\} \\ &= \exp\{\phi_\lambda(x_{n-2})\} + \underline{\epsilon} = \exp\{\phi_\lambda(x_{n-2})\} \cdot [1 + \underline{\epsilon} \cdot \exp\{-\phi_\lambda(x_{n-2})\}]. \end{aligned}$$

By employing

$$\inf_{x \in \mathbb{N}_0} \left\{ \exp\{\phi_\lambda(x)\} \cdot [1 + \underline{\epsilon} \cdot \exp\{-\phi_\lambda(x)\}] \right\} =: \underline{\delta} > 1, \tag{A12}$$

one can show analogously to (A4) that

$$H_\lambda(P_{\mathcal{A},n}||P_{\mathcal{H},n}) \geq \dots \geq \underline{\delta}^{\lfloor n/2 \rfloor} \xrightarrow{n \rightarrow \infty} \infty.$$

Notice that this method does not work for the parameter cases $\mathcal{P}_{SP,4a} \cup \mathcal{P}_{SP,4b}$, since there the infimum in (A12) is equal to one. □

Proof of Proposition 13. In the setup $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P}_{SP,4a} \times (\mathbb{R} \setminus [0, 1])$ we require $\beta_\bullet := \beta_A = \beta_H < 1$. As in the proof of Proposition 9, we stick to the tangent line $\phi_{\lambda,y}^{\text{tan}}(\cdot)$ at $y \geq 0$ (cf. (52)) as a linear lower bound for $\phi_\lambda(\cdot)$, i.e., we use the function

$$\phi_{\lambda,y}^{\text{tan}}(x) := (p_y - \alpha_\lambda) + (q_y - \beta_\bullet) \cdot x := (p_{\lambda,y}^{\text{tan}} - \alpha_\lambda) + (q_{\lambda,y}^{\text{tan}} - \beta_\lambda) \cdot x := (\phi_\lambda(y) - y \cdot \phi'_\lambda(y)) + \phi'_\lambda(y) \cdot x. \tag{A13}$$

As already mentioned in Section 3.21, on $\mathcal{P}_{SP,4a}$ the function $\phi_\lambda(\cdot)$ is strictly decreasing and converges to 0. Thus, for all $y \geq 0$ the slope $\phi'_\lambda(y)$ of the tangent line at y is negative, which implies that $q_y < \beta_\lambda = \beta_\bullet$. For $\lambda \in \mathbb{R} \setminus [0, 1]$ there clearly may hold $q_y < 0$ for some $y \in \mathbb{R}$. However, there exists a sufficiently large $y_1 > 0$ such that $q_y > 0$ for all $y > y_1$, since $\lim_{y \rightarrow \infty} \phi'_\lambda(y) = 0$ and hence $q_y \nearrow \beta_\bullet > 0$ for $y \rightarrow \infty$. Thus, let us suppose that $y > y_1$. Then, the sequence $(a_n^{(q_y)})_{n \in \mathbb{N}}$ is strictly negative, strictly decreasing and converges to $x_0^{(q_y)} \in]-\beta_\bullet, q_y - \beta_\bullet[$ (cf. Properties 1 (P1)). If there is some $y \geq y_1$ such that $h(y) > 0$ with

$$h(y) := \lim_{n \rightarrow \infty} \frac{1}{n} \log(\tilde{B}_{\lambda, X_0, n}^{(p_y, q_y)}) = p_y \cdot e^{x_0^{(q_y)}} - \alpha_\lambda,$$

then one can conclude that $\lim_{n \rightarrow \infty} H_\lambda(P_{A,n} || P_{H,n}) = \infty$. Let us at first consider the case $\alpha_\lambda \geq 0$. By employing $p_y \searrow \alpha_\lambda$ for $y \rightarrow \infty$, one gets $p_y > 0$ for all $y \geq 0$. Analogously to the proof of Proposition 9, we now look for a lower bound $\underline{x}_0^{(q_y)}$ of the fixed point $x_0^{(q_y)}$. Notice that $x_0^{(q_y)} > -\beta_\bullet$ implies

$$\underline{Q}_\lambda^{(q_y)}(x) := \frac{e^{-\beta_\bullet}}{2} \cdot q_y \cdot x^2 + q_y \cdot x + q_y - \beta_\bullet \leq q_y \cdot e^x - \beta_\bullet = \tilde{\zeta}_\lambda^{(q_y)}(x), \tag{A14}$$

since $\underline{Q}_\lambda^{(q_y)}(0) = \tilde{\zeta}_\lambda^{(q_y)}(0) < 0$, $\underline{Q}_\lambda^{(q_y)'}(0) = \tilde{\zeta}_\lambda^{(q_y)'}(0) > 0$ and $0 < \underline{Q}_\lambda^{(q_y)''}(x) < \tilde{\zeta}_\lambda^{(q_y)''}(x)$ for $x \in]-\beta_\bullet, 0]$. Thus, the negative solution $\underline{x}_0^{(q_y)}$ of the equation $\underline{Q}_\lambda^{(q_y)}(x) = x$ (which definitely exists) implies that there holds $\underline{x}_0^{(q_y)} \leq x_0^{(q_y)}$. We easily obtain

$$\begin{aligned} \underline{x}_0^{(q_y)} &= \frac{e^{\beta_\bullet}}{q_y} \left[(1 - q_y) - \sqrt{(1 - q_y)^2 - 2e^{-\beta_\bullet} q_y (q_y - \beta_\bullet)} \right] \\ &= \frac{e^{\beta_\bullet}}{\phi'_\lambda(y) + \beta_\bullet} \left[(1 - \phi'_\lambda(y) - \beta_\bullet) - \sqrt{(1 - \phi'_\lambda(y) - \beta_\bullet)^2 - 2 \cdot e^{-\beta_\bullet} q_y \cdot \phi'_\lambda(y)} \right] < 0. \end{aligned} \tag{A15}$$

Since

$$h(y) = p_y \cdot e^{x_0^{(q_y)}} - \alpha_\lambda \geq p_y \cdot e^{\underline{x}_0^{(q_y)}} - \alpha_\lambda =: \underline{h}(y), \tag{A16}$$

it is sufficient to show $\underline{h}(y) > 0$ for some $y > y_1$. We recall from Properties 3 (P15), (P17) and (P19) that

$$\begin{aligned} \phi_\lambda(y) &= (\alpha_A + \beta_\bullet \cdot y)^\lambda (\alpha_H + \beta_\bullet \cdot y)^{1-\lambda} - \lambda (\alpha_A + \beta_\bullet \cdot y) - (1 - \lambda) (\alpha_H + \beta_\bullet \cdot y) > 0, \\ \phi'_\lambda(y) &= \lambda \cdot \beta_\bullet \cdot \left(\frac{\alpha_A + \beta_\bullet \cdot y}{\alpha_H + \beta_\bullet \cdot y} \right)^{\lambda-1} + (1 - \lambda) \cdot \beta_\bullet \cdot \left(\frac{\alpha_A + \beta_\bullet \cdot y}{\alpha_H + \beta_\bullet \cdot y} \right)^\lambda - \beta_\bullet < 0 \quad \text{and} \\ \phi''_\lambda(y) &= - \left(\frac{\alpha_A + \beta_\bullet \cdot y}{\alpha_H + \beta_\bullet \cdot y} \right)^\lambda \cdot \frac{\lambda(1 - \lambda) \cdot \beta_\bullet^2 \cdot (\alpha_A - \alpha_H)^2}{(\alpha_A + \beta_\bullet \cdot y)^2 (\alpha_H + \beta_\bullet \cdot y)} > 0, \end{aligned} \tag{A17}$$

which immediately implies $\lim_{y \rightarrow \infty} \phi_\lambda(y) = \lim_{y \rightarrow \infty} \phi'_\lambda(y) = \lim_{y \rightarrow \infty} \phi''_\lambda(y) = 0$, and by means of l'Hospital's rule

$$\begin{aligned} \lim_{y \rightarrow \infty} y \cdot \phi_\lambda(y) &= \lim_{y \rightarrow \infty} -y^2 \cdot \phi'_\lambda(y) = \lim_{y \rightarrow \infty} \frac{y^3}{2} \cdot \phi''_\lambda(y) \\ &= -\frac{1}{2} \lim_{y \rightarrow \infty} \left(\frac{\alpha_A + \beta_\bullet \cdot y}{\alpha_H + \beta_\bullet \cdot y} \right)^\lambda \cdot \frac{\lambda(1 - \lambda) \cdot \beta_\bullet^2 \cdot (\alpha_A - \alpha_H)^2}{(\alpha_A/y + \beta_\bullet)^2 (\alpha_H/y + \beta_\bullet)} = -\frac{1}{2} \lambda(1 - \lambda) \cdot \frac{(\alpha_A - \alpha_H)^2}{\beta_\bullet}. \end{aligned} \tag{A18}$$

The Formulas (A13), (A15), (A17) imply the limits $\lim_{y \rightarrow \infty} p_y = \alpha_\lambda$, $\lim_{y \rightarrow \infty} q_y = \beta_\bullet$ and $\lim_{y \rightarrow \infty} x_0^{(qy)} = 0$ iff $\beta_\bullet \leq 1$. The latter is due to the fact that for $\beta_\bullet > 1$ one gets with (A15) $\lim_{y \rightarrow \infty} x_0^{(qy)} = \frac{e^{\beta_\bullet}}{\beta_\bullet} [(1 - \beta_\bullet) - \sqrt{(1 - \beta_\bullet)^2}] = \frac{e^{\beta_\bullet}}{\beta_\bullet} [2 - 2\beta_\bullet] \neq 0$. In the following, let us assume $\beta_\bullet < 1$ (the reason why we exclude the case $\beta_\bullet = 1$ is explained below). One gets $\lim_{y \rightarrow \infty} h(y) \geq \lim_{y \rightarrow \infty} \underline{h}(y) = 0$. Since we have to prove that $\underline{h}(y) > 0$ for some $y > y_1$, it is sufficient to show that $\lim_{y \rightarrow \infty} y \cdot \underline{h}(y) > 0$. To verify the latter, we first derive with l'Hospital's rule and with (A17), (A18)

$$\begin{aligned} \lim_{y \rightarrow \infty} y \cdot \left(1 - e^{x_0^{(qy)}}\right) &= \lim_{y \rightarrow \infty} y^2 \cdot e^{x_0^{(qy)}} \cdot \left(\frac{\partial}{\partial y} x_0^{(qy)}\right) \\ &= \lim_{y \rightarrow \infty} \left\{ y^2 \cdot \frac{-e^{\beta_\bullet} \cdot \phi_\lambda''(y)}{(\phi_\lambda'(y) + \beta_\bullet)^2} \cdot \left[(1 - q_y) - \sqrt{(1 - q_y)^2 - 2e^{-\beta_\bullet} q_y (q_y - \beta_\bullet)} \right] \right. \\ &\quad \left. + \frac{e^{\beta_\bullet}}{q_y} \cdot \left[-y^2 \cdot \phi_\lambda''(y) - \frac{-2y^2 \phi_\lambda''(y)(1 - q_y) - 2y^2 \phi_\lambda''(y)e^{-\beta_\bullet} q_y - 2y^2 \phi_\lambda''(y)e^{-\beta_\bullet} \phi_\lambda'(y)}{2 \cdot \sqrt{(1 - q_y)^2 - 2e^{-\beta_\bullet} q_y (q_y - \beta_\bullet)}} \right] \right\} \\ &= 0. \end{aligned} \tag{A19}$$

Notice that without further examination this limit would not necessarily hold for $\beta_\bullet = 1$, since then the denominator in (A19) converges to zero. With (A13), (A16), (A18) and (A19) we finally obtain

$$\begin{aligned} \lim_{y \rightarrow \infty} y \cdot \underline{h}(y) &= \lim_{y \rightarrow \infty} \left\{ \left(y \cdot \phi_\lambda(y) - y^2 \cdot \phi_\lambda'(y) \right) \cdot e^{x_0^{(qy)}} - y \cdot \left(1 - e^{x_0^{(qy)}}\right) \alpha_\lambda \right\} \\ &= -\lambda(1 - \lambda) \frac{(\alpha_A - \alpha_H)^2}{\beta_\bullet} > 0. \end{aligned} \tag{A20}$$

Let us now consider the case $\alpha_\lambda < 0$. The proof works out almost completely analogous to the case $\alpha_\lambda \geq 0$. We indicate the main differences. Since $p_y \searrow \alpha_\lambda < 0$ and $q_y \nearrow \beta_\bullet \in]0, 1[$ for $y \rightarrow \infty$, there is a sufficiently large $y_2 > y_1$, such that $p_y < 0$ and $q_y > 0$. Thus,

$$\overline{Q}_\lambda^{(qy)}(x) := \frac{q_y}{2} \cdot x^2 + q_y \cdot x + q_y - \beta_\bullet \geq \xi_\lambda^{(qy)}(x) = q_y e^x - \beta_\bullet \quad \text{for } x \in]-\infty, 0].$$

The corresponding (existing) smaller solution of $\overline{Q}_\lambda^{(qy)}(x) = x$ is

$$\overline{x}_0^{(qy)} = \frac{1}{q_y} \left[(1 - q_y) - \sqrt{(1 - q_y)^2 - 2q_y(q_y - \beta_\bullet)} \right],$$

having the same form as the solution (A15) with $e^{-\beta_\bullet}$ substituted by 1. Notice that there clearly holds $x_0^{(qy)} < \overline{x}_0^{(qy)} < 0$. However, since $p_y < 0$, we now get $h(y) = p_y \cdot e^{x_0^{(qy)}} - \alpha_\lambda \geq p_y \cdot e^{\overline{x}_0^{(qy)}} - \alpha_\lambda =: \underline{h}(y)$, as in (A16). Since all calculations (A17) to (A20) remain valid (with $e^{-\beta_\bullet}$ substituted by 1), this proof is finished. \square

Appendix A.2. Proofs and Auxiliary Lemmas for Section 5

We start with two lemmas which will be useful for the proof of Theorem 3. They deal with the sequence $(a_n^{(q\lambda)})_{n \in \mathbb{N}}$ from (36).

Lemma A2. For arbitrarily fixed parameter constellation $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P} \times]0, 1[$, suppose that $q_\lambda > 0$ and $\lim_{\lambda \nearrow 1} q_\lambda = \beta_A$ holds. Then one gets the limit

$$\forall n \in \mathbb{N} : \lim_{\lambda \nearrow 1} a_n^{(q\lambda)} = 0. \tag{A21}$$

Proof. This can be easily seen by induction: for $n = 1$ there clearly holds

$$\lim_{\lambda \nearrow 1} a_1^{(q_\lambda)} = \lim_{\lambda \nearrow 1} (q_\lambda - \beta_\lambda) = \beta_{\mathcal{A}} - \beta_{\mathcal{A}} = 0.$$

Assume now that $\lim_{\lambda \nearrow 1} a_k^{(q_\lambda)} = 0$ holds for all $k \in \mathbb{N}, k \leq n - 1$, then

$$\lim_{\lambda \nearrow 1} a_n^{(q_\lambda)} = \lim_{\lambda \nearrow 1} (q_\lambda \cdot e^{a_{n-1}^{(q_\lambda)}} - \beta_\lambda) = \beta_{\mathcal{A}} \cdot 1 - \beta_{\mathcal{A}} = 0. \quad \square$$

Lemma A3. In addition to the assumptions of Lemma A2, suppose that $\lambda \mapsto q_\lambda$ is continuously differentiable on $]0, 1[$ and that the limit $l := \lim_{\lambda \nearrow 1} \frac{\partial q_\lambda}{\partial \lambda}$ is finite. Then, for all $n \in \mathbb{N}$ one obtains

$$\lim_{\lambda \nearrow 1} \frac{\partial a_n^{(q_\lambda)}}{\partial \lambda} = u_n := \begin{cases} \frac{l + \beta_{\mathcal{H}} - \beta_{\mathcal{A}}}{1 - \beta_{\mathcal{A}}} \cdot (1 - (\beta_{\mathcal{A}})^n), & \text{if } \beta_{\mathcal{A}} \neq 1, \\ n \cdot (l + \beta_{\mathcal{H}} - 1), & \text{if } \beta_{\mathcal{A}} = 1, \end{cases} \quad (\text{A22})$$

which is the unique solution of the linear recursion equation

$$u_n = l + \beta_{\mathcal{H}} - \beta_{\mathcal{A}} + \beta_{\mathcal{A}} \cdot u_{n-1}, \quad u_0 = 0. \quad (\text{A23})$$

Furthermore, for all $n \in \mathbb{N}$ there holds

$$\sum_{k=1}^n \lim_{\lambda \nearrow 1} \frac{\partial a_k^{(q_\lambda)}}{\partial \lambda} = \sum_{k=1}^n u_k = \begin{cases} \frac{l + \beta_{\mathcal{H}} - \beta_{\mathcal{A}}}{1 - \beta_{\mathcal{A}}} \cdot \left[n - \frac{\beta_{\mathcal{A}}}{1 - \beta_{\mathcal{A}}} (1 - (\beta_{\mathcal{A}})^n) \right], & \text{if } \beta_{\mathcal{A}} \neq 1, \\ \frac{n \cdot (n+1)}{2} \cdot (l + \beta_{\mathcal{H}} - 1), & \text{if } \beta_{\mathcal{A}} = 1. \end{cases}$$

Proof. Clearly, u_n defined by (A22) is the unique solution of (A23). We prove by induction that $\lim_{\lambda \nearrow 1} \frac{\partial a_n^{(q_\lambda)}}{\partial \lambda} = u_n$ holds. For $n = 1$ one gets

$$\lim_{\lambda \nearrow 1} \frac{\partial a_1^{(q_\lambda)}}{\partial \lambda} = \lim_{\lambda \nearrow 1} \frac{\partial (q_\lambda - \beta_\lambda)}{\partial \lambda} = l - (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) = u_1.$$

Suppose now that (A22) holds for all $k \in \mathbb{N}, k \leq n - 1$. Then, by incorporating (A21) we obtain

$$\begin{aligned} \lim_{\lambda \nearrow 1} \frac{\partial a_n^{(q_\lambda)}}{\partial \lambda} &= \lim_{\lambda \nearrow 1} \frac{\partial}{\partial \lambda} \left(q_\lambda \cdot e^{a_{n-1}^{(q_\lambda)}} - \beta_\lambda \right) = \lim_{\lambda \nearrow 1} e^{a_{n-1}^{(q_\lambda)}} \cdot \left(\frac{\partial q_\lambda}{\partial \lambda} + q_\lambda \frac{\partial a_{n-1}^{(q_\lambda)}}{\partial \lambda} \right) - (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) \\ &= l - (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) + \beta_{\mathcal{A}} \cdot u_{n-1} = u_n. \end{aligned}$$

The remaining assertions follow immediately. \square

We are now ready to give the

Proof of Theorem 3. (a) Recall that for the setup $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in (\mathcal{P}_{\text{NI}} \cup \mathcal{P}_{\text{SP},1})$ we chose the intercept as $p_\lambda := p_\lambda^E := \alpha_{\mathcal{A}}^\lambda \alpha_{\mathcal{H}}^{1-\lambda}$ and the slope as $q_\lambda := q_\lambda^E := \beta_{\mathcal{A}}^\lambda \beta_{\mathcal{H}}^{1-\lambda}$, which in (39) lead to the exact

value $V_{\lambda, X_0, n}$ of the Hellinger integral. Because of $\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda = 0$ as well as $\lim_{\lambda \nearrow 1} q_\lambda = \beta_{\mathcal{A}}$, we obtain by using (38) and Lemma A2 for all $X_0 \in \mathbb{N}$ and for all $n \in \mathbb{N}$

$$\lim_{\lambda \nearrow 1} V_{\lambda, X_0, n} := \lim_{\lambda \nearrow 1} \exp \left\{ a_n^{(q_\lambda)} \cdot X_0 + \sum_{k=1}^n b_k^{(p_\lambda, q_\lambda)} \right\} = \lim_{\lambda \nearrow 1} \exp \left\{ a_n^{(q_\lambda)} \cdot X_0 + \frac{\alpha_{\mathcal{A}}}{\beta_{\mathcal{A}}} \sum_{k=1}^n a_k^{(q_\lambda)} \right\} = 1,$$

which leads by (68) to

$$\begin{aligned} I(P_{\mathcal{A}, n} || P_{\mathcal{H}, n}) &= \lim_{\lambda \nearrow 1} \frac{1 - H_\lambda(P_{\mathcal{A}, n} || P_{\mathcal{H}, n})}{\lambda \cdot (1 - \lambda)} = \lim_{\lambda \nearrow 1} \frac{1 - V_{\lambda, X_0, n}}{\lambda \cdot (1 - \lambda)} \\ &= \lim_{\lambda \nearrow 1} \frac{-V_{\lambda, X_0, n}}{1 - 2\lambda} \cdot \frac{\partial}{\partial \lambda} \left[a_n^{(q_\lambda)} \cdot X_0 + \frac{p_\lambda}{q_\lambda} \sum_{k=1}^n a_k^{(q_\lambda)} \right] \\ &= \lim_{\lambda \nearrow 1} \left[\frac{\partial a_n^{(q_\lambda)}}{\partial \lambda} \cdot X_0 + \left(\frac{\partial p_\lambda}{\partial \lambda} \frac{1}{q_\lambda} \right) \cdot \sum_{k=1}^n a_k^{(q_\lambda)} + \frac{p_\lambda}{q_\lambda} \cdot \sum_{k=1}^n \frac{\partial a_k^{(q_\lambda)}}{\partial \lambda} \right]. \end{aligned} \tag{A24}$$

For further analysis, we use the obvious derivatives

$$\frac{\partial p_\lambda}{\partial \lambda} = p_\lambda \log \left(\frac{\alpha_{\mathcal{A}}}{\alpha_{\mathcal{H}}} \right), \quad \frac{\partial p_\lambda}{\partial \lambda} \frac{1}{q_\lambda} = \frac{p_\lambda}{q_\lambda} \log \left(\frac{\alpha_{\mathcal{A}} \beta_{\mathcal{H}}}{\alpha_{\mathcal{H}} \beta_{\mathcal{A}}} \right), \quad \frac{\partial q_\lambda}{\partial \lambda} = q_\lambda \log \left(\frac{\beta_{\mathcal{A}}}{\beta_{\mathcal{H}}} \right), \tag{A25}$$

where the subcase $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{\text{NI}}$ (with $p_\lambda \equiv 0$) is consistently covered. From (A25) and Lemma A3 we deduce

$$\lim_{\lambda \nearrow 1} \frac{\partial a_n^{(q_\lambda)}}{\partial \lambda} \cdot X_0 = \begin{cases} \left(\beta_{\mathcal{A}} \log \left(\frac{\beta_{\mathcal{A}}}{\beta_{\mathcal{H}}} \right) - (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) \right) \cdot \frac{1 - (\beta_{\mathcal{A}})^n}{1 - \beta_{\mathcal{A}}} \cdot X_0, & \text{if } \beta_{\mathcal{A}} \neq 1, \\ n \cdot \left(\beta_{\mathcal{A}} \log \left(\frac{\beta_{\mathcal{A}}}{\beta_{\mathcal{H}}} \right) - (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) \right) \cdot X_0, & \text{if } \beta_{\mathcal{A}} = 1, \end{cases}$$

and by means of (A21)

$$\forall n \in \mathbb{N} : \lim_{\lambda \nearrow 1} \left[\left(\frac{\partial p_\lambda}{\partial \lambda} \frac{1}{q_\lambda} \right) \cdot \sum_{k=1}^n a_k^{(q_\lambda)} \right] = 0.$$

For the last expression in (A24) we again apply Lemma A3 to end up with

$$\lim_{\lambda \nearrow 1} \frac{p_\lambda}{q_\lambda} \cdot \sum_{k=1}^n \frac{\partial}{\partial \lambda} a_k^{(q_\lambda)} = \begin{cases} \frac{\alpha_{\mathcal{A}} \cdot \left[\beta_{\mathcal{A}} \log \left(\frac{\beta_{\mathcal{A}}}{\beta_{\mathcal{H}}} \right) - (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) \right]}{\beta_{\mathcal{A}} (1 - \beta_{\mathcal{A}})} \cdot \left[n - \frac{\beta_{\mathcal{A}}}{1 - \beta_{\mathcal{A}}} (1 - (\beta_{\mathcal{A}})^n) \right], & \text{if } \beta_{\mathcal{A}} \neq 1, \\ n \cdot (n + 1) \frac{\alpha_{\mathcal{A}}}{2\beta_{\mathcal{A}}} \cdot \left[\beta_{\mathcal{A}} \log \left(\frac{\beta_{\mathcal{A}}}{\beta_{\mathcal{H}}} \right) - (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) \right], & \text{if } \beta_{\mathcal{A}} = 1, \end{cases} \tag{A26}$$

which finishes the proof of part (a). To show part (b), for the corresponding setup $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP}, 1}$ let us first choose – according to (45) in Section 3.4 – the intercept as $p_\lambda := p_\lambda^L := \alpha_{\mathcal{A}}^\lambda \alpha_{\mathcal{H}}^{1-\lambda}$ and the slope as $q_\lambda := q_\lambda^L := \beta_{\mathcal{A}}^\lambda \beta_{\mathcal{H}}^{1-\lambda}$, which in part (b) of Proposition 6 lead to the lower bounds $B_{\lambda, X_0, n}^L$ of the Hellinger integral. This is formally the same choice as in part (a) satisfying $\lim_{\lambda \nearrow 1} p_\lambda = \alpha_{\mathcal{A}}$, $\lim_{\lambda \nearrow 1} q_\lambda = \beta_{\mathcal{A}}$ but in contrast to (a) we now have $\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda \neq 0$ but nevertheless

$$\lim_{\lambda \nearrow 1} \frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda = 0.$$

From this, (38), part (b) of Proposition 6 and Lemma A2 we obtain

$$\lim_{\lambda \nearrow 1} B_{\lambda, X_0, n}^L = \lim_{\lambda \nearrow 1} \exp \left\{ a_n^{(q_\lambda)} \cdot X_0 + \frac{p_\lambda}{q_\lambda} \sum_{k=1}^n a_k^{(q_\lambda)} + n \cdot \left(\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda \right) \right\} = 1 \tag{A27}$$

and hence

$$\begin{aligned}
 I(P_{\mathcal{A},n}||P_{\mathcal{H},n}) &\leq \lim_{\lambda \nearrow 1} \frac{1 - B_{\lambda, X_0, n}^L}{\lambda \cdot (1 - \lambda)} = \lim_{\lambda \nearrow 1} \frac{-B_{\lambda, X_0, n}^L}{1 - 2\lambda} \cdot \frac{\partial}{\partial \lambda} \left[a_n^{(q_\lambda)} X_0 + \frac{p_\lambda}{q_\lambda} \sum_{k=1}^n a_k^{(q_\lambda)} + n \left(\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda \right) \right] \\
 &= \lim_{\lambda \nearrow 1} \left[\frac{\partial a_n^{(q_\lambda)}}{\partial \lambda} X_0 + \left(\frac{\partial}{\partial \lambda} \frac{p_\lambda}{q_\lambda} \right) \sum_{k=1}^n a_k^{(q_\lambda)} + \frac{p_\lambda}{q_\lambda} \sum_{k=1}^n \frac{\partial a_k^{(q_\lambda)}}{\partial \lambda} + n \frac{\partial}{\partial \lambda} \left(\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda \right) \right]. \tag{A28}
 \end{aligned}$$

In the current setup, the first three expressions in (A28) can be evaluated in exactly the same way as in (A25) to (A26), and for the last expression one has the limit

$$\begin{aligned}
 \frac{\partial}{\partial \lambda} \left(\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda \right) &= \frac{p_\lambda}{q_\lambda} \log \left(\frac{\alpha_{\mathcal{A}} \beta_{\mathcal{H}}}{\alpha_{\mathcal{H}} \beta_{\mathcal{A}}} \right) \cdot \beta_\lambda + \frac{p_\lambda}{q_\lambda} \cdot (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) - (\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}}) \\
 &\xrightarrow{\lambda \nearrow 1} \alpha_{\mathcal{A}} \left[\log \left(\frac{\alpha_{\mathcal{A}} \beta_{\mathcal{H}}}{\alpha_{\mathcal{H}} \beta_{\mathcal{A}}} \right) - \frac{\beta_{\mathcal{H}}}{\beta_{\mathcal{A}}} \right] + \alpha_{\mathcal{H}},
 \end{aligned}$$

which finishes the proof of part (b). □

Proof of Theorem 4. Let us fix $(\beta_{\mathcal{A}}, \beta_{\mathcal{H}}, \alpha_{\mathcal{A}}, \alpha_{\mathcal{H}}) \in \mathcal{P}_{\text{SP}} \setminus \mathcal{P}_{\text{SP},1}$, $X_0 \in \mathbb{N}$, $n \in \mathbb{N}$ and $y \in [0, \infty[$. The lower bound $E_{y, X_0, n}^{L, \text{tan}}$ of the Kullback-Leibler information divergence (relative entropy) is derived by using $\phi_\lambda^U \equiv \phi_{\lambda, y}^{\text{tan}}$ (cf. (52)), which corresponds to the tangent line of ϕ_λ at y , as a linear upper bound for ϕ_λ ($\lambda \in]0, 1[$). More precisely, one gets $\phi_\lambda^U(x) := (p_\lambda^U - \alpha_\lambda) + (q_\lambda^U - \beta_\lambda) x$ ($x \in [0, \infty[$) with $p_\lambda := p_\lambda(y) := \phi_\lambda(y) - y \phi'_\lambda(y) + \alpha_\lambda$ and $q_\lambda := q_\lambda(y) := \phi'_\lambda(y) + \beta_\lambda$, implying $q_\lambda > 0$ because of Properties 3 (P17). Analogously to (A27) and (A28), we obtain from (38) and (40) the convergence $\lim_{\lambda \nearrow 1} B_{\lambda, X_0, n}^U = 1$ and thus

$$I(P_{\mathcal{A},n}||P_{\mathcal{H},n}) \geq \lim_{\lambda \nearrow 1} \left[\frac{\partial a_n^{(q_\lambda)}}{\partial \lambda} X_0 + \left(\frac{\partial}{\partial \lambda} \frac{p_\lambda}{q_\lambda} \right) \sum_{k=1}^n a_k^{(q_\lambda)} + \frac{p_\lambda}{q_\lambda} \sum_{k=1}^n \frac{\partial a_k^{(q_\lambda)}}{\partial \lambda} + n \frac{\partial}{\partial \lambda} \left(\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda \right) \right]. \tag{A29}$$

As before, we compute the involved derivatives. From (30) to (32) as well as (P17) we get

$$\begin{aligned}
 \frac{\partial p_\lambda}{\partial \lambda} &= \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^\lambda f_{\mathcal{H}}(y) \log \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right) - \beta_{\mathcal{A}y} \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^{\lambda-1} - \lambda \beta_{\mathcal{A}y} \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^{\lambda-1} \log \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right) \\
 &\quad + \beta_{\mathcal{H}y} \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^\lambda - (1 - \lambda) \beta_{\mathcal{H}y} \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^\lambda \log \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right) \\
 &\xrightarrow{\lambda \nearrow 1} \alpha_{\mathcal{A}} \log \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right) + \frac{y \cdot (\alpha_{\mathcal{A}} \beta_{\mathcal{H}} - \alpha_{\mathcal{H}} \beta_{\mathcal{A}})}{f_{\mathcal{H}}(y)}, \tag{A30}
 \end{aligned}$$

and

$$\begin{aligned}
 \frac{\partial q_\lambda}{\partial \lambda} &= \beta_{\mathcal{A}} \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^{\lambda-1} + \lambda \beta_{\mathcal{A}} \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^{\lambda-1} \log \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right) - \beta_{\mathcal{H}} \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^\lambda \\
 &\quad + (1 - \lambda) \beta_{\mathcal{H}} \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right)^\lambda \log \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right) \\
 &\xrightarrow{\lambda \nearrow 1} \beta_{\mathcal{A}} \left(1 + \log \left(\frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right) \right) - \beta_{\mathcal{H}} \frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} =: l. \tag{A31}
 \end{aligned}$$

Combining these two limits we get

$$\begin{aligned}
 \frac{\partial}{\partial \lambda} \left(\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda \right) &= \frac{q_\lambda \left(\frac{\partial p_\lambda}{\partial \lambda} \right) - p_\lambda \left(\frac{\partial q_\lambda}{\partial \lambda} \right)}{(q_\lambda)^2} \cdot \beta_\lambda + \frac{p_\lambda}{q_\lambda} \cdot (\beta_{\mathcal{A}} - \beta_{\mathcal{H}}) - (\alpha_{\mathcal{A}} - \alpha_{\mathcal{H}}) \\
 &\xrightarrow{\lambda \nearrow 1} \left[\frac{y \cdot (\alpha_{\mathcal{A}} \beta_{\mathcal{H}} - \alpha_{\mathcal{H}} \beta_{\mathcal{A}})}{f_{\mathcal{H}}(y)} - \alpha_{\mathcal{A}} \left(1 - \frac{\beta_{\mathcal{H}} f_{\mathcal{A}}(y)}{\beta_{\mathcal{A}} f_{\mathcal{H}}(y)} \right) \right] + \alpha_{\mathcal{H}} - \frac{\alpha_{\mathcal{A}} \beta_{\mathcal{H}}}{\beta_{\mathcal{A}}} \\
 &= \left(\alpha_{\mathcal{H}} - \alpha_{\mathcal{A}} \frac{\beta_{\mathcal{H}}}{\beta_{\mathcal{A}}} \right) \left(1 - \frac{f_{\mathcal{A}}(y)}{f_{\mathcal{H}}(y)} \right). \tag{A32}
 \end{aligned}$$

The above calculation also implies that $\lim_{\lambda \nearrow 1} \left(\frac{\partial p_\lambda}{\partial \lambda} \frac{p_\lambda}{q_\lambda} \right)$ is finite and thus $\lim_{\lambda \nearrow 1} \left(\frac{\partial p_\lambda}{\partial \lambda} \frac{p_\lambda}{q_\lambda} \right) \sum_{k=1}^n a_k^{(q_\lambda)} = 0$ by means of Lemma A2. The proof of $I(P_{A,n} || P_{H,n}) \geq E_{y, X_{0,n}}^{L, \tan}$ is finished by using Lemma A3 with l defined in (A31) and by plugging the limits (A30) to (A32) in (A29).

To derive the lower bound $E_{k, X_{0,n}}^{L, \sec}$ (cf. (73)) for fixed $k \in \mathbb{N}_0$, we use as a linear upper bound ϕ_λ^U for $\phi_\lambda(\cdot)$ ($\lambda \in]0, 1[$) the secant line $\phi_{\lambda,k}^{\sec}$ (cf. (53)) of ϕ_λ across its arguments k and $k + 1$, corresponding to the choices $p_\lambda := p_{\lambda,k}^{\sec} = (k + 1) \cdot \phi_\lambda(k) - k \cdot \phi_\lambda(k + 1) + \alpha_\lambda$ and $q_\lambda := q_{\lambda,k}^{\sec} := \phi_\lambda(k + 1) - \phi_\lambda(k) + \beta_\lambda$, implying $q_\lambda > 0$ because of Properties 3 (P18). As a side remark, notice that this $\phi_\lambda^U(x)$ may become positive for some $x \in [0, \infty[$ (which is not always consistent with Goal (G1) for fixed λ , but leads to a tractable limit bound as λ tends to 1). Analogously to (A27) and (A28) we get again $\lim_{\lambda \nearrow 1} B_{\lambda, X_{0,n}}^U = 1$, which leads to the lower bound given in (A29) with appropriately plugged-in quantities. As in the above proof of the lower bound $E_{y, X_{0,n}}^{L, \tan}$, the inequality $I(P_{A,n} || P_{H,n}) \geq E_{k, X_{0,n}}^{L, \sec}$ follows straightforwardly from Lemma A2, Lemma A3 and the three limits

$$\begin{aligned} \frac{\partial p_\lambda}{\partial \lambda} &= \left(\frac{f_A(k)}{f_H(k)} \right)^\lambda f_H(k) \cdot (k+1) \log \left(\frac{f_A(k)}{f_H(k)} \right) - \left(\frac{f_A(k+1)}{f_H(k+1)} \right)^\lambda f_H(k+1) \cdot k \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) \\ \xrightarrow{\lambda \nearrow 1} & f_A(k)(k+1) \log \left(\frac{f_A(k)}{f_H(k)} \right) - f_A(k+1)k \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right), \\ \frac{\partial q_\lambda}{\partial \lambda} &= \left(\frac{f_A(k+1)}{f_H(k+1)} \right)^\lambda f_H(k+1) \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) - \left(\frac{f_A(k)}{f_H(k)} \right)^\lambda f_H(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) \\ \xrightarrow{\lambda \nearrow 1} & f_A(k+1) \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) - f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) =: l, \quad \text{and} \\ \frac{\partial}{\partial \lambda} \left(\frac{p_\lambda}{q_\lambda} \beta_\lambda - \alpha_\lambda \right) &= \frac{q_\lambda \left(\frac{\partial p_\lambda}{\partial \lambda} \right) - p_\lambda \left(\frac{\partial q_\lambda}{\partial \lambda} \right)}{(q_\lambda)^2} \cdot \beta_\lambda + \frac{p_\lambda}{q_\lambda} \cdot (\beta_A - \beta_H) - (\alpha_A - \alpha_H) \\ \xrightarrow{\lambda \nearrow 1} & f_A(k) \log \left(\frac{f_A(k)}{f_H(k)} \right) \left(k+1 + \frac{\alpha_A}{\beta_A} \right) - f_A(k+1) \log \left(\frac{f_A(k+1)}{f_H(k+1)} \right) \left(k + \frac{\alpha_A}{\beta_A} \right) - \frac{\alpha_A \beta_H}{\beta_A} + \alpha_H. \end{aligned}$$

To construct the third lower bound $E_{X_{0,n}}^{L, hor}$ (cf. (74)), we start by using the horizontal line $\phi_\lambda^{\text{hor}}(\cdot)$ (cf. (54)) as an upper bound of ϕ_λ . For each fixed $\lambda \in]0, 1[$, it is defined by the intercept $\sup_{x \in \mathbb{N}_0} \phi_\lambda(x)$. On $\mathcal{P}_{\text{SP},3a} \cup \mathcal{P}_{\text{SP},3b}$, this supremum is achieved at the finite integer point $z_\lambda^* := \arg \max_{x \in \mathbb{N}_0} \phi_\lambda(x)$ (since $\lim_{x \rightarrow \infty} \phi_\lambda(x) = -\infty$) and there holds $\phi_\lambda(z_\lambda^*) < 0$ which leads with the parameters $q_\lambda = \beta_\lambda$, $p_\lambda = \phi_\lambda(z_\lambda^*) + \alpha_\lambda$ to the Hellinger integral upper bound $B_{\lambda, X_{0,n}}^U = \exp \{ \phi_\lambda(z_\lambda^*) \cdot n \} < 1$ (cf. Remark 1 (b)). We strive for computing the limit $\lim_{\lambda \nearrow 1} \frac{1 - B_{\lambda, X_{0,n}}^U}{\lambda(1-\lambda)}$, which is not straightforward to solve since in general it seems to be intractable to express z_λ^* explicitly in terms of λ . To circumvent this problem, we notice that it is sufficient to determine z_λ^* in a small ϵ -environment $]1 - \epsilon, 1[$. To accomplish this, we incorporate $\lim_{\lambda \nearrow 1} \phi_\lambda(x) = 0$ for all $x \in [0, \infty[$ and calculate by using l'Hospital's rule

$$\lim_{\lambda \nearrow 1} \frac{\phi_\lambda(x)}{1-\lambda} = (\alpha_A + \beta_A x) \left[-\log \left(\frac{\alpha_A + \beta_A x}{\alpha_H + \beta_H x} \right) + 1 \right] - (\alpha_H + \beta_H x).$$

Accordingly, let us define $z^* := \arg \max_{x \in \mathbb{N}_0} \left\{ (\alpha_A + \beta_A x) \left[-\log \left(\frac{\alpha_A + \beta_A x}{\alpha_H + \beta_H x} \right) + 1 \right] - (\alpha_H + \beta_H x) \right\}$ (note that the maximum exists since $\lim_{x \rightarrow \infty} \left\{ (\alpha_A + \beta_A x) \left[-\log \left(\frac{\alpha_A + \beta_A x}{\alpha_H + \beta_H x} \right) + 1 \right] - (\alpha_H + \beta_H x) \right\} = -\infty$). Due to continuity of the function $(\lambda, x) \mapsto \frac{\phi_\lambda(x)}{1-\lambda}$, there exists an $\epsilon > 0$ such that for all $\lambda \in]1 - \epsilon, 1[$ there holds $z_\lambda^* = z^*$. Applying these considerations, we get with l'Hospital's rule

$$I(P_{A,n} || P_{H,n}) \geq \lim_{\lambda \nearrow 1} \frac{1 - \exp \{ \phi_\lambda(z^*) \cdot n \}}{\lambda(1-\lambda)} = \left[f_A(z^*) \cdot \left[\log \left(\frac{f_A(z^*)}{f_H(z^*)} \right) - 1 \right] + f_H(z^*) \right] \cdot n \geq 0. \quad (\text{A33})$$

In fact, for the current parameter constellation $\mathcal{P}_{\text{SP},3a} \cup \mathcal{P}_{\text{SP},3b}$ we have $\phi_\lambda(x) < 0$ for all $\lambda \in]0, 1[$ and all $x \in \mathbb{N}_0$ which implies $f_A(z^*) \neq f_H(z^*)$ by Lemma A1; thus, we even get $E_{X_{0,n}}^{L, hor} > 0$ for all $n \in \mathbb{N}$ by virtue of the inequality $-\log \left(\frac{f_H(z^*)}{f_A(z^*)} \right) > -\frac{f_H(z^*)}{f_A(z^*)} + 1$.

For the case $\mathcal{P}_{SP,2}$, the above-mentioned procedure leads to $z_\lambda^* = 0 = z^*$ ($\lambda \in]0, 1[$) which implies $\phi_\lambda(z_\lambda^*) = 0$, $B_{\lambda, X_0, n}^U \equiv 1$ and thus the trivial lower bound $E_{X_0, n}^{L, hor} = \lim_{\lambda \nearrow 1} \frac{1 - B_{\lambda, X_0, n}^U}{\lambda(1-\lambda)} = 0$ follows for all $n \in \mathbb{N}$. In contrast, for the case $\mathcal{P}_{SP,3c}$ one gets $z_\lambda^* = \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A} = z^*$ ($\lambda \in]0, 1[$) which nevertheless also implies $\phi_\lambda(z_\lambda^*) = 0$ and hence $E_{X_0, n}^{L, hor} \equiv 0$. On $\mathcal{P}_{SP,4}$, we have $\sup_{x \in \mathbb{N}_0} \phi_\lambda(x) = \lim_{x \rightarrow \infty} \phi_\lambda(x) = 0$ and hence we set $E_{X_0, n}^{L, hor} \equiv 0$.

To show the strict positivity $E_{X_0, n}^L > 0$ in the parameter case $\mathcal{P}_{SP,2}$, we inspect the bound $E_{0, X_0, n}^{L, sec}$. With $\alpha := \alpha_\bullet := \alpha_A = \alpha_H$ (the bullet will be omitted in this proof) and the auxiliary variable $x := \frac{\beta_H}{\beta_A} > 0$, the definition (73) respectively its special case (76) rewrites for all $n \in \mathbb{N}$ as

$$E_{0, X_0, n}^{L, sec} := E_{0, X_0, n}^{L, sec}(x) := \begin{cases} \left[-(\alpha + \beta_A) \cdot \log\left(\frac{\alpha + \beta_A x}{\alpha + \beta_A}\right) + \beta_A(x - 1) \right] \cdot \frac{1 - (\beta_A)^n}{1 - \beta_A} \cdot \left[X_0 - \frac{\alpha}{1 - \beta_A} \right] \\ + \left[\frac{\alpha}{\beta_A(1 - \beta_A)} \left(-(\alpha + \beta_A) \cdot \log\left(\frac{\alpha + \beta_A x}{\alpha + \beta_A}\right) + \beta_A(x - 1) \right) \right. \\ \left. + \frac{\alpha}{\beta_A} (\alpha + \beta_A) \cdot \log\left(\frac{\alpha + \beta_A x}{\alpha + \beta_A}\right) - \alpha(x - 1) \right] \cdot n, & \text{if } \beta_A \neq 1, \\ \left[-(\alpha + 1) \cdot \log\left(\frac{\alpha + x}{\alpha + 1}\right) + x - 1 \right] \cdot \left[\frac{\alpha}{2} \cdot n^2 + \left(X_0 + \frac{\alpha}{2}\right) \cdot n \right] \\ + \left[(\alpha + 1) \cdot \log\left(\frac{\alpha + x}{\alpha + 1}\right) - x + 1 \right] \cdot \alpha \cdot n, & \text{if } \beta_A = 1. \end{cases} \quad (A34)$$

To prove that $E_{0, X_0, n}^{L, sec} > 0$ for all $X_0 \in \mathbb{N}$ and all $n \in \mathbb{N}$ it suffices to show that $E_{0, X_0, n}^{L, sec}(1) = \left(\frac{\partial}{\partial x} E_{0, X_0, n}^{L, sec}\right)(1) = 0$ and $\left(\frac{\partial^2}{\partial x^2} E_{0, X_0, n}^{L, sec}\right)(x) > 0$ for all $x \in]0, \infty[\setminus \{1\}$. The assertion $E_{0, X_0, n}^{L, sec}(1) = 0$ is trivial from (A34). Moreover, we obtain

$$\left(\frac{\partial}{\partial x} E_{0, X_0, n}^{L, sec}\right)(x) = \begin{cases} \beta_A \cdot \left[1 - \frac{\alpha + \beta_A}{\alpha + \beta_A x} \right] \cdot \frac{1 - (\beta_A)^n}{1 - \beta_A} \cdot \left[X_0 - \frac{\alpha}{1 - \beta_A} \right] \\ + \alpha \cdot \left(1 - \frac{\alpha + \beta_A}{\alpha + \beta_A x} \right) \cdot \frac{\beta_A}{1 - \beta_A} \cdot n, & \text{if } \beta_A \neq 1, \\ \left[1 - \frac{\alpha + 1}{\alpha + x} \right] \cdot \left[\frac{\alpha}{2} \cdot n^2 + \left(X_0 - \frac{\alpha}{2}\right) \cdot n \right], & \text{if } \beta_A = 1, \end{cases}$$

which immediately yields $\left(\frac{\partial}{\partial x} E_{0, X_0, n}^{L, sec}\right)(1) = 0$. For the second derivative we get

$$\left(\frac{\partial^2}{\partial x^2} E_{0, X_0, n}^{L, sec}\right)(x) = \begin{cases} \frac{(\alpha + \beta_A) \cdot \beta_A^2}{(\alpha + \beta_A x)^2} \cdot \frac{1 - (\beta_A)^n}{1 - \beta_A} \cdot \left[X_0 - \frac{\alpha}{1 - \beta_A} \right] \\ + \alpha \frac{\alpha + \beta_A}{(\alpha + \beta_A x)^2} \cdot \frac{\beta_A^2}{1 - \beta_A} \cdot n > 0, & \text{if } \beta_A \neq 1, \\ \frac{\alpha + 1}{(\alpha + x)^2} \cdot \left[\frac{\alpha}{2} \cdot n^2 + \left(X_0 - \frac{\alpha}{2}\right) \cdot n \right] > 0, & \text{if } \beta_A = 1, \end{cases} \quad (A35)$$

where the strict positivity of $E_{0, X_0, n}^{L, sec}$ in the case $\beta_A \neq 1$ follows immediately by replacing X_0 with 0 and by using the obvious relation $\frac{1}{1 - \beta_A} \cdot \left[n - \frac{1 - \beta_A^n}{1 - \beta_A} \right] = \frac{1}{1 - \beta_A} \sum_{k=0}^{n-1} (1 - \beta_A^k) > 0$. The strict positivity in the case $\beta_A = 1$ is trivial by inspection.

For the constellation $\mathcal{P}_{SP,4}$ with parameters $\beta := \beta_\bullet := \beta_A = \beta_H$, $\alpha_A \neq \alpha_H$, the strict positivity of $E_{X_0, n}^L > 0$ follows by showing that $E_{y, X_0, n}^{L, tan}$ converges from above to zero as y tends to infinity. This is done by proving $\lim_{y \rightarrow \infty} y \cdot E_{y, X_0, n}^{L, tan} \in]0, \infty[$. To see this, let us first observe that by l'Hospital's rule we get

$$\lim_{y \rightarrow \infty} y \cdot \log\left(\frac{\alpha_A + \beta y}{\alpha_H + \beta y}\right) = \frac{\alpha_A - \alpha_H}{\beta} \quad \text{as well as} \quad \lim_{y \rightarrow \infty} y \cdot \left(1 - \frac{\alpha_A + \beta y}{\alpha_H + \beta y} \right) = -\frac{\alpha_A - \alpha_H}{\beta}.$$

From this and (72), we obtain $\lim_{y \rightarrow \infty} y \cdot E_{y, X_0, n}^{L, tan} = \frac{(\alpha_A - \alpha_H)^2}{\beta} \cdot n > 0$ in both cases $\beta \neq 1$ and $\beta = 1$.

Finally, for the parameter case $\mathcal{P}_{SP,3c}$ we consider the bound $E_{y^*, X_0, n}^{L, tan}$, with $y^* = \frac{\alpha_A - \alpha_H}{\beta_H - \beta_A}$. Since $\alpha_A + \beta_A y^* = \alpha_H + \beta_H y^*$, it is easy to see that $E_{y^*, X_0, n}^{L, tan} = 0$ for all $n \in \mathbb{N}$. However, the condition $\left(\frac{\partial}{\partial y} E_{y, X_0, n}^{L, tan}\right)(y^*) \neq 0$ implies that $\sup_{y \geq 0} E_{y, X_0, n}^{L, tan} > 0$. The explicit form (75) of this condition follows from

$$\left(\frac{\partial}{\partial y} E_{y, X_0, n}^{L, tan}\right)(y) = \begin{cases} \frac{(\alpha_A \beta_H - \alpha_H \beta_A)^2}{f_A(y)(f_H(y))^2} \cdot \frac{1 - (\beta_A)^n}{1 - \beta_A} \cdot \left[X_0 - \frac{\alpha_A}{1 - \beta_A}\right] \\ + \frac{\alpha_A \beta_H - \alpha_H \beta_A}{(f_H(y))^2} \cdot \left[\frac{\alpha_A}{\beta_A(1 - \beta_A)f_A(y)} - \frac{\alpha_A \beta_H - \alpha_H \beta_A}{\beta_A}\right] \cdot n, & \text{if } \beta_A \neq 1, \\ \frac{(\alpha_A \beta_H - \alpha_H \beta_A)^2}{f_A(y)(f_H(y))^2} \cdot \left[\frac{\alpha_A}{2} \cdot n^2 + (X_0 + \frac{\alpha_A}{2}) \cdot n\right] - \frac{(\alpha_A \beta_H - \alpha_H \beta_A)^2}{(f_H(y))^2} \cdot n, & \text{if } \beta_A = 1, \end{cases}$$

$y \geq 0$, by using the particular choice $y = y^*$ together with $f_A(y^*) = f_H(y^*) = -\frac{\alpha_A \beta_H - \alpha_H \beta_A}{\beta_A - \beta_H}$. \square

Appendix A.3. Proofs and Auxiliary Lemmas for Section 6

Proof of Lemma 2. A closed-form representation of a sequence $(\tilde{a}_n)_{n \in \mathbb{N}_0}$ defined in (83) to (85) is given by the formula

$$\tilde{a}_n = \sum_{k=0}^{n-1} (c + \rho_k) d^{n-1-k}. \tag{A36}$$

This can be seen by induction: from (83) we obtain with $\tilde{a}_0 = 0$ for the first element $\tilde{a}_1 = c + \rho_0 = \sum_{k=0}^0 (c + \rho_k) d^{-k}$. Supposing that (A36) holds for the n -th element, the induction step is

$$\tilde{a}_{n+1} = c + d \cdot \tilde{a}_n + \rho_n = c + d \cdot \sum_{k=0}^{n-1} (c + \rho_k) d^{n-1-k} + \rho_n = \sum_{k=0}^n (c + \rho_k) d^{n-k}.$$

In order to obtain the explicit representation of \tilde{a}_n , we consider first the case $0 \leq v < \varkappa < d$ and $\rho_n = K_1 \cdot \varkappa^n + K_2 \cdot v^n$, which leads to

$$\begin{aligned} \tilde{a}_n &= d^{n-1} \sum_{k=0}^{n-1} \left(c \cdot d^{-k} + K_1 \cdot \left(\frac{\varkappa}{d}\right)^k + K_2 \cdot \left(\frac{v}{d}\right)^k \right) \\ &= d^{n-1} \cdot \left[c \cdot \frac{1 - d^{-n}}{1 - d^{-1}} + K_1 \cdot \frac{1 - \left(\frac{\varkappa}{d}\right)^n}{1 - \frac{\varkappa}{d}} + K_2 \cdot \frac{1 - \left(\frac{v}{d}\right)^n}{1 - \frac{v}{d}} \right] \\ &= \frac{c}{1 - d} (1 - d^n) + K_1 \cdot \frac{d^n - \varkappa^n}{d - \varkappa} + K_2 \cdot \frac{d^n - v^n}{d - v}. \end{aligned} \tag{A37}$$

Hence, for the corresponding sum we get

$$\begin{aligned} \sum_{k=1}^n \tilde{a}_k &= \sum_{k=1}^n \left[\frac{c}{1 - d} + \left(\frac{K_1}{d - \varkappa} + \frac{K_2}{d - v} - \frac{c}{1 - d} \right) \cdot d^k - \frac{K_1}{d - \varkappa} \cdot \varkappa^k - \frac{K_2}{d - v} \cdot v^k \right] \\ &= \frac{c}{1 - d} \cdot n + \left(\frac{K_1}{d - \varkappa} + \frac{K_2}{d - v} - \frac{c}{1 - d} \right) \cdot \frac{d \cdot (1 - d^n)}{1 - d} - \frac{K_1 \cdot \varkappa \cdot (1 - \varkappa^n)}{(d - \varkappa)(1 - \varkappa)} - \frac{K_2 \cdot v \cdot (1 - v^n)}{(d - v)(1 - v)}. \end{aligned} \tag{A38}$$

Consider now the case $0 \leq v < \varkappa = d$. Then some expressions in (A37) and (A38) have a zero denominator. In this case, the evaluation of (A36) becomes

$$\begin{aligned} \tilde{a}_n &= d^{n-1} \sum_{k=0}^{n-1} \left(c \cdot d^{-k} + K_1 + K_2 \cdot \left(\frac{v}{d}\right)^k \right) = d^{n-1} \cdot \left[c \cdot \frac{1 - d^{-n}}{1 - d^{-1}} + K_1 \cdot n + K_2 \cdot \frac{1 - \left(\frac{v}{d}\right)^n}{1 - \frac{v}{d}} \right] \\ &= \frac{c}{1 - d} (1 - d^n) + K_1 \cdot n \cdot d^{n-1} + K_2 \cdot \frac{d^n - v^n}{d - v}. \end{aligned} \tag{A39}$$

Before we calculate the corresponding sum $\sum_{k=1}^n \tilde{a}_k$, we notice that

$$\sum_{k=1}^n k \cdot d^{k-1} = \sum_{k=1}^n \frac{\partial}{\partial d} d^k = \frac{\partial}{\partial d} \sum_{k=1}^n d^k = \frac{\partial}{\partial d} \left(\frac{d \cdot (1 - d^n)}{1 - d} \right) = \frac{1 - n \cdot d^n(1 - d) - d^n}{(1 - d)^2}.$$

Using this fact, we obtain

$$\begin{aligned} \sum_{k=1}^n \tilde{a}_k &= \sum_{k=1}^n \left[\frac{c}{1-d}(1-d^k) + K_1 \cdot k \cdot d^{k-1} + K_2 \cdot \frac{d^k - v^k}{d-v} \right] \\ &= \frac{c}{1-d} \cdot n + \sum_{k=1}^n \left(\frac{K_2}{d-v} - \frac{c}{1-d} \right) d^k + K_1 \sum_{k=1}^n k \cdot d^{k-1} - \frac{K_2}{d-v} \sum_{k=1}^n v^k \\ &= \left(\frac{K_2}{d-v} - \frac{c}{1-d} \right) \frac{d \cdot (1 - d^n)}{1-d} + K_1 \cdot \frac{1 - n \cdot d^n(1 - d) - d^n}{(1 - d)^2} - \frac{K_2 \cdot v(1 - v^n)}{(d - v)(1 - v)} + \frac{c}{1-d} \cdot n \\ &= \left(\frac{K_1}{d(1-d)} + \frac{K_2}{d-v} - \frac{c}{1-d} \right) \frac{d \cdot (1 - d^n)}{1-d} - \frac{K_2 \cdot v(1 - v^n)}{(d - v)(1 - v)} + \left(\frac{c}{1-d} - \frac{K_1 \cdot d^n}{1-d} \right) \cdot n. \quad \square \end{aligned}$$

Proof of Lemma 3. (a) In this case we have $0 < q < \beta_\lambda$. To prove part (i), we consider the function $\zeta_\lambda^{(q)}(\cdot)$ on $[x_0^{(q)}, 0]$, the range of the sequence $(a_n^{(q)})_{n \in \mathbb{N}}$ (recall Properties 1 (P1)). For tackling the left-hand inequality in (i), we compare $\zeta_\lambda^{(q)}(x) = q \cdot e^x - \beta_\lambda$ with the quadratic function

$$\underline{Y}_\lambda^{(q)}(x) := \frac{q}{2} e^{x_0^{(q)}} \cdot x^2 + qe^{x_0^{(q)}} (1 - x_0^{(q)}) \cdot x + x_0^{(q)} \left(1 - qe^{x_0^{(q)}} + \frac{q}{2} e^{x_0^{(q)}} x_0^{(q)} \right). \tag{A40}$$

Clearly, one has the relations $\underline{Y}_\lambda^{(q)}(x_0^{(q)}) = x_0^{(q)} = \zeta_\lambda^{(q)}(x_0^{(q)})$, $\underline{Y}_\lambda^{(q)'}(x_0^{(q)}) = q \cdot e^{x_0^{(q)}} = \zeta_\lambda^{(q)'}(x_0^{(q)})$, and $\underline{Y}_\lambda^{(q)''}(x) < \zeta_\lambda^{(q)''}(x)$ for all $x \in]x_0^{(q)}, 0]$. Hence, $\underline{Y}_\lambda^{(q)}(\cdot)$ is on $]x_0^{(q)}, 0]$ a strict lower functional bound of $\zeta_\lambda^{(q)}(\cdot)$. We are now ready to prove the left-hand inequality in (i) by induction. For $n = 1$, we easily see that $\underline{a}_1^{(q)} < a_1^{(q)}$ iff $x_0^{(q)} \left(1 - qe^{x_0^{(q)}} + \frac{q}{2} e^{x_0^{(q)}} x_0^{(q)} \right) < q - \beta_\lambda$ iff $\underline{Y}_\lambda^{(q)}(0) < \zeta_\lambda^{(q)}(0)$, and the latter is obviously true. Let us assume that $\underline{a}_n^{(q)} \leq a_n^{(q)}$ holds. From this, (93), (78) and (80) we obtain

$$\begin{aligned} 0 < \underline{\rho}_n^{(q)} &= \frac{q}{2} e^{x_0^{(q)}} \left(x_0^{(q)} \cdot \left(q \cdot e^{x_0^{(q)}} \right)^n \right)^2 = \frac{q}{2} e^{x_0^{(q)}} \left(a_n^{(q),T} - x_0^{(q)} \right)^2 \\ &< \frac{q}{2} e^{x_0^{(q)}} \left(a_n^{(q)} - x_0^{(q)} \right)^2 = \underline{Y}_\lambda^{(q)} \left(a_n^{(q)} \right) - d^{(q),T} \cdot a_n^{(q)} - x_0^{(q)} \cdot \left(1 - d^{(q),T} \right) \\ &< \zeta_\lambda^{(q)} \left(a_n^{(q)} \right) - d^{(q),T} \cdot a_n^{(q)} - x_0^{(q)} \cdot \left(1 - d^{(q),T} \right) \\ &< a_{n+1}^{(q)} - d^{(q),T} \cdot \underline{a}_n^{(q)} - x_0^{(q)} \cdot \left(1 - d^{(q),T} \right) = a_{n+1}^{(q)} - \zeta_\lambda^{(q),T}(\underline{a}_n^{(q)}). \end{aligned}$$

Thus, there holds $\underline{a}_{n+1}^{(q)} < a_{n+1}^{(q)}$. For the right-hand inequality in (i), we proceed analogously:

$$\bar{Y}_\lambda^{(q)}(x) := \frac{q}{2} e^{x_0^{(q)}} \cdot x^2 + \left(1 - \frac{q}{2} e^{x_0^{(q)}} x_0^{(q)} - \frac{q - \beta_\lambda}{x_0^{(q)}} \right) \cdot x + q - \beta_\lambda \tag{A41}$$

satisfies $\bar{Y}_\lambda^{(q)}(x_0^{(q)}) = x_0^{(q)} = \zeta_\lambda^{(q)}(x_0^{(q)})$, $\bar{Y}_\lambda^{(q)}(0) = q - \beta_\lambda = \zeta_\lambda^{(q)}(0)$ as well as $\bar{Y}_\lambda^{(q)''}(x) < \zeta_\lambda^{(q)''}(x)$ for all $x \in]x_0^{(q)}, 0]$. Hence, $\bar{Y}_\lambda^{(q)}(\cdot)$ is on $]x_0^{(q)}, 0]$ a strict upper functional bound of $\zeta_\lambda^{(q)}(\cdot)$. Let us first observe the

obvious relation $\bar{a}_1^{(q)} = q - \beta_\lambda = a_1^{(q)} < 0$, and assume that $\bar{a}_n^{(q)} \geq a_n^{(q)}$ ($n \in \mathbb{N}$) holds. From this, (95), (79), and (80) we obtain the desired inequality $\bar{a}_{n+1}^{(q)} > a_{n+1}^{(q)}$ by

$$\begin{aligned} 0 > \bar{\rho}_n^{(q)} &= -\Gamma_{<}^{(q)} \left(d^{(q),T} \right)^n \cdot \frac{a_n^{(q),S}}{x_0^{(q)}} = \frac{q}{2} e^{x_0^{(q)}} \left(a_n^{(q),T} - x_0^{(q)} \right) \cdot a_n^{(q),S} \\ &\geq \frac{q}{2} e^{x_0^{(q)}} \left(a_n^{(q)} - x_0^{(q)} \right) \cdot a_n^{(q)} = \bar{Y}_\lambda^{(q)} \left(a_n^{(q)} \right) - d^{(q),S} \cdot a_n^{(q)} - (q - \beta_\lambda) \\ &> \bar{\zeta}_\lambda^{(q)} \left(a_n^{(q)} \right) - d^{(q),S} \cdot a_n^{(q)} - (q - \beta_\lambda) \geq a_{n+1}^{(q)} - d^{(q),S} \cdot \bar{a}_n^{(q)} - (q - \beta_\lambda) = a_{n+1}^{(q)} - \bar{\zeta}_\lambda^{(q),S} \left(\bar{a}_n^{(q)} \right). \end{aligned}$$

The explicit representations of the sequences $\left(a_n^{(q)} \right)_{n \in \mathbb{N}}$, $\left(\underline{a}_n^{(q)} \right)_{n \in \mathbb{N}}$ and $\left(\bar{a}_n^{(q)} \right)_{n \in \mathbb{N}}$ follow from (86) by incorporating the appropriate constants mentioned in the prelude of Lemma 3. With (83) to (85) and (86) we immediately achieve $\underline{a}_n^{(q)} > a_n^{(q),T}$ for all $n \in \mathbb{N}$. Analogously, for all $n \geq 2$, we get $\bar{\rho}_{n-1} < 0$, which implies that $\bar{a}_n^{(q)} < a_n^{(q),S}$ for all $n \geq 2$. For $n = 1$ one obtains $\bar{\rho}_0 = 0$ as well as $\bar{a}_1^{(q)} = a_1^{(q),S} = a_1^{(q)} = q - \beta_\lambda$.

For the second part (ii), we employ the representation (A36) which leads to

$$\begin{aligned} \underline{a}_n^{(q)} &= \sum_{k=0}^{n-1} \left(d^{(q),T} \right)^{n-1-k} \cdot \left(\underline{\rho}_k^{(q)} + x_0^{(q)} \cdot (1 - d^{(q),T}) \right) \\ \text{as well as} \quad \bar{a}_n^{(q)} &= \sum_{k=0}^{n-1} \left(d^{(q),S} \right)^{n-1-k} \cdot \left(\bar{\rho}_k^{(q)} + (q - \beta_\lambda) \right). \end{aligned}$$

The strict decreasingness of both sequences follows from

$$\underline{\rho}_k^{(q)} + x_0^{(q)} (1 - d^{(q),T}) = \frac{q e^{x_0^{(q)}}}{2} \left(x_0^{(q)} \right)^2 \left(d^{(q),T} \right)^{2n} + x_0^{(q)} \left(1 - d^{(q),T} \right) \leq \underline{Y}_\lambda^{(q)}(0) < \bar{\zeta}_\lambda^{(q)}(0) = q - \beta_\lambda < 0$$

and from the fact that $\bar{\rho}_k^{(q)} \leq 0$ for all $k \in \mathbb{N}_0$ and $q < \beta_\lambda$. Part (iii) follows directly from (i), since $d^{(q),T}, d^{(q),S} \in]0, 1[$.

Let us now prove part (b), where $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}$ is assumed. To tackle part (i), we compare $\bar{\zeta}_\lambda^{(q)}(x) = q \cdot e^x - \beta_\lambda$ with the quadratic function

$$\underline{v}_\lambda^{(q)}(x) := \frac{q}{2} \cdot x^2 + q \cdot \left(e^{x_0^{(q)}} - x_0^{(q)} \right) \cdot x + x_0^{(q)} \left(1 - q e^{x_0^{(q)}} + \frac{q}{2} x_0^{(q)} \right) > 0 \tag{A42}$$

on the interval $[0, x_0^{(q)}]$. Clearly, we have $\underline{v}_\lambda^{(q)}(x_0^{(q)}) = \bar{\zeta}_\lambda^{(q)}(x_0^{(q)}) = x_0^{(q)}$, $\underline{v}_\lambda^{(q)'}(x_0^{(q)}) = \bar{\zeta}_\lambda^{(q)'}(x_0^{(q)}) = q e^{x_0^{(q)}}$ and $0 < \underline{v}_\lambda^{(q)''}(x) < \bar{\zeta}_\lambda^{(q)''}(x)$ for all $x \in]0, x_0^{(q)}[$. Thus, $\underline{v}_\lambda^{(q)}(\cdot)$ constitutes a positive functional lower bound for $\bar{\zeta}_\lambda^{(q)}(\cdot)$ on $[0, x_0^{(q)}]$. Let us now prove the left-hand inequality of (i) by induction: for $n = 1$ we get $\underline{a}_1^{(q)} = \underline{v}_\lambda^{(q)}(0) < \bar{\zeta}_\lambda^{(q)}(0) = a_1^{(q)}$. Moreover, by assuming $\underline{a}_n^{(q)} \leq a_n^{(q)}$ for $n \in \mathbb{N}$, we obtain with the above-mentioned considerations and (93), (80) and (82)

$$\begin{aligned} 0 < \underline{\rho}_n^{(q)} &= \Gamma_{>}^{(q)} \left(d^{(q),S} \right)^{2n} = \frac{q}{2} \cdot \left(a_n^{(q),S} - x_0^{(q)} \right)^2 < \frac{q}{2} \cdot \left(a_n^{(q)} - x_0^{(q)} \right)^2 \\ &= \frac{q}{2} \left(a_n^{(q)} \right)^2 + q \cdot \left(e^{x_0^{(q)}} - x_0^{(q)} \right) \cdot a_n^{(q)} + x_0^{(q)} \cdot \left(1 - q e^{x_0^{(q)}} + \frac{q}{2} x_0^{(q)} \right) - d^{(q),T} a_n^{(q)} - c^{(q),T} \\ &= \underline{v}_\lambda^{(q)} \left(a_n^{(q)} \right) - d^{(q),T} a_n^{(q)} - c^{(q),T} < \bar{\zeta}_\lambda^{(q)} \left(a_n^{(q)} \right) - d^{(q),T} a_n^{(q)} - c^{(q),T} \\ &< a_{n+1}^{(q)} - d^{(q),T} \underline{a}_n^{(q)} - c^{(q),T} = a_{n+1}^{(q)} - \bar{\zeta}_\lambda^{(q),T} \left(\underline{a}_n^{(q)} \right). \end{aligned}$$

Hence, $\underline{a}_{n+1}^{(q)} < a_{n+1}^{(q)}$. For the right-hand inequality in part (i), we define the quadratic function

$$\bar{v}_\lambda^{(q)}(x) := \frac{q}{2} \cdot x^2 + \left(1 - \frac{q}{2}x_0^{(q)} - \frac{q - \beta_\lambda}{x_0^{(q)}}\right) \cdot x + q - \beta_\lambda, \tag{A43}$$

which is a functional upper bound for $\zeta_\lambda^{(q)}(\cdot)$ on the interval $[0, x_0^{(q)}]$ since there holds $\bar{v}_\lambda^{(q)}(0) = \zeta_\lambda^{(q)}(0) = q - \beta_\lambda$, $\bar{v}_\lambda^{(q)}(x_0^{(q)}) = \zeta_\lambda^{(q)}(x_0^{(q)}) = x_0^{(q)}$ and additionally $\bar{v}_\lambda^{(q)''}(x) = q < qe^x = \zeta_\lambda^{(q)''}(x)$ on $]0, x_0^{(q)}[$. Obviously, $\bar{a}_1^{(q)} = q - \beta_\lambda = a_1^{(q)}$. By assuming $\bar{a}_n^{(q)} \geq a_n^{(q)}$ for $n \in \mathbb{N}$, we obtain with (80), (82) and (95)

$$\begin{aligned} 0 > \bar{\rho}_n^{(q)} &= -\Gamma_{>}^{(q)} \cdot (d^{(q),S})^n \cdot (1 - (d^{(q),T})^n) = -\frac{q}{2} \cdot (x_0 - a_n^{(q),S}) \cdot a_n^{(q),T} \\ &> -\frac{q}{2} \cdot (x_0 - a_n^{(q)}) \cdot a_n^{(q)} = \bar{v}_\lambda^{(q)}(a_n^{(q)}) - \frac{x_0^{(q)} - (q - \beta_\lambda)}{x_0^{(q)}} \cdot a_n^{(q)} - (q - \beta_\lambda) \\ &> \zeta_\lambda^{(q)}(a_n^{(q)}) - d^{(q),S} a_n^{(q)} - c^{(q),S} > \zeta_\lambda^{(q)}(a_n^{(q)}) - d^{(q),S} \bar{a}_n^{(q),S} - c^{(q),S} = a_{n+1}^{(q)} - \zeta_\lambda^{(q),S}(\bar{a}_n^{(q)}), \end{aligned} \tag{A44}$$

which implies $\bar{a}_{n+1}^{(q)} > a_{n+1}^{(q)}$. The explicit representations of the sequences $(a_n^{(q)})_{n \in \mathbb{N}}$ and $(\bar{a}_n^{(q)})_{n \in \mathbb{N}}$ follow from (86) by employing the appropriate constants mentioned in the prelude of Lemma 3. By means of (83) to (85) and (86), we directly get $a_n^{(q)} > a_n^{(q),T}$ for all $n \in \mathbb{N}$, whereas $\bar{a}_n^{(q)} < a_n^{(q),S}$ holds only for all $n \geq 2$, since $\bar{\rho}_0 = 0$ implies that $\bar{a}_1^{(q)} = a_1^{(q),S} = a_1^{(q)} = q - \beta_\lambda$.

The second part (ii) can be proved in the same way as part (ii) of (a), by employing the representation (A36). For the lower bound one has

$$a_n^{(q)} = \sum_{k=0}^{n-1} (d^{(q),T})^{n-1-k} \cdot [c^{(q),T} + \rho_k^{(q)}], \quad \text{with } c^{(q),T} > 0 \quad \text{and } \rho_k^{(q)} > 0.$$

For the upper bound we get

$$\bar{a}_n^{(q)} = \sum_{k=0}^{n-1} (d^{(q),S})^{n-1-k} \cdot [c^{(q),S} + \bar{\rho}_k^{(q)}],$$

hence it is enough to show $c^{(q),S} + \bar{\rho}_n^{(q)} > 0$ for all $n \in \mathbb{N}_0$. Considering the first two lines of calculation (A44) and incorporating $c^{(q),S} = q - \beta_\lambda$, this can be seen from

$$c^{(q),S} + \bar{\rho}_n^{(q)} > \bar{v}_\lambda^{(q)}(a_n^{(q)}) - \frac{x_0^{(q)} - (q - \beta_\lambda)}{x_0^{(q)}} \cdot a_n^{(q)} = \bar{v}_\lambda^{(q)}(a_n^{(q)}) - d^{(q),S} \cdot a_n^{(q)} > 0,$$

because on $[0, x_0^{(q)}]$ there holds $d^{(q),S} \cdot x < x < \bar{v}_\lambda^{(q)}(x)$. The last part (iii) can be easily deduced from (i) together with $\lim_{n \rightarrow \infty} n \cdot (d^{(q),S})^{n-1} = 0$. \square

The proofs of all Theorems 5–9 are mainly based on the following

Lemma A4. Recall the quantity $\tilde{B}_{\lambda, X_0, n}^{(p,q)}$ from (42) for general $p \geq 0$, $q > 0$ (notice that we do not consider parameters $p < 0$, $q \leq 0$ in Section 6) as well as the constants $d^{(q),T}$, $d^{(q),S}$ and $\Gamma_{<}^{(q)}$, $\Gamma_{>}^{(q)}$ defined in (76), (77) and (91). For all $(\beta_A, \beta_H, \alpha_A, \alpha_H, \lambda) \in \mathcal{P} \times \mathbb{R} \setminus \{0, 1\}$, all initial population sizes $X_0 \in \mathbb{N}$ and all observation horizons $n \in \mathbb{N}$ there holds

(a) in the case $p \geq 0$ and $0 < q < \beta_\lambda$

$$\begin{aligned} \tilde{B}_{\lambda, X_0, n}^{(p, q)} &\geq \exp \left\{ x_0^{(q)} \cdot \left[X_0 - \frac{p}{q} \cdot \frac{d^{(q), T}}{1 - d^{(q), T}} \right] \cdot \left(1 - \left(d^{(q), T} \right)^n \right) + \left(\frac{p}{q} \cdot \left(\beta_\lambda + x_0^{(q)} \right) - \alpha_\lambda \right) \cdot n \right. \\ &\quad \left. + \underline{\zeta}_n^{(q)} \cdot X_0 + \frac{p}{q} \cdot \underline{\vartheta}_n^{(q)} \right\} =: C_{\lambda, X_0, n}^{(p, q), L}, \end{aligned} \tag{A45}$$

$$\begin{aligned} \tilde{B}_{\lambda, X_0, n}^{(p, q)} &\leq \exp \left\{ x_0^{(q)} \cdot \left[X_0 - \frac{p}{q} \cdot \frac{d^{(q), S}}{1 - d^{(q), S}} \right] \cdot \left(1 - \left(d^{(q), S} \right)^n \right) + \left(\frac{p}{q} \cdot \left(\beta_\lambda + x_0^{(q)} \right) - \alpha_\lambda \right) \cdot n \right. \\ &\quad \left. - \bar{\zeta}_n^{(q)} \cdot X_0 - \frac{p}{q} \cdot \bar{\vartheta}_n^{(q)} \right\} =: C_{\lambda, X_0, n}^{(p, q), U}, \end{aligned} \tag{A46}$$

where $\underline{\zeta}_n^{(q)} := \Gamma_{<}^{(q)} \cdot \frac{\left(d^{(q), T} \right)^{n-1}}{1 - d^{(q), T}} \cdot \left(1 - \left(d^{(q), T} \right)^n \right) > 0,$ (A47)

$$\underline{\vartheta}_n^{(q)} := \Gamma_{<}^{(q)} \cdot \frac{1 - \left(d^{(q), T} \right)^n}{\left(1 - d^{(q), T} \right)^2} \cdot \left[1 - \frac{d^{(q), T} \left(1 + \left(d^{(q), T} \right)^n \right)}{1 + d^{(q), T}} \right] > 0, \tag{A48}$$

$$\bar{\zeta}_n^{(q)} := \Gamma_{<}^{(q)} \cdot \left[\frac{\left(d^{(q), S} \right)^n - \left(d^{(q), T} \right)^n}{d^{(q), S} - d^{(q), T}} - \left(d^{(q), S} \right)^{n-1} \cdot \frac{1 - \left(d^{(q), T} \right)^n}{1 - d^{(q), T}} \right] > 0, \tag{A49}$$

$$\bar{\vartheta}_n^{(q)} := \Gamma_{<}^{(q)} \cdot \frac{d^{(q), T}}{1 - d^{(q), T}} \cdot \left[\frac{1 - \left(d^{(q), S} d^{(q), T} \right)^n}{1 - d^{(q), S} d^{(q), T}} - \frac{\left(d^{(q), S} \right)^n - \left(d^{(q), T} \right)^n}{d^{(q), S} - d^{(q), T}} \right] > 0. \tag{A50}$$

(b) in the case $p \geq 0$ and $0 < q = \beta_\lambda$

$$\tilde{B}_{\lambda, X_0, n}^{(p, q)} = \exp \left\{ \left(\frac{p}{q} \cdot \left(\beta_\lambda + x_0^{(q)} \right) - \alpha_\lambda \right) \cdot n \right\} = \exp \{ (p - \alpha_\lambda) \cdot n \}.$$

(c) in the case $p \geq 0$ and $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}$ the bounds $C_{\lambda, X_0, n}^{(p, q), L}$ and $C_{\lambda, X_0, n}^{(p, q), U}$ from (96) and (97) remain valid, but with

$$\underline{\zeta}_n^{(q)} := \Gamma_{>}^{(q)} \cdot \frac{\left(d^{(q), T} \right)^n - \left(d^{(q), S} \right)^{2n}}{d^{(q), T} - \left(d^{(q), S} \right)^2} > 0, \tag{A51}$$

$$\underline{\vartheta}_n^{(q)} := \frac{\Gamma_{>}^{(q)}}{d^{(q), T} - \left(d^{(q), S} \right)^2} \cdot \left[\frac{d^{(q), T} \cdot \left(1 - \left(d^{(q), T} \right)^n \right)}{1 - d^{(q), T}} - \frac{\left(d^{(q), S} \right)^2 \cdot \left(1 - \left(d^{(q), S} \right)^{2n} \right)}{1 - \left(d^{(q), S} \right)^2} \right] > 0, \tag{A52}$$

$$\bar{\zeta}_n^{(q)} := \Gamma_{>}^{(q)} \cdot \left(d^{(q), S} \right)^{n-1} \cdot \left[n - \frac{1 - \left(d^{(q), T} \right)^n}{1 - d^{(q), T}} \right] > 0, \tag{A53}$$

$$\begin{aligned} \bar{\vartheta}_n^{(q)} &:= \Gamma_{>}^{(q)} \cdot \left[\frac{d^{(q), S} - d^{(q), T}}{\left(1 - d^{(q), S} \right)^2 \left(1 - d^{(q), T} \right)} \cdot \left(1 - \left(d^{(q), S} \right)^n \right) \right. \\ &\quad \left. + \frac{d^{(q), T} \left(1 - \left(d^{(q), S} d^{(q), T} \right)^n \right)}{\left(1 - d^{(q), T} \right) \left(1 - d^{(q), S} d^{(q), T} \right)} - \frac{\left(d^{(q), S} \right)^n}{1 - d^{(q), S}} \cdot n \right]. \end{aligned} \tag{A54}$$

(d) for the special choices $p := p_\lambda^E := \alpha_\mathcal{A}^\lambda \alpha_\mathcal{H}^{1-\lambda} > 0$, $q := q_\lambda^E := \beta_\mathcal{A}^\lambda \beta_\mathcal{H}^{1-\lambda} > 0$ in the parameter setup $(\beta_\mathcal{A}, \beta_\mathcal{H}, \alpha_\mathcal{A}, \alpha_\mathcal{H}, \lambda) \in (\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}) \times]\lambda_-, \lambda_+[\setminus \{0, 1\}$ we obtain

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log (V_{\lambda, X_0, n}) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^E, q_\lambda^E), L} \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p_\lambda^E, q_\lambda^E), U} \right) = \frac{\alpha_\mathcal{A}}{\beta_\mathcal{A}} \cdot x_0^{(q_\lambda^E)}.$$

(e) for all general $p \geq 0$ with either $0 < q < \beta_\lambda$ or $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}$ we get

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\tilde{B}_{\lambda, X_0, n}^{(p, q)} \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p, q), L} \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(C_{\lambda, X_0, n}^{(p, q), U} \right) = \frac{p}{q} \cdot \left(\beta_\lambda + x_0^{(q)} \right) - \alpha_\lambda.$$

Proof of Lemma A4. The closed-form bounds $C_{\lambda, X_0, n}^{(p, q), L}$ and $C_{\lambda, X_0, n}^{(p, q), U}$ are obtained by substituting in the representation (42) (for $\tilde{B}_{\lambda, X_0, n}^{(p, q)}$, cf. Theorem 1) the recursive sequence member $a_n^{(q)}$ by the explicit sequence member $\underline{a}_n^{(q)}$ respectively $\bar{a}_n^{(q)}$. From the definitions of these sequences (92) to (95) and from (83) to (85) one can see that we basically have to evaluate the term

$$\exp \left\{ \left(\tilde{a}_n^{hom} + \tilde{c}_n \right) \cdot X_0 + \frac{p}{q} \cdot \sum_{k=1}^n \left(\tilde{a}_k^{hom} + \tilde{c}_k \right) + \left(\frac{p}{q} \cdot \beta_\lambda - \alpha_\lambda \right) \cdot n \right\}, \tag{A55}$$

where $\tilde{a}_n^{hom} + \tilde{c}_n = \tilde{a}_n$ is either interpreted as the lower approximate $\underline{a}_n^{(q)}$ or as the upper approximate $\bar{a}_n^{(q)}$. After rearranging and incorporating that $\frac{c^{(q), S}}{1 - d^{(q), S}} = \frac{c^{(q), T}}{1 - d^{(q), T}} = x_0^{(q)}$ in both approximate cases, we obtain with the help of (86), (87) for the expression (A55) in the case $0 \leq \nu < \varkappa < d$

$$\begin{aligned} & \exp \left\{ x_0^{(q)} \cdot (1 - d^n) \cdot \left[X_0 - \frac{p}{q} \cdot \frac{d}{1 - d} \right] + \left(\frac{p}{q} \cdot \left(\beta_\lambda + x_0^{(q)} \right) - \alpha_\lambda \right) \cdot n \right. \\ & \quad + \left[K_1 \cdot \frac{d^n - \varkappa^n}{d - \varkappa} + K_2 \cdot \frac{d^n - \nu^n}{d - \nu} \right] \cdot X_0 \\ & \quad \left. + \frac{p}{q} \cdot \left[\left(\frac{K_1}{d - \varkappa} + \frac{K_2}{d - \nu} \right) \cdot \frac{d \cdot (1 - d^n)}{1 - d} - \frac{K_1 \cdot \varkappa \cdot (1 - \varkappa^n)}{(d - \varkappa)(1 - \varkappa)} - \frac{K_2 \cdot \nu \cdot (1 - \nu^n)}{(d - \nu)(1 - \nu)} \right] \right\}. \tag{A56} \end{aligned}$$

In the other case $0 \leq \nu < \varkappa = d$, the application of (88), (89) turns (A55) into

$$\begin{aligned} & \exp \left\{ x_0^{(q)} \cdot (1 - d^n) \cdot \left[X_0 - \frac{p}{q} \cdot \frac{d}{1 - d} \right] + \left(\frac{p}{q} \cdot \left(\beta_\lambda + x_0^{(q)} \right) - \alpha_\lambda \right) \cdot n \right. \\ & \quad + \left[K_1 \cdot n \cdot d^{n-1} + K_2 \cdot \frac{d^n - \nu^n}{d - \nu} \right] \cdot X_0 \\ & \quad \left. + \frac{p}{q} \cdot \left[\left(\frac{K_1}{d(1 - d)} + \frac{K_2}{d - \nu} \right) \cdot \frac{d \cdot (1 - d^n)}{1 - d} - \frac{K_2 \cdot \nu \cdot (1 - \nu^n)}{(d - \nu)(1 - \nu)} - \frac{K_1 \cdot d^n}{1 - d} \cdot n \right] \right\}. \tag{A57} \end{aligned}$$

After these preparatory considerations let us now begin with elaboration of the details.

(a) Let $0 < q < \beta_\lambda$. We obtain a closed-form lower bound for $\tilde{B}_{\lambda, X_0, n}^{(p, q)}$ by employing the parameters $c \hat{=} c^{(q), T}$, $d \hat{=} d^{(q), T}$, $K_2 = \nu = 0$, $K_1 = \Gamma_{<}^{(q)}$, and $\varkappa = \left(d^{(q), T} \right)^2$, cf. (93) in combination with (85). Since $\varkappa < d^{(q), T}$, we have to plug in these parameters into (A56). The representations of $\underline{z}_n^{(q)}$ and $\underline{\vartheta}_n^{(q)}$ in (A47) and (A48) follow immediately. For a closed-form upper bound, we employ the parameters $c \hat{=} c^{(q), S}$, $d \hat{=} d^{(q), S}$, $-K_1 = K_2 = \Gamma_{<}^{(q)}$, $\varkappa = d^{(q), T}$ and $\nu = d^{(q), S} d^{(q), T}$ (in particular, $\varkappa < d^{(q), S}$ implying that

we have to use (A56). From this, (A49) can be deduced directly; the representation (A50) comes from the expressions in the squared brackets in the last line of (A56) and from

$$\begin{aligned}
 & - \left(\frac{\Gamma_{<}^{(q)}}{d^{(q),S} - d^{(q),T}} - \frac{\Gamma_{<}^{(q)}}{d^{(q),S} - d^{(q),S}d^{(q),T}} \right) \cdot \frac{d^{(q),S} \cdot \left(1 - \left(d^{(q),S}\right)^n\right)}{1 - d^{(q),S}} + \frac{\Gamma_{<}^{(q)} \cdot d^{(q),T} \cdot \left(1 - \left(d^{(q),T}\right)^n\right)}{\left(d^{(q),S} - d^{(q),T}\right) \left(1 - d^{(q),T}\right)} \\
 & \quad - \frac{\Gamma_{<}^{(q)} \cdot d^{(q),S}d^{(q),T} \cdot \left(1 - \left(d^{(q),S}d^{(q),T}\right)^n\right)}{\left(d^{(q),S} - d^{(q),S}d^{(q),T}\right) \left(1 - d^{(q),S}d^{(q),T}\right)} \\
 & = - \frac{\Gamma_{<}^{(q)} \cdot d^{(q),T} \left(1 - d^{(q),S}\right)}{d^{(q),S} \left(d^{(q),S} - d^{(q),T}\right) \left(1 - d^{(q),T}\right)} \cdot \frac{d^{(q),S} \cdot \left(1 - \left(d^{(q),S}\right)^n\right)}{1 - d^{(q),S}} + \frac{\Gamma_{<}^{(q)} \cdot d^{(q),T} \cdot \left(1 - \left(d^{(q),T}\right)^n\right)}{\left(d^{(q),S} - d^{(q),T}\right) \left(1 - d^{(q),T}\right)} \\
 & \quad - \frac{\Gamma_{<}^{(q)} \cdot d^{(q),T} \cdot \left(1 - \left(d^{(q),S}d^{(q),T}\right)^n\right)}{\left(1 - d^{(q),T}\right) \left(1 - d^{(q),S}d^{(q),T}\right)} \\
 & = - \frac{\Gamma_{<}^{(q)} \cdot d^{(q),T}}{1 - d^{(q),T}} \cdot \left[\frac{1 - \left(d^{(q),S}d^{(q),T}\right)^n}{1 - d^{(q),S}d^{(q),T}} + \frac{1 - \left(d^{(q),S}\right)^n}{d^{(q),S} - d^{(q),T}} - \frac{1 - \left(d^{(q),T}\right)^n}{d^{(q),S} - d^{(q),T}} \right] \\
 & = - \frac{\Gamma_{<}^{(q)} \cdot d^{(q),T}}{1 - d^{(q),T}} \cdot \left[\frac{1 - \left(d^{(q),S}d^{(q),T}\right)^n}{1 - d^{(q),S}d^{(q),T}} - \frac{\left(d^{(q),S}\right)^n - \left(d^{(q),T}\right)^n}{d^{(q),S} - d^{(q),T}} \right] = -\bar{\vartheta}_n^{(q)}.
 \end{aligned}$$

Part (b) has already been mentioned in Remark 1 (b) and is due to the fact that for $0 < q = \beta_\lambda$, the sequence $\left(a_n^{(q)}\right)_{n \in \mathbb{N}}$ is itself explicitly representable by $a_n^{(q)} = 0$ for all $n \in \mathbb{N}$ (cf. Properties 1 (P2)). Plugging this into (42) gives the desired result.

(c) Let us now consider $\max\{0, \beta_\lambda\} < q < \min\{1, e^{\beta_\lambda - 1}\}$. For a closed-form lower bound for $\tilde{B}_{\lambda, X_0, n}^{(p, q)}$ we have to employ the parameters $c \hat{=} c^{(q),T}$, $d \hat{=} d^{(q),T}$, $K_2 = \nu = 0$, $K_1 = \Gamma_{>}^{(q)}$ and $\varkappa = \left(d^{(q),S}\right)^2$, cf. (93) in combination with (85). The representations of $\zeta_n^{(q)}$ and $\vartheta_n^{(q)}$ in (A51) and (A52) follow immediately from (A56). For a closed-form upper bound, we use the parameters $c \hat{=} c^{(q),S}$, $d \hat{=} d^{(q),S}$, $-K_1 = K_2 = \Gamma_{>}^{(q)}$, $\varkappa = d^{(q),S}$ and $\nu = d^{(q),S}d^{(q),T}$. Notice that in this case we stick to the representation (A57). The formula (104) is obviously valid, and (105) is implied by

$$\begin{aligned}
 & \left(\frac{-\Gamma_{>}^{(q)}}{d^{(q),S} \left(1 - d^{(q),S}\right)} + \frac{\Gamma_{>}^{(q)}}{d^{(q),S} - d^{(q),S}d^{(q),T}} \right) \cdot \frac{d^{(q),S} \cdot \left(1 - \left(d^{(q),S}\right)^n\right)}{1 - d^{(q),S}} \\
 & = -\Gamma_{>}^{(q)} \cdot \frac{d^{(q),S} - d^{(q),T}}{\left(1 - d^{(q),S}\right)^2 \left(1 - d^{(q),T}\right)} \cdot \left(1 - \left(d^{(q),S}\right)^n\right).
 \end{aligned}$$

The parts (d) and (e) are trivial by incorporating that in all respective cases one has $d^{(q),S} \in]0, 1[$, $d^{(q),T} \in]0, 1[$ and $\lim_{n \rightarrow \infty} n \cdot d^{(q),S} = 0$. \square

Proof of Theorem 5. (a) For $\lambda \in]0, 1[$, we get $0 < q_\lambda^E < \beta_\lambda$ and the assertion follows by applying part (a) of Lemma A4. Notice that in the current subcase $\mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$ there holds $\frac{p_\lambda^E}{q_\lambda^E} \beta_\lambda - \alpha_\lambda = 0$ as well as $\frac{p_\lambda^E}{q_\lambda^E} = \frac{\alpha_A}{\beta_A} = \frac{\alpha_H}{\beta_H}$. For the case $\lambda \in \mathbb{R} \setminus]0, 1[$, one gets from Lemma A1 that $\max\{0, \beta_\lambda\} < q_\lambda^E$, and there holds $q_\lambda^E < \min\{1, e^{\beta_\lambda - 1}\}$ iff $\lambda \in]\lambda_-, \lambda_+[\setminus]0, 1[$, cf. Lemma 1. Thus, an application of part (c) of Lemma A4 proves the desired result. The assertion (b) is equivalent to part (d) of Lemma A4. \square

Proof of Theorem 6. The assertions follow immediately from (A45), Lemma A4(b),(e), Proposition 6(d) as well as the incorporation of the fact that for $\lambda \in]0, 1[$ there holds $q_\lambda^L = \beta_A^\lambda \beta_H^{1-\lambda} < \beta_\lambda$ in the case $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in (\mathcal{P}_{SP} \setminus (\mathcal{P}_{SP,1} \cup \mathcal{P}_{SP,4}))$ (i.e., $\beta_A \neq \beta_H$) respectively $q_\lambda^L = \beta_\lambda$ in the case $(\beta_A, \beta_H, \alpha_A, \alpha_H) \in \mathcal{P}_{SP,4}$ (i.e., $\beta_A = \beta_H$). \square

Proof of Theorem 7. This can be deduced from (A46), from the parts (b), (c) and (e) of Lemma A4 as well as the incorporation of $p_\lambda^U \geq \alpha_A^\lambda \alpha_H^{1-\lambda} > 0$ for $\lambda \in]0, 1[$. Notice that an inadequate choice of p_λ^U, q_λ^U may lead to $\frac{p_\lambda^U}{q_\lambda^U}(\beta_\lambda + x_0^{(q_\lambda^U)}) - \alpha_\lambda > 0$. \square

Proof of Theorem 8. The assertions follow immediately from (A45) and from the parts (b), (c) and (e) of Lemma A4. Notice that an inadequate choice of p_λ^L, q_λ^L may lead to $\frac{p_\lambda^L}{q_\lambda^L}(\beta_\lambda + x_0^{(q_\lambda^L)}) - \alpha_\lambda < 0$. \square

Proof of Theorem 9. Let $p_\lambda^U = \alpha_A^\lambda \alpha_H^{1-\lambda} > \max\{0, \alpha_\lambda\}$ and $q_\lambda^U = \beta_A^\lambda \beta_H^{1-\lambda} > \max\{0, \beta_\lambda\}$. Since $q_\lambda^U < \min\{1, e^{\beta_\lambda - 1}\}$ iff $\lambda \in]\lambda_-, \lambda_+[\setminus]0, 1[$ (cf. Lemma 1 for $q_\lambda := q_\lambda^U$), this theorem follows from (A46) of Lemma A4, from the parts (b), (e) of Lemma A4 and from part (d) of Proposition 14. \square

Appendix A.4. Proofs and Auxiliary Lemmas for Section 7

Proof of Theorem 10. As already mentioned above, one can adapt the proof of Theorem 9.1.3 in Ethier & Kurtz [138] who deal with drift-parameters $\eta = 0, \kappa_\bullet = 0$, and the different setup of σ -independent time-scale and a sequence of critical Galton-Watson processes without immigration with general offspring distribution. For the sake of brevity, we basically outline here only the main differences to their proof; for similar limit investigations involving offspring/immigration distributions and parametrizations which are incompatible to ours, see e.g., Sriram [142].

As a first step, let us define the generator

$$A_\bullet f(x) := (\eta - \kappa_\bullet \cdot x) \cdot f'(x) + \frac{\sigma^2}{2} \cdot x \cdot f''(x), \quad f \in C_c^\infty([0, \infty)),$$

which corresponds to the diffusion process \tilde{X} governed by (133). In connection with (130), we study

$$T_\bullet^{(m)} f(x) := EP_\bullet \left[f \left(\frac{1}{m} \left(\sum_{k=1}^{mx} Y_{0,k}^{(m)} + \tilde{Y}_0^{(m)} \right) \right) \right], \quad x \in E^{(m)} := \frac{1}{m} \mathbb{N}_0, \quad f \in C_c^\infty([0, \infty),$$

where the $Y_{0,k}^{(m)}, \tilde{Y}_0^{(m)}$ are independent and (Poisson- $\beta_\bullet^{(m)}$) respectively Poisson- $\alpha_\bullet^{(m)}$) distributed as the members of the collection $Y^{(m)}$ respectively $\tilde{Y}^{(m)}$. By the Theorems 8.2.1 and 1.6.5 as well as Corollary 4.8.9 of [138] it is sufficient to show

$$\lim_{m \rightarrow \infty} \sup_{x \in E^{(m)}} \left| \sigma^2 m \left(T_\bullet^{(m)} f(x) - f(x) \right) - A_\bullet f(x) \right| = 0, \quad f \in C_c^\infty([0, \infty)). \tag{A58}$$

But (A58) follows mainly from the next

Lemma A5. Let

$$S_n^{(m)} := \frac{1}{\sqrt{n}} \left(\sum_{k=1}^n (Y_{0,k}^{(m)} - \beta_{\bullet}^{(m)}) + \tilde{Y}_0^{(m)} - \alpha_{\bullet}^{(m)} \right), \quad n \in \mathbb{N}, m \in \bar{\mathbb{N}},$$

with the usual convention $S_0^{(m)} := 0$. Then for all $m \in \bar{\mathbb{N}}, x \in E^{(m)}$ and all $f \in C_c^\infty([0, \infty))$

$$\begin{aligned} \epsilon^{(m)}(x) &:= EP_{\bullet} \left[\int_0^1 (S_{mx}^{(m)})^2 x(1-v) \left(f'' \left(\beta_{\bullet}^{(m)} x + \frac{\alpha_{\bullet}^{(m)}}{m} + v \sqrt{\frac{x}{m}} S_{mx}^{(m)} \right) - f''(x) \right) dv \right] \\ &= \frac{1}{\sigma^2} \cdot \left[\sigma^2 m \cdot (T_{\bullet}^{(m)} f(x) - f(x)) - A_{\bullet} f(x) \right] + R^{(m)}, \quad \text{where } \lim_{m \rightarrow \infty} R^{(m)} = 0. \end{aligned} \tag{A59}$$

Proof of Lemma A5. Let us fix $f \in C_c^\infty([0, \infty))$. From the involved Poissonian expectations it is easy to see that

$$\lim_{m \rightarrow \infty} \left| \sigma^2 m (T_{\bullet}^{(m)} f(0) - f(0)) - A_{\bullet} f(0) \right| = 0,$$

and thus (A59) holds for $x = 0$. Accordingly, we next consider the case $x \in E^{(m)} \setminus \{0\}$, with fixed $m \in \bar{\mathbb{N}}$.

From $EP_{\bullet} \left[(S_{mx}^{(m)})^2 \right] = \beta_{\bullet}^{(m)} + \frac{\alpha_{\bullet}^{(m)}}{mx}$ we obtain

$$EP_{\bullet} \left[(S_{mx}^{(m)})^2 x f''(x) \int_0^1 (1-v) dv \right] = \frac{1}{2} \left(\beta_{\bullet}^{(m)} \cdot x + \frac{\alpha_{\bullet}^{(m)}}{m} \right) f''(x) =: a_{mx} \frac{f''(x)}{2} =: a \frac{f''(x)}{2}. \tag{A60}$$

Furthermore, with $b_{mx} := b := a + \sqrt{x/m} \cdot S_{mx}^{(m)} = \frac{1}{m} (\sum_{k=1}^{mx} Y_{0,k}^{(m)} + \tilde{Y}_0^{(m)})$ we get on $\{S_{mx}^{(m)} \neq 0\}$

$$\int_0^1 f'' \left(\beta_{\bullet}^{(m)} x + \frac{\alpha_{\bullet}^{(m)}}{m} + v \sqrt{\frac{x}{m}} S_{mx}^{(m)} \right) dv = \sqrt{\frac{m}{x}} \cdot \frac{1}{S_{mx}^{(m)}} \int_a^b f''(y) dy = \sqrt{\frac{m}{x}} \cdot \frac{f'(b) - f'(a)}{S_{mx}^{(m)}} \tag{A61}$$

as well as

$$\begin{aligned} \int_0^1 v f'' \left(\beta_{\bullet}^{(m)} x + \frac{\alpha_{\bullet}^{(m)}}{m} + v \sqrt{\frac{x}{m}} S_{mx}^{(m)} \right) dv &= \frac{m}{x (S_{mx}^{(m)})^2} \left[\int_a^b y f''(y) dy - a \int_a^b f''(y) dy \right] \\ &= \sqrt{\frac{m}{x}} \cdot \frac{f'(b)}{S_{mx}^{(m)}} + \frac{m}{x} \cdot \frac{f(a) - f(b)}{(S_{mx}^{(m)})^2}. \end{aligned} \tag{A62}$$

With our choice $\beta_{\bullet}^{(m)} = 1 - \frac{\kappa_{\bullet}}{\sigma^2 m}$ and $\alpha_{\bullet}^{(m)} = \beta_{\bullet}^{(m)} \cdot \frac{\eta}{\sigma^2}$, a Taylor expansion of f at x gives

$$f(a) = f(x) + \frac{1}{\sigma^2 m} \cdot f'(x) (\beta_{\bullet}^{(m)} \cdot \eta - \kappa_{\bullet} \cdot x) + o\left(\frac{1}{m}\right), \tag{A63}$$

where for the case $\eta = \kappa = 0$ we use the convention $o\left(\frac{1}{m}\right) \equiv 0$. Combining (A60) to (A63) and the centering $EP_{\bullet}\left[S_{mx}^{(m)}\right] = 0$, the left hand side of Equation (A59) becomes

$$\begin{aligned} & EP_{\bullet}\left[\int_0^1\left(S_{mx}^{(m)}\right)^2x(1-v)\left(f''\left(\beta_{\bullet}^{(m)}x+\frac{\alpha_{\bullet}^{(m)}}{m}+v\sqrt{\frac{x}{m}}S_{mx}^{(m)}\right)-f''(x)\right)dv\right] \\ &= EP_{\bullet}\left[\sqrt{mx}\cdot S_{mx}^{(m)}\cdot\left(f'(b)-f'(a)\right)\right]-EP_{\bullet}\left[\sqrt{mx}\cdot S_{mx}^{(m)}\cdot f'(b)+m\cdot\left(f(a)-f(b)\right)\right] \\ &\quad -\frac{1}{2}\left(\beta_{\bullet}^{(m)}\cdot x+\frac{\alpha_{\bullet}^{(m)}}{m}\right)\cdot f''(x) \\ &= m\cdot\left(EP_{\bullet}\left[f(b)\right]-f(a)\right)-\frac{1}{2}\left(\beta_{\bullet}^{(m)}\cdot x+\frac{\alpha_{\bullet}^{(m)}}{m}\right)\cdot f''(x) \\ &= m\cdot\left\{EP_{\bullet}\left[f\left(\frac{1}{m}\left(\sum_{k=1}^{mx}Y_{0,k}^{(m)}+\tilde{Y}_0\right)\right)\right]-f(x)\right\}-\frac{1}{\sigma^2}A_{\bullet}f(x) \\ &\quad +\frac{1}{\sigma^2}\left[\left(\eta-\kappa_{\bullet}\cdot x\right)-\beta_{\bullet}^{(m)}\cdot\eta+\kappa_{\bullet}\cdot x\right]\cdot f'(x)+\frac{x}{2}\left[1-\beta_{\bullet}^{(m)}-\frac{\alpha_{\bullet}^{(m)}}{m}\right]\cdot f''(x)-m\cdot o\left(\frac{1}{m}\right) \end{aligned}$$

which immediately leads to the right hand side of (A59). \square

To proceed with the proof of Theorem 10, we obtain for $m \geq 2\kappa_{\bullet}/\sigma^2$ the inequality $\beta_{\bullet}^{(m)} \geq 1/2$ and accordingly for all $v \in]0, 1[$, $x \in E^{(m)}$

$$\beta_{\bullet}^{(m)}x+\frac{\alpha_{\bullet}^{(m)}}{m}+v\sqrt{\frac{x}{m}}S_{mx}^{(m)}=(1-v)\cdot x\cdot\beta_{\bullet}^{(m)}+(1-v)\frac{\alpha_{\bullet}^{(m)}}{m}+v\left(\sum_{k=1}^{mx}Y_{0,k}^{(m)}+\tilde{Y}_0\right)\geq x\cdot\frac{1-v}{2}.$$

Suppose that the support of f is contained in the interval $[0, c]$. Correspondingly, for $v \leq 1 - 2c/x$ the integrand in $\epsilon^{(m)}(x)$ is zero and hence with (A64) we obtain the bounds

$$\begin{aligned} & \left|\int_0^1\left(S_{mx}^{(m)}\right)^2x(1-v)\left(f''\left(\beta_{\bullet}^{(m)}x+\frac{\alpha_{\bullet}^{(m)}}{m}+v\sqrt{\frac{x}{m}}S_{mx}^{(m)}\right)-f''(x)\right)dv\right| \\ & \leq \int_{0v(1-2c/x)}^1\left(S_{mx}^{(m)}\right)^2x(1-v)\cdot 2\|f''\|_{\infty}dv \leq x\cdot\left(S_{mx}^{(m)}\right)^2\left(1\wedge\frac{2c}{x}\right)^2\|f''\|_{\infty}. \end{aligned}$$

From this, one can deduce $\lim_{m \rightarrow \infty} \sup_{x \in E^{(m)}} \epsilon^{(m)}(x) = 0$ —and thus (A58)—in the same manner as at the end of the proof of Theorem 9.1.3 in [138] (by means of the dominated convergence theorem). \square

Proof of Proposition 15. Let $(\kappa_A, \kappa_H, \eta) \in \tilde{\mathcal{P}}_{NI} \cup \tilde{\mathcal{P}}_{SP,1}$ be fixed. We have to find those orders $\lambda \in \mathbb{R} \setminus]0, 1[$ which satisfy for all sufficiently large $m \in \mathbb{N}$

$$q_{\lambda}^{(m)} = \left(1 - \frac{\kappa_A}{\sigma^2 m}\right)^{\lambda} \left(1 - \frac{\kappa_H}{\sigma^2 m}\right)^{1-\lambda} < \min\left\{1, e^{\beta_{\lambda}^{(m)}-1}\right\}. \tag{A64}$$

In order to achieve this, we interpret $q_{\lambda}^{(m)} = q_{\lambda}\left(\frac{1}{m}\right)$ in terms of the function

$$q_{\lambda}(x) := \left(1 - \frac{\kappa_A}{\sigma^2} \cdot x\right)^{\lambda} \left(1 - \frac{\kappa_H}{\sigma^2} \cdot x\right)^{1-\lambda}, \quad x \in]-\epsilon, \epsilon[, \tag{A65}$$

for some small enough $\epsilon > 0$ such that (A65) is well-defined. Since $\beta_\lambda^{(m)} - 1 = -\frac{\kappa_\lambda}{\sigma^2 \cdot m} = -\frac{\kappa_\lambda}{\sigma^2} \cdot x = -\frac{\lambda \kappa_A + (1-\lambda)\kappa_H}{\sigma^2} \cdot x$, for the verification of (A64) it suffices to show

$$\lim_{x \searrow 0} \frac{1 - q_\lambda(x)}{x} > 0, \tag{A66}$$

$$\text{and } \lim_{x \searrow 0} \frac{e^{-\frac{\kappa_\lambda}{\sigma^2} \cdot x} - q_\lambda(x)}{x^2} > 0. \tag{A67}$$

By l'Hospital's rule, one gets $\lim_{x \searrow 0} \frac{1 - q_\lambda(x)}{x} = \frac{\lambda \kappa_A + (1-\lambda)\kappa_H}{\sigma^2} = \frac{\kappa_\lambda}{\sigma^2}$ and hence

$$(A66) \iff \begin{cases} \lambda < \frac{\kappa_H}{\kappa_H - \kappa_A}, & \text{if } \kappa_A < \kappa_H, \\ \lambda > -\frac{\kappa_H}{\kappa_A - \kappa_H}, & \text{if } \kappa_A > \kappa_H. \end{cases} \tag{A68}$$

To find a condition that guarantees (A67), we use l'Hospital's rule twice to deduce

$$\lim_{x \searrow 0} \frac{e^{-\frac{\kappa_\lambda}{\sigma^2} \cdot x} - q_\lambda(x)}{x^2} = \frac{1}{2\sigma^4} [\kappa_\lambda^2 - \lambda(\lambda - 1)(\kappa_A - \kappa_H)^2] = \frac{1}{2\sigma^4} [\lambda \kappa_A^2 + (1 - \lambda)\kappa_H^2]$$

and hence we obtain

$$(A67) \iff \begin{cases} \lambda < \frac{\kappa_H^2}{\kappa_H^2 - \kappa_A^2}, & \text{if } \kappa_A < \kappa_H, \\ \lambda > -\frac{\kappa_H^2}{\kappa_A^2 - \kappa_H^2}, & \text{if } \kappa_A > \kappa_H. \end{cases} \tag{A69}$$

To compare both the lower and upper bounds in (A68) and (A69), let us calculate

$$\frac{\kappa_H^2}{\kappa_H^2 - \kappa_A^2} - \frac{\kappa_H}{\kappa_H - \kappa_A} = -\frac{\kappa_A \kappa_H}{(\kappa_H - \kappa_A)(\kappa_H + \kappa_A)} \begin{cases} < 0, & \text{if } \kappa_A < \kappa_H, \\ > 0, & \text{if } \kappa_A > \kappa_H. \end{cases} \tag{A70}$$

Incorporating this, we observe that both conditions (A66) and (A67) are satisfied simultaneously iff

$$\begin{aligned} \lambda &< \min \left\{ \frac{\kappa_H}{\kappa_H - \kappa_A}, \frac{\kappa_H^2}{\kappa_H^2 - \kappa_A^2} \right\} &= \frac{\kappa_H^2}{\kappa_H^2 - \kappa_A^2} && \text{if } \kappa_A < \kappa_H, \\ \lambda &> \max \left\{ -\frac{\kappa_H}{\kappa_A - \kappa_H}, -\frac{\kappa_H^2}{\kappa_A^2 - \kappa_H^2} \right\} &= -\frac{\kappa_H^2}{\kappa_A^2 - \kappa_H^2} && \text{if } \kappa_A > \kappa_H, \end{aligned}$$

which finishes the proof. \square

The following lemma is the main tool for the proof of Theorem 11.

Lemma A6. Let $(\kappa_A, \kappa_H, \eta, \lambda) \in (\tilde{\mathcal{P}}_{NI} \cup \tilde{\mathcal{P}}_{SP,1}) \times (]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus\{0,1\})$. By using the quantities $\kappa_\lambda := \lambda\kappa_A + (1-\lambda)\kappa_H$ and $\Lambda_\lambda := \sqrt{\lambda\kappa_A^2 + (1-\lambda)\kappa_H^2}$ from (150) (which is well-defined, cf. (138)), one gets for all $t > 0$

- (a) $\lim_{m \rightarrow \infty} m \cdot (1 - q_\lambda^{(m)}) = \lim_{m \rightarrow \infty} m \cdot (1 - \beta_\lambda^{(m)}) = \frac{\kappa_\lambda}{\sigma^2}.$
- (b) $\lim_{m \rightarrow \infty} m^2 \cdot a_1^{(m)} = \lim_{m \rightarrow \infty} m^2 \cdot (q_\lambda^{(m)} - \beta_\lambda^{(m)}) = -\frac{\lambda(1-\lambda)(\kappa_A - \kappa_H)^2}{2\sigma^4} = -\frac{\Lambda_\lambda^2 - \kappa_\lambda^2}{2\sigma^4}.$
- (c) $\lim_{m \rightarrow \infty} m \cdot x_0^{(m)} = -\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \begin{cases} < 0, & \text{if } \lambda \in]0,1[, \\ > 0, & \text{if } \lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus[0,1]. \end{cases}$
- (d) $\lim_{m \rightarrow \infty} m^2 \cdot \Gamma_{<}^{(m)} = \lim_{m \rightarrow \infty} m^2 \cdot \Gamma_{>}^{(m)} = \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2\sigma^4} > 0.$
- (e) $\lim_{m \rightarrow \infty} m \cdot (1 - d^{(m),S}) = \frac{\Lambda_\lambda + \kappa_\lambda}{2\sigma^2} > 0.$
- (f) $\lim_{m \rightarrow \infty} m \cdot (1 - d^{(m),T}) = \frac{\Lambda_\lambda}{\sigma^2} > 0.$
- (g) $\lim_{m \rightarrow \infty} m \cdot (1 - d^{(m),S}d^{(m),T}) = \frac{3\Lambda_\lambda + \kappa_\lambda}{2\sigma^2} > 0.$
- (h) $\lim_{m \rightarrow \infty} (d^{(m),S})^{\sigma^2 mt} = \exp\left\{-\frac{\Lambda_\lambda + \kappa_\lambda}{2} \cdot t\right\} < 1.$
- (i) $\lim_{m \rightarrow \infty} (d^{(m),T})^{\sigma^2 mt} = \exp\{-\Lambda_\lambda \cdot t\} < 1.$
- (j) $\lim_{m \rightarrow \infty} (d^{(m),S}d^{(m),T})^{\sigma^2 mt} = \exp\left\{-\frac{3\Lambda_\lambda + \kappa_\lambda}{2} \cdot t\right\} < 1.$
- (k) for $\lambda \in]0,1[$, there holds for the respective quantities defined in (142) to (145)
 - $\lim_{m \rightarrow \infty} m \cdot \underline{\zeta}_{[\sigma^2 mt]}^{(m)} = \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2\sigma^2 \cdot \Lambda_\lambda} \cdot e^{-\Lambda_\lambda \cdot t} \cdot (1 - e^{-\Lambda_\lambda \cdot t}) > 0,$
 - $\lim_{m \rightarrow \infty} \underline{\vartheta}_{[\sigma^2 mt]}^{(m)} = \frac{1}{4} \cdot \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda}\right)^2 \cdot (1 - e^{-\Lambda_\lambda \cdot t})^2 > 0,$
 - $\lim_{m \rightarrow \infty} m \cdot \bar{\zeta}_{[\sigma^2 mt]}^{(m)} = \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\sigma^2} \cdot \left[\frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda - \kappa_\lambda} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} (1 - e^{-\Lambda_\lambda \cdot t})}{2 \cdot \Lambda_\lambda}\right] > 0,$
 - $\lim_{m \rightarrow \infty} \bar{\vartheta}_{[\sigma^2 mt]}^{(m)} = \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\Lambda_\lambda} \cdot \left[\frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda) \cdot t}}{3\Lambda_\lambda + \kappa_\lambda} + \frac{e^{-\Lambda_\lambda \cdot t} - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda - \kappa_\lambda}\right] > 0.$
- (l) for $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus[0,1]$, there holds for the respective quantities defined in (146) to (149)
 - $\lim_{m \rightarrow \infty} m \cdot \underline{\zeta}_{[\sigma^2 mt]}^{(m)} = \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2\sigma^2 \cdot \kappa_\lambda} \cdot e^{-\Lambda_\lambda \cdot t} \cdot (1 - e^{-\kappa_\lambda \cdot t}) > 0,$
 - $\lim_{m \rightarrow \infty} \underline{\vartheta}_{[\sigma^2 mt]}^{(m)} = \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2 \cdot \kappa_\lambda} \cdot \left[\frac{1 - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda} - \frac{1 - e^{-(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda + \kappa_\lambda}\right] > 0,$
 - $\lim_{m \rightarrow \infty} m \cdot \bar{\zeta}_{[\sigma^2 mt]}^{(m)} = \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2 \cdot \sigma^2} \cdot e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \cdot \left[t - \frac{1 - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda}\right] > 0,$
 - $\lim_{m \rightarrow \infty} \bar{\vartheta}_{[\sigma^2 mt]}^{(m)} = (\Lambda_\lambda - \kappa_\lambda)^2 \cdot \left[\frac{(\Lambda_\lambda - \kappa_\lambda) \left(1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t}\right)}{\Lambda_\lambda \cdot (\Lambda_\lambda + \kappa_\lambda)^2} + \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda \cdot (3\Lambda_\lambda + \kappa_\lambda)} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda + \kappa_\lambda} \cdot t\right] > 0.$

Proof of Lemma A6. For each of the assertions (a) to (l), we will make use of l’Hospital’s rule. To begin with, we obtain for arbitrary $\mu, \nu \in \mathbb{R}$

$$\begin{aligned} \lim_{m \rightarrow \infty} m \left[1 - (\beta_A^{(m)})^\mu (\beta_H^{(m)})^\nu\right] &= \lim_{m \rightarrow \infty} m^2 \left[\mu \cdot (\beta_A^{(m)})^{\mu-1} (\beta_H^{(m)})^\nu \frac{\kappa_A}{\sigma^2 m^2} + \nu \cdot (\beta_A^{(m)})^\mu (\beta_H^{(m)})^{\nu-1} \frac{\kappa_H}{\sigma^2 m^2}\right] \\ &= \mu \frac{\kappa_A}{\sigma^2} + \nu \frac{\kappa_H}{\sigma^2}. \end{aligned} \tag{A71}$$

From this, the first part of (a) follows immediately and the second part is a direct consequence of the definition of $\beta_\lambda^{(m)}$. Part (b) can be deduced from (A71):

$$\begin{aligned} \lim_{m \rightarrow \infty} m^2 \cdot a_1^{(m)} &= \lim_{m \rightarrow \infty} \frac{m}{2\sigma^2} \cdot \left[\lambda \cdot \kappa_{\mathcal{A}} \left(1 - (\beta_{\mathcal{A}}^{(m)})^{\lambda-1} (\beta_{\mathcal{H}}^{(m)})^{1-\lambda} \right) \right. \\ &\quad \left. + (1-\lambda) \cdot \kappa_{\mathcal{H}} \left(1 - (\beta_{\mathcal{A}}^{(m)})^\lambda (\beta_{\mathcal{H}}^{(m)})^{-\lambda} \right) \right] = -\frac{\lambda(1-\lambda)(\kappa_{\mathcal{A}} - \kappa_{\mathcal{H}})^2}{2\sigma^4} = -\frac{\Lambda_\lambda^2 - \kappa_\lambda^2}{2\sigma^4}. \end{aligned}$$

For the proof of (c), we rely on the inequalities $x_0^{(m)} \leq x_0^{(m)} \leq \bar{x}_0^{(m)}$ ($m \in \mathbb{N}$), where $x_0^{(m)}$ and $\bar{x}_0^{(m)}$ are the obvious notational adaptations of (124) and (126), respectively. Notice that $x_0^{(m)}$ and $\bar{x}_0^{(m)}$ are solutions of the (again adapted) quadratic equations $Q_\lambda^{(m)}(x) = x$ resp. $\bar{Q}_\lambda^{(m)}(x) = x$ (cf. (127) and (128)). These solutions clearly exist in the case $\lambda \in]0, 1[$. For sufficiently large approximations steps $m \in \mathbb{N}$, these solutions also exist in the case $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus]0, 1[$ since (138) together with parts (a) and (b) imply

$$\lim_{m \rightarrow \infty} \left(m \cdot (1 - q_\lambda^{(m)})^2 - 2 \cdot q_\lambda^{(m)} \cdot m^2 \cdot a_1^{(m)} \right) = \sigma^{-2} \cdot \left[\lambda \kappa_{\mathcal{A}}^2 + (1-\lambda) \kappa_{\mathcal{H}}^2 \right] > 0, \quad \text{for } \lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus]0, 1[.$$

To prove part (c), we show that the limits of $x_0^{(m)}$ and $\bar{x}_0^{(m)}$ coincide. Assume first that $\lambda \in]0, 1[$. Using (a) and (b), we obtain together with the obvious limit $\lim_{m \rightarrow \infty} q_\lambda^{(m)} = 1$

$$\begin{aligned} \lim_{m \rightarrow \infty} m \cdot \bar{x}_0^{(m)} &= \lim_{m \rightarrow \infty} \left(q_\lambda^{(m)} \right)^{-1} \cdot \left[m \cdot (1 - q_\lambda^{(m)}) - \sqrt{\left(m \cdot (1 - q_\lambda^{(m)}) \right)^2 - 2 \cdot q_\lambda^{(m)} \cdot m^2 \cdot a_1^{(m)}} \right] \\ &= \frac{\kappa_\lambda}{\sigma^2} - \sqrt{\left(\frac{\kappa_\lambda}{\sigma^2} \right)^2 + \frac{\Lambda_\lambda^2 - \kappa_\lambda^2}{\sigma^4}} = -\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2}. \end{aligned} \tag{A72}$$

Let $\underline{x}_0^{(m)}$ be the adapted version of the auxiliary fixed-point lower bound defined in (125). By incorporating $\lim_{m \rightarrow \infty} \beta_\lambda^{(m)} = 1$ we obtain with (a) and (b)

$$\lim_{m \rightarrow \infty} \underline{x}_0^{(m)} = \lim_{m \rightarrow \infty} \max \left\{ -\beta_\lambda^{(m)}, \frac{q_\lambda^{(m)} - \beta_\lambda^{(m)}}{1 - q_\lambda^{(m)}} \right\} = \lim_{m \rightarrow \infty} \frac{1}{m} \cdot \frac{m^2 \cdot a_1^{(m)}}{m \cdot (1 - q_\lambda^{(m)})} = 0,$$

which implies

$$\begin{aligned} \lim_{m \rightarrow \infty} m \cdot \underline{x}_0^{(m)} &= \lim_{m \rightarrow \infty} \frac{e^{-\underline{x}_0^{(m)}}}{q_\lambda^{(m)}} \cdot \left[m \cdot (1 - q_\lambda^{(m)}) - \sqrt{\left(m \cdot (1 - q_\lambda^{(m)}) \right)^2 - 2 \cdot e^{\underline{x}_0^{(m)}} q_\lambda^{(m)} \cdot m^2 \cdot a_1^{(m)}} \right] \\ &= \frac{\kappa_\lambda}{\sigma^2} - \sqrt{\left(\frac{\kappa_\lambda}{\sigma^2} \right)^2 + \frac{\Lambda_\lambda^2 - \kappa_\lambda^2}{\sigma^4}} = -\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2}. \end{aligned} \tag{A73}$$

Combining (A72) and (A73), the desired result (c) follows for $\lambda \in]0, 1[$. Assume now that $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus]0, 1[$. In this case the approximates $x_0^{(m)}$ and $\bar{x}_0^{(m)}$ have a different form, given in (124) and (126). However, the calculations work out in the same way: with parts (a) and (b) we get

$$\begin{aligned} \lim_{m \rightarrow \infty} m \cdot \underline{x}_0^{(m)} &= \lim_{m \rightarrow \infty} \frac{1}{q_\lambda^{(m)}} \cdot \left[m \cdot (1 - q_\lambda^{(m)}) - \sqrt{\left(m \cdot (1 - q_\lambda^{(m)}) \right)^2 - 2 \cdot q_\lambda^{(m)} \cdot m^2 \cdot a_1^{(m)}} \right] \\ &= \frac{\kappa_\lambda}{\sigma^2} - \sqrt{\left(\frac{\kappa_\lambda}{\sigma^2} \right)^2 + \frac{\Lambda_\lambda^2 - \kappa_\lambda^2}{\sigma^4}} = -\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2}, \end{aligned}$$

as well as

$$\begin{aligned} \lim_{m \rightarrow \infty} m \cdot \bar{x}_0^{(m)} &= \lim_{m \rightarrow \infty} m \cdot \left(1 - q_\lambda^{(m)}\right) - \sqrt{\left(m \cdot \left(1 - q_\lambda^{(m)}\right)\right)^2 - 2 \cdot m^2 \cdot a_1^{(m)}} \\ &= \frac{\kappa_\lambda}{\sigma^2} - \sqrt{\left(\frac{\kappa_\lambda}{\sigma^2}\right)^2 + \frac{\Lambda_\lambda^2 - \kappa_\lambda^2}{\sigma^4}} = -\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2}, \end{aligned}$$

which finally finishes the proof of part (c). Assertion (d) is a direct consequence of (c). Since the representations of the parameters $c^{(m),S}, d^{(m),S}, c^{(m),T}, d^{(m),T}$ are the same in both cases $\lambda \in]0, 1[$ and $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus]0, 1[$, the following considerations hold generally. Part (e) follows from (b) and (c) by

$$\lim_{m \rightarrow \infty} m \cdot (1 - d^{(m),S}) = \lim_{m \rightarrow \infty} \frac{m^2 \cdot a_1^{(m)}}{m \cdot x_0^{(m)}} = \frac{\Lambda_\lambda + \kappa_\lambda}{2\sigma^2} > 0.$$

Notice that this term is positive since on $]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus]0, 1[$ there holds $\kappa_\lambda > 0$ as well as $\Lambda_\lambda > 0$, cf. (A70). To prove (f), we apply the general limit $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1$ and get with (a), (c)

$$\lim_{m \rightarrow \infty} m \cdot (1 - d^{(m),T}) = \lim_{m \rightarrow \infty} \left(m \cdot \left(1 - q_\lambda^{(m)}\right) - q_\lambda^{(m)} \cdot m \cdot x_0^{(m)} \cdot \frac{e^{x_0^{(m)}} - 1}{x_0^{(m)}} \right) = \frac{\Lambda_\lambda}{\sigma^2}.$$

The limit (g) can be obtained from (e) and (f):

$$\lim_{m \rightarrow \infty} m \cdot (1 - d^{(m),S})d^{(m),T} = \lim_{m \rightarrow \infty} \left\{ m \cdot (1 - d^{(m),S}) + d^{(m),S} \cdot m \cdot (1 - d^{(m),T}) \right\} = \frac{3\Lambda_\lambda + \kappa_\lambda}{2\sigma^2}.$$

The assertions (h) resp. (i) resp. (j) follow from (e) resp. (f) resp. (g) by using the general relation $\lim_{m \rightarrow \infty} \left(1 + \frac{x_m}{m}\right)^m = \exp\{\lim_{m \rightarrow \infty} x_m\}$. To get the last two parts (k) and (l), we make repeatedly use of the results (a) to (j) and combine them with the formulas (142) to (149) of Corollary 14. More detailed, for $\lambda \in]0, 1[$ (and thus $q_\lambda^{(m)} < \beta_\lambda^{(m)}$) we obtain

$$\begin{aligned} m \cdot \bar{\zeta}_{[\sigma^2 mt]}^{(m)} &= m^2 \cdot \Gamma_{<}^{(m)} \cdot \frac{\left(d^{(m),T}\right)^{[\sigma^2 mt] - 1}}{m \cdot \left(1 - d^{(m),T}\right)} \cdot \left(1 - \left(d^{(m),T}\right)^{[\sigma^2 mt]}\right) \\ &\xrightarrow{m \rightarrow \infty} \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2\sigma^2 \cdot \Lambda_\lambda} \cdot e^{-\Lambda_\lambda \cdot t} \cdot \left(1 - e^{-\Lambda_\lambda \cdot t}\right) > 0, \\ \vartheta_{[\sigma^2 mt]}^{(m)} &= m^2 \cdot \Gamma_{<}^{(m)} \cdot \frac{1 - \left(d^{(m),T}\right)^{[\sigma^2 mt]}}{\left(m \cdot \left(1 - d^{(m),T}\right)\right)^2} \cdot \left[1 - \frac{d^{(m),T} \left(1 + \left(d^{(m),T}\right)^{[\sigma^2 mt]}\right)}{1 + d^{(m),T}}\right] \\ &\xrightarrow{m \rightarrow \infty} \frac{1}{4} \cdot \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda}\right)^2 \cdot \left(1 - e^{-\Lambda_\lambda \cdot t}\right)^2 > 0, \\ m \cdot \bar{\zeta}_{[\sigma^2 mt]}^{(m)} &= m^2 \cdot \Gamma_{<}^{(m)} \cdot \left[\frac{\left(d^{(m),S}\right)^{[\sigma^2 mt]} - \left(d^{(m),T}\right)^{[\sigma^2 mt]}}{m \cdot \left(1 - d^{(m),T}\right) - m \cdot \left(1 - d^{(m),S}\right)} \right. \\ &\quad \left. - \left(d^{(m),S}\right)^{[\sigma^2 mt] - 1} \cdot \frac{1 - \left(d^{(m),T}\right)^{[\sigma^2 mt]}}{m \cdot \left(1 - d^{(m),T}\right)} \right] \\ &\xrightarrow{m \rightarrow \infty} \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\sigma^2} \cdot \left[\frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda - \kappa_\lambda} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \left(1 - e^{-\Lambda_\lambda \cdot t}\right)}{2 \cdot \Lambda_\lambda} \right] > 0, \end{aligned}$$

$$\begin{aligned} \bar{\vartheta}_{[\sigma^2 mt]}^{(m)} &= \frac{m^2 \cdot \Gamma_{<}^{(m)} \cdot d^{(m),T}}{m \cdot (1 - d^{(m),T})} \cdot \left[\frac{1 - (d^{(m),S} d^{(m),T})^{[\sigma^2 mt]}}{m \cdot (1 - d^{(m),S} d^{(m),T})} - \frac{(d^{(m),S})^{[\sigma^2 mt]} - (d^{(m),T})^{[\sigma^2 mt]}}{m \cdot (1 - d^{(m),T}) - m \cdot (1 - d^{(m),S})} \right] \\ &\xrightarrow{m \rightarrow \infty} \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\Lambda_\lambda} \cdot \left[\frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda) \cdot t}}{3\Lambda_\lambda + \kappa_\lambda} + \frac{e^{-\Lambda_\lambda \cdot t} - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda - \kappa_\lambda} \right] > 0. \end{aligned}$$

For $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus [0, 1]$ (and thus $q_\lambda^{(m)} > \beta_\lambda^{(m)}$) we get

$$\begin{aligned} m \cdot \bar{\zeta}_{[\sigma^2 mt]}^{(m)} &= m^2 \cdot \Gamma_{>}^{(m)} \cdot \frac{(d^{(m),T})^{[\sigma^2 mt]} - (d^{(m),S})^{2 \cdot [\sigma^2 mt]}}{m \cdot (1 - d^{(m),S}) (1 + d^{(m),S}) - m \cdot (1 - d^{(m),T})} \\ &\xrightarrow{m \rightarrow \infty} \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2\sigma^2 \cdot \kappa_\lambda} \cdot e^{-\Lambda_\lambda \cdot t} \cdot (1 - e^{-\kappa_\lambda \cdot t}) > 0, \\ \bar{\vartheta}_{[\sigma^2 mt]}^{(m)} &= \frac{m^2 \cdot \Gamma_{>}^{(m)}}{m \cdot (1 - d^{(m),S}) (1 + d^{(m),S}) - m \cdot (1 - d^{(m),T})} \\ &\quad \cdot \left[\frac{d^{(m),T} \cdot (1 - (d^{(m),T})^{[\sigma^2 mt]})}{m \cdot (1 - d^{(m),T})} - \frac{(d^{(m),S})^2 \cdot (1 - (d^{(m),S})^{2 \cdot [\sigma^2 mt]})}{m \cdot (1 - d^{(m),S}) (1 + d^{(m),S})} \right] \\ &\xrightarrow{m \rightarrow \infty} \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2 \cdot \kappa_\lambda} \cdot \left[\frac{1 - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda} - \frac{1 - e^{-(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda + \kappa_\lambda} \right] > 0, \\ m \cdot \bar{\zeta}_{[\sigma^2 mt]}^{(m)} &= m^2 \cdot \Gamma_{>}^{(m)} \cdot (d^{(m),S})^{[\sigma^2 mt] - 1} \cdot \left[\frac{[\sigma^2 mt]}{m} - \frac{1 - (d^{(m),T})^{[\sigma^2 mt]}}{m \cdot (1 - d^{(m),T})} \right] \\ &\xrightarrow{m \rightarrow \infty} \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2 \cdot \sigma^2} \cdot e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \cdot \left[t - \frac{1 - e^{-\Lambda_\lambda \cdot t}}{\Lambda_\lambda} \right] > 0, \end{aligned}$$

$$\begin{aligned} \bar{\vartheta}_{[\sigma^2 mt]}^{(m)} &= m^2 \cdot \Gamma_{>}^{(m)} \cdot \left[\frac{m \cdot (1 - d^{(m),T}) - m \cdot (1 - d^{(m),S})}{m^2 \cdot (1 - d^{(m),S})^2 \cdot m \cdot (1 - d^{(m),T})} \cdot (1 - (d^{(m),S})^{[\sigma^2 mt]}) \right. \\ &\quad \left. + \frac{d^{(m),T} (1 - (d^{(m),S} d^{(m),T})^{[\sigma^2 mt]})}{m \cdot (1 - d^{(m),T}) \cdot m \cdot (1 - d^{(m),S} d^{(m),T})} - \frac{(d^{(m),S})^{[\sigma^2 mt]}}{m \cdot (1 - d^{(m),S})} \cdot \frac{[\sigma^2 mt]}{m} \right] \\ &\xrightarrow{m \rightarrow \infty} (\Lambda_\lambda - \kappa_\lambda)^2 \cdot \left[\frac{(\Lambda_\lambda - \kappa_\lambda) (1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t})}{\Lambda_\lambda \cdot (\Lambda_\lambda + \kappa_\lambda)^2} + \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda \cdot (3\Lambda_\lambda + \kappa_\lambda)} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t}}{\Lambda_\lambda + \kappa_\lambda} \cdot t \right] > 0. \quad \square \end{aligned}$$

Proof of Theorem 11. It suffices to compute the limits of the bounds given in Corollary 14 as m tends to infinity. This is done by applying Lemma A6 which provides corresponding limits of all quantities

of interest. Accordingly, for all $t > 0$ the lower bound (153) in the case $\lambda \in]0, 1[$ can be obtained from (140), (142) and (143) by

$$\begin{aligned} & \lim_{m \rightarrow \infty} \exp \left\{ x_0^{(m)} \cdot \left[X_0^{(m)} - \frac{\eta}{\sigma^2} \cdot \frac{d^{(m),T}}{1 - d^{(m),T}} \right] \left(1 - \left(d^{(m),T} \right)^{\lfloor \sigma^2 m t \rfloor} \right) \right. \\ & \quad \left. + x_0^{(m)} \frac{\eta}{\sigma^2} \cdot \lfloor \sigma^2 m t \rfloor + \underline{\zeta}_{\lfloor \sigma^2 m t \rfloor}^{(m)} \cdot X_0^{(m)} + \underline{\vartheta}_{\lfloor \sigma^2 m t \rfloor}^{(m)} \right\} \\ &= \lim_{m \rightarrow \infty} \exp \left\{ m \cdot x_0^{(m)} \cdot \left[\frac{X_0^{(m)}}{m} - \frac{\eta}{\sigma^2} \cdot \frac{d^{(m),T}}{m \cdot (1 - d^{(m),T})} \right] \left(1 - \left(d^{(m),T} \right)^{\lfloor \sigma^2 m t \rfloor} \right) \right. \\ & \quad \left. + m \cdot x_0^{(m)} \frac{\eta}{\sigma^2} \cdot \frac{\lfloor \sigma^2 m t \rfloor}{m} + m \cdot \underline{\zeta}_{\lfloor \sigma^2 m t \rfloor}^{(m)} \cdot \frac{X_0^{(m)}}{m} + \underline{\vartheta}_{\lfloor \sigma^2 m t \rfloor}^{(m)} \right\} \\ &= \exp \left\{ - \frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \cdot \left[\tilde{X}_0 - \frac{\eta}{\sigma^2} \cdot \frac{\sigma^2}{\Lambda_\lambda} \right] \left(1 - e^{-\Lambda_\lambda t} \right) - \frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \cdot \frac{\eta}{\sigma^2} \cdot \sigma^2 t \right. \\ & \quad \left. + \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{2\sigma^2 \cdot \Lambda_\lambda} \cdot e^{-\Lambda_\lambda t} \cdot \left(1 - e^{-\Lambda_\lambda t} \right) \cdot \tilde{X}_0 + \frac{\eta}{4\sigma^2} \cdot \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \right)^2 \cdot \left(1 - e^{-\Lambda_\lambda t} \right)^2 \right\} \\ &= \exp \left\{ - \frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \left[\tilde{X}_0 - \frac{\eta}{\Lambda_\lambda} \right] \left(1 - e^{-\Lambda_\lambda t} \right) - \frac{\eta}{\sigma^2} (\Lambda_\lambda - \kappa_\lambda) \cdot t + L_\lambda^{(1)}(t) \cdot \tilde{X}_0 + \frac{\eta}{\sigma^2} \cdot L_\lambda^{(2)}(t) \right\}. \end{aligned}$$

For all $t > 0$, the upper bound (154) in the case $\lambda \in]0, 1[$ follows analogously from (141), (144), (145) by

$$\begin{aligned} & \lim_{m \rightarrow \infty} \exp \left\{ x_0^{(m)} \cdot \left[X_0^{(m)} - \frac{\eta}{\sigma^2} \cdot \frac{d^{(m),S}}{1 - d^{(m),S}} \right] \left(1 - \left(d^{(m),S} \right)^{\lfloor \sigma^2 m t \rfloor} \right) \right. \\ & \quad \left. + x_0^{(m)} \frac{\eta}{\sigma^2} \cdot \lfloor \sigma^2 m t \rfloor - \bar{\zeta}_{\lfloor \sigma^2 m t \rfloor}^{(m)} \cdot X_0^{(m)} - \bar{\vartheta}_{\lfloor \sigma^2 m t \rfloor}^{(m)} \right\} \\ &= \lim_{m \rightarrow \infty} \exp \left\{ m \cdot x_0^{(m)} \cdot \left[\frac{X_0^{(m)}}{m} - \frac{\eta}{\sigma^2} \cdot \frac{d^{(m),S}}{m \cdot (1 - d^{(m),S})} \right] \left(1 - \left(d^{(m),S} \right)^{\lfloor \sigma^2 m t \rfloor} \right) \right. \\ & \quad \left. + m \cdot x_0^{(m)} \frac{\eta}{\sigma^2} \cdot \frac{\lfloor \sigma^2 m t \rfloor}{m} - m \cdot \bar{\zeta}_{\lfloor \sigma^2 m t \rfloor}^{(m)} \cdot \frac{X_0^{(m)}}{m} - \bar{\vartheta}_{\lfloor \sigma^2 m t \rfloor}^{(m)} \right\} \\ &= \exp \left\{ - \frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \left[\tilde{X}_0 - \frac{\eta}{\sigma^2} \cdot \frac{2\sigma^2}{\Lambda_\lambda + \kappa_\lambda} \right] \left(1 - \left(e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} \right) \right) - \frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \cdot \frac{\eta}{\sigma^2} \cdot \sigma^2 t \right. \\ & \quad \left. - \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\sigma^2} \cdot \left[\frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} - e^{-\Lambda_\lambda t}}{\Lambda_\lambda - \kappa_\lambda} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} (1 - e^{-\Lambda_\lambda t})}{2 \cdot \Lambda_\lambda} \right] \cdot \tilde{X}_0 \right. \\ & \quad \left. - \frac{\eta}{\sigma^2} \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\Lambda_\lambda} \cdot \left[\frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} + \frac{e^{-\Lambda_\lambda t} - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda - \kappa_\lambda} \right] \right\} \\ &= \exp \left\{ - \frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \left[\tilde{X}_0 - \frac{\eta}{\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)} \right] \left(1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} \right) - \frac{\eta}{\sigma^2} (\Lambda_\lambda - \kappa_\lambda) \cdot t \right. \\ & \quad \left. - U_\lambda^{(1)}(t) \cdot \tilde{X}_0 - \frac{\eta}{\sigma^2} \cdot U_\lambda^{(2)}(t) \right\}. \end{aligned}$$

In the case $\lambda \in]\tilde{\lambda}_-, \tilde{\lambda}_+[\setminus]0, 1[$, the lower bound as well as the upper bound of the Hellinger integral limit is obtained analogously, by taking into account that the quantities $\underline{\zeta}_n^{(m)}, \underline{\vartheta}_n^{(m)}, \bar{\zeta}_n^{(m)}, \bar{\vartheta}_n^{(m)}$ now have the form (146) to (149) instead of (142) to (145). Thus, the functions $L_\lambda^{(1)}(t), U_\lambda^{(1)}(t), L_\lambda^{(2)}(t), U_\lambda^{(2)}(t)$ are obtained by employing the limits of part (l) of Lemma A6 instead of part (k). □

The next Lemma (and parts of its proof) will be useful for the verification of Theorem 12:

Lemma A7. Recall the bounds on the Hellinger integral m -limit given in (153) and (154) of Theorem 11, in terms of $L_\lambda^{(i)}(t)$ and $U_\lambda^{(i)}(t)$ ($i = 1, 2$) defined by (155) to (158). Correspondingly, one gets the following λ -limits for all $t \in [0, \infty[$:

(a) for all $\kappa_A \in]0, \infty[$ and all $\kappa_H \in [0, \infty[$ with $\kappa_A \neq \kappa_H$

$$\lim_{\lambda \nearrow 1} \frac{\partial L_\lambda^{(1)}(t)}{\partial \lambda} = \lim_{\lambda \nearrow 1} \frac{\partial L_\lambda^{(2)}(t)}{\partial \lambda} = \lim_{\lambda \nearrow 1} \frac{\partial U_\lambda^{(1)}(t)}{\partial \lambda} = \lim_{\lambda \nearrow 1} \frac{\partial U_\lambda^{(2)}(t)}{\partial \lambda} = 0. \tag{A74}$$

(b) for $\kappa_A = 0$ and all $\kappa_H \in]0, \infty[$

$$\lim_{\lambda \nearrow 1} \frac{\partial L_\lambda^{(1)}(t)}{\partial \lambda} = -\frac{\kappa_H^2 \cdot t}{2\sigma^2}, \tag{A75}$$

$$\lim_{\lambda \nearrow 1} \frac{\partial L_\lambda^{(2)}(t)}{\partial \lambda} = -\frac{\kappa_H^2 \cdot t^2}{4}, \tag{A76}$$

$$\lim_{\lambda \nearrow 1} \frac{\partial U_\lambda^{(1)}(t)}{\partial \lambda} = \lim_{\lambda \nearrow 1} \frac{\partial U_\lambda^{(2)}(t)}{\partial \lambda} = 0. \tag{A77}$$

Proof of Lemma A7. For all $\kappa_A, \kappa_H \in [0, \infty[$ with $\kappa_A \neq \kappa_H$ one can deduce from (150) as well as (155) to (158) the following derivatives:

$$\begin{aligned} \frac{\partial L_\lambda^{(1)}(t)}{\partial \lambda} &= \frac{1}{2\sigma^2} \left\{ \frac{t}{2} \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \right)^2 (\kappa_A^2 - \kappa_H^2) [2e^{-2\Lambda_\lambda t} - e^{-\Lambda_\lambda t}] \right. \\ &\quad \left. + e^{-\Lambda_\lambda t} \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} \left[\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} (\kappa_A^2 - \kappa_H^2 - 2\Lambda_\lambda(\kappa_A - \kappa_H)) - \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \right)^2 \frac{\kappa_A^2 - \kappa_H^2}{2} \right] \right\}, \tag{A78} \end{aligned}$$

$$\begin{aligned} \frac{\partial L_\lambda^{(2)}(t)}{\partial \lambda} &= \frac{1}{4} \left\{ \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \cdot \left(\frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} \right)^2 \cdot (\kappa_A^2 - \kappa_H^2 - 2\Lambda_\lambda(\kappa_A - \kappa_H) - \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} (\kappa_A^2 - \kappa_H^2)) \right. \\ &\quad \left. + t \cdot e^{-\Lambda_\lambda t} \cdot \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \right)^2 \cdot \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} \cdot (\kappa_A^2 - \kappa_H^2) \right\}, \tag{A79} \end{aligned}$$

$$\begin{aligned} \frac{\partial U_\lambda^{(1)}(t)}{\partial \lambda} &= \frac{1}{\sigma^2} \left\{ \frac{\Lambda_\lambda - \kappa_\lambda}{2\Lambda_\lambda} \left[t e^{-\Lambda_\lambda t} (\kappa_A^2 - \kappa_H^2) - \frac{t}{2} e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} (\kappa_A^2 - \kappa_H^2 + 2\Lambda_\lambda(\kappa_A - \kappa_H)) \right] \right. \\ &\quad - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} - e^{-\Lambda_\lambda t}}{2\Lambda_\lambda} \cdot (\kappa_A^2 - \kappa_H^2 - 2\Lambda_\lambda(\kappa_A - \kappa_H)) \\ &\quad + \left(\frac{\Lambda_\lambda - \kappa_\lambda}{2\Lambda_\lambda} \right)^2 \left[\frac{t}{2} e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} (\kappa_A^2 - \kappa_H^2 + 2\Lambda_\lambda(\kappa_A - \kappa_H)) \right. \\ &\quad \left. - \frac{t}{2} e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t} (3(\kappa_A^2 - \kappa_H^2) + 2\Lambda_\lambda(\kappa_A - \kappa_H)) \right. \\ &\quad \left. + e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} \cdot \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} \cdot (\kappa_A^2 - \kappa_H^2) \right] \\ &\quad \left. + \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} (\kappa_A^2 - \kappa_H^2 - 2\Lambda_\lambda(\kappa_A - \kappa_H)) \left[\frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} - e^{-\Lambda_\lambda t}}{\Lambda_\lambda - \kappa_\lambda} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} (1 - e^{-\Lambda_\lambda t})}{2\Lambda_\lambda} \right] \right\}, \tag{A80} \end{aligned}$$

$$\begin{aligned}
 \frac{\partial U_\lambda^{(2)}(t)}{\partial \lambda} &= \frac{(\Lambda_\lambda - \kappa_\lambda)^2}{\Lambda_\lambda(3\Lambda_\lambda + \kappa_\lambda)} \left[\frac{t}{2} e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t} \left(3 \frac{\kappa_A^2 - \kappa_H^2}{2\Lambda_\lambda} + \kappa_A - \kappa_H \right) \right. \\
 &\quad \left. - \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} \cdot \left(3 \frac{\kappa_A^2 - \kappa_H^2}{2\Lambda_\lambda} + \kappa_A - \kappa_H \right) \right] \\
 &\quad + \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \left[\frac{t}{2} e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} \left(\frac{\kappa_A^2 - \kappa_H^2}{2\Lambda_\lambda} + \kappa_A - \kappa_H \right) - t e^{-\Lambda_\lambda t} \frac{\kappa_A^2 - \kappa_H^2}{2\Lambda_\lambda} \right] \\
 &\quad + \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} \left(\frac{\kappa_A^2 - \kappa_H^2}{2\Lambda_\lambda} - \kappa_A + \kappa_H \right) \\
 &\quad + \left[2 \left(\frac{\kappa_A^2 - \kappa_H^2}{2\Lambda_\lambda} - \kappa_A + \kappa_H \right) - \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda^2} \cdot \frac{\kappa_A^2 - \kappa_H^2}{2} \right] \\
 &\quad \cdot \frac{1}{\Lambda_\lambda} \left[\frac{\Lambda_\lambda - \kappa_\lambda}{3\Lambda_\lambda + \kappa_\lambda} \left(1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t} \right) - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} + e^{-\Lambda_\lambda t} \right]. \tag{A81}
 \end{aligned}$$

If $\kappa_A \in]0, \infty[$ and $\kappa_H \in [0, \infty[$ with $\kappa_A \neq \kappa_H$, then one gets $\lim_{\lambda \nearrow 1} \Lambda_\lambda = \lim_{\lambda \nearrow 1} \kappa_\lambda = \kappa_A > 0$ which implies (A74) from (A78) to (A81). For the proof of part (b), let us correspondingly assume $\kappa_A = 0$ and $\kappa_H \in]0, \infty[$, which by (150) leads to $\kappa_\lambda = \kappa_H \cdot (1 - \lambda)$, $\Lambda_\lambda = \kappa_H \cdot \sqrt{1 - \lambda}$ and the convergences $\lim_{\lambda \nearrow 1} \Lambda_\lambda = \lim_{\lambda \nearrow 1} \kappa_\lambda = 0$. From this, the assertions (A75), (A76), (A77) follow in a straightforward manner from (A78), (A79), (A80) – respectively – by using (parts of) the obvious relations

$$\lim_{\lambda \nearrow 1} \frac{\kappa_\lambda}{\Lambda_\lambda} = 0, \quad \lim_{\lambda \nearrow 1} \frac{\Lambda_\lambda \pm \kappa_\lambda}{\Lambda_\lambda} = \lim_{\lambda \nearrow 1} \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda + \kappa_\lambda} = 1, \tag{A82}$$

$$\lim_{\lambda \nearrow 1} \frac{1 - e^{-c_\lambda t}}{c_\lambda} = t \quad \text{for all } c_\lambda \in \left\{ \Lambda_\lambda, \frac{\Lambda_\lambda + \kappa_\lambda}{2}, \frac{3\Lambda_\lambda + \kappa_\lambda}{2} \right\}. \tag{A83}$$

In order to get the last assertion in (A77), we make use of the following limits

$$\lim_{\lambda \nearrow 1} \frac{1}{\Lambda_\lambda - \kappa_\lambda} - \frac{3}{3\Lambda_\lambda + \kappa_\lambda} = \lim_{\lambda \nearrow 1} \frac{4\kappa_H}{(\kappa_H - \kappa_H \cdot \sqrt{1 - \lambda}) \cdot (3\kappa_H + \kappa_H \cdot \sqrt{1 - \lambda})} = \frac{4}{3\kappa_H} \tag{A84}$$

and

$$\lim_{\lambda \nearrow 1} \frac{1}{\Lambda_\lambda} \left[\frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} - \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda - \kappa_\lambda} + \frac{1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda - \kappa_\lambda} \right] = 0. \tag{A85}$$

To see (A85), let us first observe that the involved limit can be rewritten as

$$\lim_{\lambda \nearrow 1} \left\{ \frac{1}{\Lambda_\lambda(\Lambda_\lambda - \kappa_\lambda)} \left[\frac{1}{3} - \frac{1}{3} e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t} + e^{-\Lambda_\lambda t} - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} \right] \right. \tag{A86}$$

$$\left. + \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda} \left[\frac{1}{3\Lambda_\lambda + \kappa_\lambda} - \frac{1}{3(\Lambda_\lambda - \kappa_\lambda)} \right] \right\}. \tag{A87}$$

Substituting $x := \sqrt{1 - \lambda}$ and applying l’Hôpital’s rule twice, we get for the first limit (A86)

$$\begin{aligned}
 &\lim_{x \searrow 0} \frac{\frac{1}{3} - \frac{1}{3} e^{-\frac{\kappa_H t}{2}(3x+x^2)} + e^{-\kappa_H t x} - e^{-\frac{\kappa_H t}{2}(x+x^2)}}{\kappa_H^2 \cdot (x^2 - x^3)} \\
 &= \lim_{x \searrow 0} \frac{\frac{\kappa_H t}{6}(3+2x) e^{-\frac{\kappa_H t}{2}(3x+x^2)} - \kappa_H t e^{-\kappa_H t x} + \frac{\kappa_H t}{2}(1+2x) e^{-\frac{\kappa_H t}{2}(x+x^2)}}{\kappa_H^2 \cdot (2x - 3x^2)} \\
 &= \lim_{x \searrow 0} \frac{\left[-\frac{\kappa_H^2 t^2}{12}(3+2x)^2 + \frac{\kappa_H t}{3} \right] e^{-\frac{\kappa_H t}{2}(3x+x^2)} + \kappa_H^2 t^2 e^{-\kappa_H t x} - \left[\frac{\kappa_H^2 t^2}{4}(1+2x)^2 - \kappa_H t \right] e^{-\frac{\kappa_H t}{2}(x+x^2)}}{\kappa_H^2 \cdot (2 - 6x)} \\
 &= \frac{1}{2\kappa_H^2} \left[-\frac{3\kappa_H^2 t^2}{4} + \frac{\kappa_H t}{3} + \kappa_H^2 t^2 - \frac{\kappa_H^2 t^2}{4} + \kappa_H t \right] = \frac{2t}{3\kappa_H}.
 \end{aligned}$$

The second limit (A87) becomes

$$\lim_{\lambda \nearrow 1} \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} \cdot \frac{3\Lambda_\lambda + \kappa_\lambda}{\Lambda_\lambda} \cdot \frac{-4\kappa_{\mathcal{H}}}{(3\kappa_{\mathcal{H}} + \sqrt{1 - \lambda\kappa_{\mathcal{H}}})(3\kappa_{\mathcal{H}} - 3\sqrt{1 - \lambda\kappa_{\mathcal{H}}})} \tag{A88}$$

and consequently (A85) follows. To proceed with the proof of (A77), we rearrange

$$\begin{aligned} \lim_{\lambda \nearrow 1} \frac{\partial U_\lambda^{(2)}(t)}{\partial \lambda} &= \lim_{\lambda \nearrow 1} \left\{ \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \right)^2 \left[\frac{\Lambda_\lambda}{3\Lambda_\lambda + \kappa_\lambda} \left(\frac{t}{2} e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t} \left(-\frac{3\kappa_{\mathcal{H}}^2}{2\Lambda_\lambda} - \kappa_{\mathcal{H}} \right) \right) \right. \right. \\ &\quad - \frac{\Lambda_\lambda}{3\Lambda_\lambda + \kappa_\lambda} \cdot \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} \left(-\frac{3\kappa_{\mathcal{H}}^2}{2\Lambda_\lambda} - \kappa_{\mathcal{H}} \right) + \frac{\Lambda_\lambda}{\Lambda_\lambda - \kappa_\lambda} \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} - e^{-\Lambda_\lambda t}}{\Lambda_\lambda - \kappa_\lambda} \left(-\frac{\kappa_{\mathcal{H}}^2}{2\Lambda_\lambda} + \kappa_{\mathcal{H}} \right) \\ &\quad \left. \left. - \frac{\Lambda_\lambda}{\Lambda_\lambda - \kappa_\lambda} \left(-\frac{t}{2} e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} \left(-\frac{\kappa_{\mathcal{H}}^2}{2\Lambda_\lambda} - \kappa_{\mathcal{H}} \right) - t e^{-\Lambda_\lambda t} \frac{\kappa_{\mathcal{H}}^2}{2\Lambda_\lambda} \right) \right] \right. \\ &\quad \left. + \left[\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \left(-\kappa_{\mathcal{H}}^2 + 2\Lambda_\lambda \kappa_{\mathcal{H}} \right) + \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \right)^2 \frac{\kappa_{\mathcal{H}}^2}{2} \right] \cdot \left[\frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda(3\Lambda_\lambda + \kappa_\lambda)} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} - e^{-\Lambda_\lambda t}}{\Lambda_\lambda(\Lambda_\lambda - \kappa_\lambda)} \right] \right\} \\ &= \lim_{\lambda \nearrow 1} \left\{ \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \right)^2 \left[\frac{\kappa_{\mathcal{H}}^2 t}{4} \left(-\frac{3e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda - \kappa_\lambda} + \frac{2e^{-\Lambda_\lambda t}}{\Lambda_\lambda - \kappa_\lambda} \right) \right. \right. \tag{A89} \end{aligned}$$

$$\begin{aligned} &\quad \left. + \frac{\kappa_{\mathcal{H}}^2}{2} \left(\frac{3 \left(1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t} \right)}{(3\Lambda_\lambda + \kappa_\lambda)^2} - \frac{1 - e^{-\Lambda_\lambda t}}{(\Lambda_\lambda - \kappa_\lambda)^2} + \frac{1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t}}{(\Lambda_\lambda - \kappa_\lambda)^2} \right) \right. \tag{A90} \\ &\quad \left. + \kappa_{\mathcal{H}} \left(-\frac{\Lambda_\lambda}{3\Lambda_\lambda + \kappa_\lambda} \cdot \frac{t e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{2} + \frac{\Lambda_\lambda}{3\Lambda_\lambda + \kappa_\lambda} \cdot \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} - \frac{\Lambda_\lambda}{\Lambda_\lambda - \kappa_\lambda} \cdot \frac{t e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t}}{2} \right. \right. \\ &\quad \left. \left. + \frac{\Lambda_\lambda}{\Lambda_\lambda - \kappa_\lambda} \cdot \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda - \kappa_\lambda} - \frac{\Lambda_\lambda}{\Lambda_\lambda - \kappa_\lambda} \cdot \frac{1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda - \kappa_\lambda} \right) \right] \\ &\quad \left. + \left[\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \left(-\kappa_{\mathcal{H}}^2 + 2\Lambda_\lambda \kappa_{\mathcal{H}} \right) + \left(\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \right)^2 \frac{\kappa_{\mathcal{H}}^2}{2} \right] \cdot \left[\frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda(3\Lambda_\lambda + \kappa_\lambda)} - \frac{e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t} - e^{-\Lambda_\lambda t}}{\Lambda_\lambda(\Lambda_\lambda - \kappa_\lambda)} \right] \right\}. \tag{A91} \end{aligned}$$

By means of (A82) to (A84), the limit of the expression after the squared brackets in (A89) becomes

$$\lim_{\lambda \nearrow 1} \left\{ \frac{\kappa_{\mathcal{H}}^2 t}{4} \left[\frac{1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda - \kappa_\lambda} - 2 \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda - \kappa_\lambda} + 3 \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} + \frac{1}{\Lambda_\lambda - \kappa_\lambda} - \frac{3}{3\Lambda_\lambda + \kappa_\lambda} \right] = \frac{\kappa_{\mathcal{H}} t}{3}, \tag{A92}$$

and the limit of the expression in (A90) becomes with (A85)

$$\begin{aligned} \lim_{\lambda \nearrow 1} \left\{ \frac{\Lambda_\lambda}{\Lambda_\lambda - \kappa_\lambda} \cdot \frac{\kappa_{\mathcal{H}}^2}{2\Lambda_\lambda} \cdot \left[\frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} - \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda - \kappa_\lambda} + \frac{1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)t}}{\Lambda_\lambda - \kappa_\lambda} \right] \right. \\ \left. - \frac{\kappa_{\mathcal{H}}^2}{2} \cdot \frac{1 - e^{-\frac{1}{2}(3\Lambda_\lambda + \kappa_\lambda)t}}{3\Lambda_\lambda + \kappa_\lambda} \cdot \left[\frac{1}{\Lambda_\lambda - \kappa_\lambda} - \frac{3}{3\Lambda_\lambda + \kappa_\lambda} \right] = -\frac{\kappa_{\mathcal{H}} t}{3}. \right. \tag{A93} \end{aligned}$$

By putting (A91)–(A93) together with (A85) we finally end up with

$$\lim_{\lambda \nearrow 1} \frac{\partial U_\lambda^{(2)}(t)}{\partial \lambda} = \left[\frac{\kappa_{\mathcal{H}} t}{3} - \frac{\kappa_{\mathcal{H}} t}{3} \right] + \kappa_{\mathcal{H}} \left(-\frac{t}{6} + \frac{t}{6} - \frac{t}{2} + t - \frac{t}{2} \right) + \left[-\kappa_{\mathcal{H}}^2 + \frac{\kappa_{\mathcal{H}}^2}{2} \right] \cdot 0 = 0,$$

which finishes the proof of Lemma A7. □

Proof of Theorem 12. Recall from (131) the approximative Poisson offspring-distribution parameter $\beta_\bullet^{(m)} := 1 - \frac{\kappa_\bullet}{\sigma^2 m}$ and Poisson immigration-distribution parameter $\alpha_\bullet^{(m)} := \beta_\bullet^{(m)} \cdot \frac{\eta}{\sigma^2}$, which is a special case of $(\beta_{\mathcal{A}}^{(m)}, \beta_{\mathcal{H}}^{(m)}, \alpha_{\mathcal{A}}^{(m)}, \alpha_{\mathcal{H}}^{(m)}) \in \mathcal{P}_{NI} \cup \mathcal{P}_{SP,1}$. Let us first calculate $\lim_{m \rightarrow \infty} I \left(P_{\mathcal{A}, [\sigma^2 m t]}^{(m)} \middle| \middle| P_{\mathcal{H}, [\sigma^2 m t]}^{(m)} \right)$ by starting

from Theorem 3(a). Correspondingly, we evaluate for all $\kappa_A \geq 0, \kappa_H \geq 0$ with $\kappa_A \neq \kappa_H$ by a twofold application of l’Hospital’s rule

$$\begin{aligned} \lim_{m \rightarrow \infty} m^2 \cdot \left[\beta_A^{(m)} \cdot \left(\log \left(\frac{\beta_A^{(m)}}{\beta_H^{(m)}} \right) - 1 \right) + \beta_H^{(m)} \right] &= \lim_{m \rightarrow \infty} \frac{-m}{2\sigma^2} \left[\kappa_A \log \left(\frac{\beta_A^{(m)}}{\beta_H^{(m)}} \right) + \kappa_H \left(1 - \frac{\beta_A^{(m)}}{\beta_H^{(m)}} \right) \right] \\ &= \frac{1}{2\sigma^4} \cdot \lim_{m \rightarrow \infty} \frac{\beta_H^{(m)} \cdot \kappa_A - \beta_A^{(m)} \cdot \kappa_H}{\left(\beta_H^{(m)} \right)^2} \cdot \left(\kappa_A \cdot \frac{\beta_H^{(m)}}{\beta_A^{(m)}} - \kappa_H \right) = \frac{(\kappa_A - \kappa_H)^2}{2\sigma^4}. \end{aligned} \tag{A94}$$

Additionally there holds

$$\lim_{m \rightarrow \infty} m \cdot (1 - \beta_A^{(m)}) = \frac{\kappa_A}{\sigma^2} \quad \text{and} \quad \lim_{m \rightarrow \infty} \left(\beta_A^{(m)} \right)^{\lfloor \sigma^2 m t \rfloor} = \lim_{m \rightarrow \infty} \left[\left(1 - \frac{\kappa_A}{\sigma^2 m} \right)^m \right]^{\lfloor \sigma^2 m t \rfloor / m} = e^{-\kappa_A \cdot t}. \tag{A95}$$

For $\kappa_A > 0$, we apply the upper part of formula (69) as well as (A94) and (A95) to derive

$$\begin{aligned} \lim_{m \rightarrow \infty} I_\lambda \left(P_{\mathcal{A}, \lfloor \sigma^2 m t \rfloor}^{(m)} \middle| \middle| P_{\mathcal{H}, \lfloor \sigma^2 m t \rfloor}^{(m)} \right) &= \lim_{m \rightarrow \infty} \left[\frac{m^2 \cdot \left[\beta_A^{(m)} \cdot \left(\log \left(\frac{\beta_A^{(m)}}{\beta_H^{(m)}} \right) - 1 \right) + \beta_H^{(m)} \right]}{m \cdot (1 - \beta_A^{(m)})} \right. \\ &\cdot \left[\frac{X_0^{(m)}}{m} - \frac{\alpha_A^{(m)}}{m \cdot (1 - \beta_A^{(m)})} \right] \cdot \left(1 - \left(\beta_A^{(m)} \right)^{\lfloor \sigma^2 m t \rfloor} \right) \\ &+ \frac{\alpha_A^{(m)}}{\beta_A^{(m)} \cdot m \cdot (1 - \beta_A^{(m)})} \cdot m^2 \cdot \left[\beta_A^{(m)} \cdot \left(\log \left(\frac{\beta_A^{(m)}}{\beta_H^{(m)}} \right) - 1 \right) + \beta_H^{(m)} \right] \cdot \frac{\lfloor \sigma^2 m t \rfloor}{m} \left. \right] \\ &= \frac{(\kappa_A - \kappa_H)^2}{2\sigma^2 \cdot \kappa_A} \cdot \left[\left(\tilde{X}_0 - \frac{\eta}{\kappa_A} \right) \cdot (1 - e^{-\kappa_A \cdot t}) + \eta \cdot t \right]. \end{aligned}$$

For $\kappa_A = 0$ (and thus $\kappa_H > 0, \beta_A^{(m)} \equiv 1, \alpha_A^{(m)} \equiv \eta / \sigma^2$), we apply the lower part of formula (69) as well as (A94) and (A95) to obtain

$$\begin{aligned} \lim_{m \rightarrow \infty} I_\lambda \left(P_{\mathcal{A}, \lfloor \sigma^2 m t \rfloor}^{(m)} \middle| \middle| P_{\mathcal{H}, \lfloor \sigma^2 m t \rfloor}^{(m)} \right) &= \left\{ \lim_{m \rightarrow \infty} m^2 \cdot \left[\beta_H^{(m)} - \log \beta_H^{(m)} - 1 \right] \right. \\ &\cdot \left[\frac{\eta}{2\sigma^2} \cdot \frac{(\lfloor \sigma^2 m t \rfloor)^2}{m^2} + \left(\frac{X_0^{(m)}}{m} + \frac{\eta}{2\sigma^2 \cdot m} \right) \cdot \frac{\lfloor \sigma^2 m t \rfloor}{m} \left. \right\} = \frac{\kappa_H^2}{2\sigma^2} \cdot \left[\frac{\eta}{2} \cdot t^2 + \tilde{X}_0 \cdot t \right]. \end{aligned}$$

Let us now calculate the “converse” double limit

$$\lim_{\lambda \nearrow 1} \lim_{m \rightarrow \infty} I_\lambda \left(P_{\mathcal{A}, \lfloor \sigma^2 m t \rfloor}^{(m)} \middle| \middle| P_{\mathcal{H}, \lfloor \sigma^2 m t \rfloor}^{(m)} \right) = \lim_{\lambda \nearrow 1} \lim_{m \rightarrow \infty} \frac{1 - H_\lambda \left(P_{\mathcal{A}, \lfloor \sigma^2 m t \rfloor}^{(m)} \middle| \middle| P_{\mathcal{H}, \lfloor \sigma^2 m t \rfloor}^{(m)} \right)}{\lambda \cdot (1 - \lambda)}.$$

This will be achieved by evaluating for each $t > 0$ the two limits

$$\lim_{\lambda \nearrow 1} \frac{1 - D_{\lambda, \tilde{X}_0, t}^L}{\lambda \cdot (1 - \lambda)} \quad \text{and} \quad \lim_{\lambda \nearrow 1} \frac{1 - D_{\lambda, \tilde{X}_0, t}^U}{\lambda \cdot (1 - \lambda)} \tag{A96}$$

which will turn out to coincide; the involved lower and upper bound $D_{\lambda, \tilde{X}_0, t}^L, D_{\lambda, \tilde{X}_0, t}^U$ defined by (153) and (154) satisfy $\lim_{\lambda \nearrow 1} D_{\lambda, \tilde{X}_0, t}^L = \lim_{\lambda \nearrow 1} D_{\lambda, \tilde{X}_0, t}^U = 1$ as an easy consequence of the limits (cf. 150)

$$\lim_{\lambda \nearrow 1} \Lambda_\lambda = \kappa_A \geq 0 \quad \text{and} \quad \lim_{\lambda \nearrow 1} \kappa_\lambda = \kappa_A \geq 0, \tag{A97}$$

as well as the formulas (A82) and (A83) for the case $\kappa_A = 0$. Accordingly, we compute

$$\begin{aligned} \lim_{\lambda \nearrow 1} \frac{1 - D_{\lambda, \tilde{X}_0, t}^L}{\lambda \cdot (1 - \lambda)} &= \lim_{\lambda \nearrow 1} \frac{-D_{\lambda, \tilde{X}_0, t}^L}{1 - 2\lambda} \frac{\partial}{\partial \lambda} \left[-\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \cdot \left[\tilde{X}_0 - \frac{\eta}{\Lambda_\lambda} \right] \cdot \left(1 - e^{-\Lambda_\lambda \cdot t} \right) - \frac{\eta}{\sigma^2} \cdot (\Lambda_\lambda - \kappa_\lambda) \cdot t \right. \\ &\quad \left. + L_\lambda^{(1)}(t) \cdot \tilde{X}_0 + \frac{\eta}{\sigma^2} \cdot L_\lambda^{(2)}(t) \right] \\ &= \lim_{\lambda \nearrow 1} \left\{ -\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \left[\left(\tilde{X}_0 - \frac{\eta}{\Lambda_\lambda} \right) \cdot t e^{-\Lambda_\lambda \cdot t} \cdot \frac{\partial \Lambda_\lambda}{\partial \lambda} + \left(1 - e^{-\Lambda_\lambda \cdot t} \right) \cdot \frac{\eta}{\Lambda_\lambda^2} \cdot \frac{\partial \Lambda_\lambda}{\partial \lambda} \right] \right. \\ &\quad \left. - \frac{1}{\sigma^2} \cdot \frac{\partial}{\partial \lambda} (\Lambda_\lambda - \kappa_\lambda) \cdot \left(\tilde{X}_0 - \frac{\eta}{\Lambda_\lambda} \right) \cdot \left(1 - e^{-\Lambda_\lambda \cdot t} \right) - \frac{\eta t}{\sigma^2} \cdot \frac{\partial}{\partial \lambda} (\Lambda_\lambda - \kappa_\lambda) \right. \\ &\quad \left. + \tilde{X}_0 \frac{\partial L_\lambda^{(1)}(t)}{\partial \lambda} + \frac{\eta}{\sigma^2} \frac{\partial L_\lambda^{(2)}(t)}{\partial \lambda} \right\}, \quad \text{with} \end{aligned} \tag{A98}$$

$$\frac{\partial \Lambda_\lambda}{\partial \lambda} = \frac{\kappa_A^2 - \kappa_H^2}{2 \Lambda_\lambda} \quad \text{and} \quad \frac{\partial \kappa_\lambda}{\partial \lambda} = \kappa_A - \kappa_H. \tag{A99}$$

For the case $\kappa_A > 0$, one can combine this with (A97) and (A74) to end up with

$$\lim_{\lambda \nearrow 1} \frac{1 - D_{\lambda, \tilde{X}_0, t}^L}{\lambda \cdot (1 - \lambda)} = \frac{(\kappa_A - \kappa_H)^2}{2\sigma^2 \cdot \kappa_A} \cdot \left[\left(\tilde{X}_0 - \frac{\eta}{\kappa_A} \right) \cdot \left(1 - e^{-\kappa_A \cdot t} \right) + \eta \cdot t \right]. \tag{A100}$$

For the case $\kappa_A = 0$, we continue the calculation (A98) by rearranging terms and by employing the Formulas (A75), (A76), (A82) and (A83) as well as the obvious relation $\frac{1}{\Lambda} - \frac{\Lambda - \kappa_\lambda}{\Lambda^2} = \frac{1}{\kappa_H}$ and obtain

$$\begin{aligned} \lim_{\lambda \nearrow 1} \frac{1 - D_{\lambda, \tilde{X}_0, t}^L}{\lambda \cdot (1 - \lambda)} &= \lim_{\lambda \nearrow 1} \left\{ \frac{\kappa_H^2 \cdot \tilde{X}_0}{2\sigma^2} \left[\frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \cdot t \cdot e^{-\Lambda_\lambda t} + \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} \right] \right. \\ &\quad \left. + \frac{\eta \cdot \kappa_H^2 \cdot t}{2\sigma^2} \left[\frac{1}{\Lambda_\lambda} - \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda^2} + \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda} \cdot \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} \right] - \frac{\eta \cdot \kappa_H^2}{2\sigma^2} \cdot \frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} \left[\frac{1}{\Lambda_\lambda} - \frac{\Lambda_\lambda - \kappa_\lambda}{\Lambda_\lambda^2} \right] \right. \\ &\quad \left. - \frac{\kappa_H \cdot \tilde{X}_0}{\sigma^2} \left(1 - e^{-\Lambda_\lambda t} \right) + \frac{\eta \cdot \kappa_H}{\sigma^2} \left[\frac{1 - e^{-\Lambda_\lambda t}}{\Lambda_\lambda} - t \right] + \frac{\partial L_\lambda^{(1)}(t)}{\partial \lambda} \cdot \tilde{X}_0 + \frac{\eta}{\sigma^2} \cdot \frac{\partial L_\lambda^{(2)}(t)}{\partial \lambda} \right\} \\ &= \frac{\kappa_H^2 \tilde{X}_0 t}{\sigma^2} + \frac{\eta \kappa_H^2 t}{2\sigma^2} \left[\frac{1}{\kappa_H} + t \right] - \frac{\eta \kappa_H t}{2\sigma^2} - \frac{\kappa_H^2 \tilde{X}_0 t}{2\sigma^2} - \frac{\eta \kappa_H^2 t^2}{4\sigma^2} = \frac{\kappa_H^2}{2\sigma^2} \cdot \left[\frac{\eta}{2} \cdot t^2 + \tilde{X}_0 \cdot t \right]. \end{aligned} \tag{A101}$$

Let us now turn to the second limit (A96) for which we compute analogously to (A98)

$$\begin{aligned} \lim_{\lambda \nearrow 1} \frac{1 - D_{\lambda, \tilde{X}_0, t}^U}{\lambda \cdot (1 - \lambda)} &= \lim_{\lambda \nearrow 1} \frac{-D_{\lambda, \tilde{X}_0, t}^U}{1 - 2\lambda} \frac{\partial}{\partial \lambda} \left[-\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \cdot \left[\tilde{X}_0 - \frac{\eta}{\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)} \right] \cdot \left(1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \right) \right. \\ &\quad \left. - \frac{\eta}{\sigma^2} \cdot (\Lambda_\lambda - \kappa_\lambda) \cdot t - U_\lambda^{(1)}(t) \cdot \tilde{X}_0 - \frac{\eta}{\sigma^2} \cdot U_\lambda^{(2)}(t) \right] \\ &= \lim_{\lambda \nearrow 1} \left\{ -\frac{\Lambda_\lambda - \kappa_\lambda}{\sigma^2} \left[\left(\tilde{X}_0 - \frac{\eta}{\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)} \right) \cdot \frac{t}{2} \cdot e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \cdot \frac{\partial}{\partial \lambda} (\Lambda_\lambda + \kappa_\lambda) \right. \right. \\ &\quad \left. \left. + \left(1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \right) \cdot \frac{2 \cdot \eta}{(\Lambda_\lambda + \kappa_\lambda)^2} \cdot \frac{\partial}{\partial \lambda} (\Lambda_\lambda + \kappa_\lambda) \right] \right. \\ &\quad \left. - \frac{1}{\sigma^2} \cdot \frac{\partial}{\partial \lambda} (\Lambda_\lambda - \kappa_\lambda) \cdot \left(\tilde{X}_0 - \frac{\eta}{\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda)} \right) \cdot \left(1 - e^{-\frac{1}{2}(\Lambda_\lambda + \kappa_\lambda) \cdot t} \right) - \frac{\eta t}{\sigma^2} \cdot \frac{\partial}{\partial \lambda} (\Lambda_\lambda - \kappa_\lambda) \right. \\ &\quad \left. - \frac{\partial U_\lambda^{(1)}(t)}{\partial \lambda} \cdot \tilde{X}_0 - \frac{\eta}{\sigma^2} \frac{\partial U_\lambda^{(2)}(t)}{\partial \lambda} \right\}. \end{aligned} \tag{A102}$$

For the case $\kappa_{\mathcal{A}} > 0$, one can combine this with (A97), (A99) and (A74) to end up with

$$\lim_{\lambda \nearrow 1} \frac{1 - D_{\lambda, \tilde{X}_0, t}^U}{\lambda \cdot (1 - \lambda)} = \frac{(\kappa_{\mathcal{A}} - \kappa_{\mathcal{H}})^2}{2\sigma^2 \cdot \kappa_{\mathcal{A}}} \cdot \left[\left(\tilde{X}_0 - \frac{\eta}{\kappa_{\mathcal{A}}} \right) \cdot (1 - e^{-\kappa_{\mathcal{A}} \cdot t}) + \eta \cdot t \right]. \quad (\text{A103})$$

For the case $\kappa_{\mathcal{A}} = 0$, we continue the calculation of (A102) by rearranging terms and by employing the formulas (A77), (A82) and (A83) as well as the obvious relation $\lim_{\lambda \nearrow 1} \frac{1}{\Lambda_{\lambda}} - \frac{\Lambda_{\lambda} - \kappa_{\lambda}}{\Lambda_{\lambda}(\Lambda_{\lambda} + \kappa_{\lambda})} = \frac{2}{\kappa_{\mathcal{H}}}$ to obtain

$$\begin{aligned} \lim_{\lambda \nearrow 1} \frac{1 - D_{\lambda, \tilde{X}_0, t}^U}{\lambda \cdot (1 - \lambda)} &= \lim_{\lambda \nearrow 1} \left\{ \frac{t \cdot \tilde{X}_0}{4\sigma^2} \cdot \frac{\Lambda_{\lambda} - \kappa_{\lambda}}{\Lambda_{\lambda}} \cdot e^{-\frac{1}{2}(\Lambda_{\lambda} + \kappa_{\lambda}) \cdot t} \left(\kappa_{\mathcal{H}}^2 + 2\Lambda_{\lambda} \kappa_{\mathcal{H}} \right) \right. \\ &+ \frac{\tilde{X}_0}{2\sigma^2} \cdot \frac{1 - e^{-\frac{1}{2}(\Lambda_{\lambda} + \kappa_{\lambda}) \cdot t}}{\Lambda_{\lambda}} \left(\kappa_{\mathcal{H}}^2 - 2\Lambda_{\lambda} \kappa_{\mathcal{H}} \right) - \frac{\eta \cdot t}{\sigma^2} \left[\kappa_{\mathcal{H}} \left(1 + e^{-\frac{1}{2}(\Lambda_{\lambda} + \kappa_{\lambda}) \cdot t} \frac{\Lambda_{\lambda} - \kappa_{\lambda}}{\Lambda_{\lambda} + \kappa_{\lambda}} \right) \right. \\ &- \left. \left. \frac{\kappa_{\mathcal{H}}^2}{2} \cdot \left(\frac{1}{\Lambda_{\lambda}} - \frac{\Lambda_{\lambda} - \kappa_{\lambda}}{\Lambda_{\lambda}(\Lambda_{\lambda} + \kappa_{\lambda})} + \frac{\Lambda_{\lambda} - \kappa_{\lambda}}{\Lambda_{\lambda} + \kappa_{\lambda}} \cdot \frac{1 - e^{-\frac{1}{2}(\Lambda_{\lambda} + \kappa_{\lambda}) \cdot t}}{\Lambda_{\lambda}} \right) \right] \right\} \\ &+ \frac{2\eta}{\sigma^2} \cdot \frac{1 - e^{-\frac{1}{2}(\Lambda_{\lambda} + \kappa_{\lambda}) \cdot t}}{\Lambda_{\lambda} + \kappa_{\lambda}} \left[\kappa_{\mathcal{H}} \left(1 + \frac{\Lambda_{\lambda} - \kappa_{\lambda}}{\Lambda_{\lambda} + \kappa_{\lambda}} \right) - \frac{\kappa_{\mathcal{H}}^2}{2} \left(\frac{1}{\Lambda_{\lambda}} - \frac{\Lambda_{\lambda} - \kappa_{\lambda}}{\Lambda_{\lambda}(\Lambda_{\lambda} + \kappa_{\lambda})} \right) \right] \\ &- \left. \frac{\partial U_{\lambda}^{(1)}(t)}{\partial \lambda} \cdot \tilde{X}_0 - \frac{\eta}{\sigma^2} \frac{\partial U_{\lambda}^{(2)}(t)}{\partial \lambda} \right\} \\ &= \frac{\kappa_{\mathcal{H}}^2 t \tilde{X}_0}{4\sigma^2} + \frac{\kappa_{\mathcal{H}}^2 t \tilde{X}_0}{4\sigma^2} - \frac{\eta t}{\sigma^2} \left[2\kappa_{\mathcal{H}} - \kappa_{\mathcal{H}} - \frac{\kappa_{\mathcal{H}}^2 t}{4} \right] + \frac{\eta t}{\sigma^2} [2\kappa_{\mathcal{H}} - \kappa_{\mathcal{H}}] = \frac{\kappa_{\mathcal{H}}^2}{2\sigma^2} \left[\frac{\eta}{2} \cdot t^2 + \tilde{X}_0 \cdot t \right]. \quad (\text{A104}) \end{aligned}$$

Since (A100) coincides with (A103) and (A101) coincides with (A104), we have finished the proof. \square

References

- Liese, F.; Vajda, I. *Convex Statistical Distances*; Teubner: Leipzig, Germany, 1987.
- Read, T.R.C.; Cressie N.A.C. *Goodness-of-Fit Statistics for Discrete Multivariate Data*; Springer: New York, NY, USA, 1988.
- Vajda, I. *Theory of Statistical Inference and Information*; Kluwer: Dordrecht, The Netherlands, 1989.
- Csiszár, I.; Shields, P.C. *Information Theory and Statistics: A Tutorial*; Now Publishers: Hanover, MA, USA, 2004.
- Stummer, W. *Exponentials, Diffusions, Finance, Entropy and Information*; Shaker: Aachen, Germany, 2004.
- Pardo, L. *Statistical Inference Based on Divergence Measures*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2006.
- Liese, F.; Miescke, K.J. *Statistical Decision Theory: Estimation, Testing, and Selection*; Springer: New York, NY, USA, 2008.
- Basu, A.; Shioya, H.; Park, C. *Statistical Inference: The Minimum Distance Approach*; CRC Press: Boca Raton, FL, USA, 2011.
- Voinov V.; Nikulin, M.; Balakrishnan N. *Chi-Squared Goodness of Fit Tests with Applications*; Academic Press: Waltham, MA, USA, 2013.
- Liese, F.; Vajda, I. On divergences and informations in statistics and information theory. *IEEE Trans. Inform. Theory* **2006**, *52*, 4394–4412. [[CrossRef](#)]
- Vajda I.; van der Meulen, E.C. Goodness-of-fit criteria based on observations quantized by hypothetical and empirical percentiles. In *Handbook of Fitting Statistical Distributions with R*; Karian, Z.A., Dudewicz, E.J., Eds.; CRC: Heidelberg, Germany, 2010; pp. 917–994.
- Stummer, W.; Vajda, I. On Bregman distances and divergences of probability measures. *IEEE Trans. Inform. Theory* **2012**, *58*, 1277–1288. [[CrossRef](#)]
- Kißlinger, A.-L.; Stummer, W. Robust statistical engineering by means of scaled Bregman distances. In *Recent Advances in Robust Statistics—Theory and Applications*; Agostinelli, C., Basu, A., Filzmoser, P., Mukherjee, D., Eds.; Springer: New Delhi, India, 2016; pp. 81–113.

14. Broniatowski, M.; Stummer, W. Some universal insights on divergences for statistics, machine learning and artificial intelligence. In *Geometric Structures of Information*; Nielsen, F., Ed.; Springer: Cham, Switzerland, 2019; pp. 149–211.
15. Stummer, W.; Vajda, I. Optimal statistical decisions about some alternative financial models. *J. Econom.* **2007**, *137*, 441–471. [[CrossRef](#)]
16. Stummer, W.; Lao, W. Limits of Bayesian decision related quantities of binomial asset price models. *Kybernetika* **2012**, *48*, 750–767.
17. Csizar, I. Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. *Publ. Math. Inst. Hungar. Acad. Sci.* **1963**, *A-8*, 85–108.
18. Ali, M.S.; Silvey, D. A general class of coefficients of divergence of one distribution from another. *J. Roy. Statist. Soc. B* **1966**, *28*, 131–140. [[CrossRef](#)]
19. Morimoto, T. Markov processes and the H-theorem. *J. Phys. Soc. Jpn.* **1963**, *18*, 328–331. [[CrossRef](#)]
20. van Erven, T.; Harremoës, P. Renyi divergence and Kullback-Leibler divergence. *IEEE Trans. Inf. Theory* **2014**, *60*, 3797–3820. [[CrossRef](#)]
21. Newman, C.M. On the orthogonality of independent increment processes. In *Topics in Probability Theory*; Courant Institute of Mathematical Sciences New York University: New York, NY, USA, 1973; pp. 93–111.
22. Liese, F. Hellinger integrals of Gaussian processes with independent increments. *Stochastics* **1982**, *6*, 81–96. [[CrossRef](#)]
23. Memin, J.; Shiryaev, A.N. Distance de Hellinger-Kakutani des lois correspondant a deux processus a accroissements indépendants. *Probab. Theory Relat. Fields* **1985**, *70*, 67–89. [[CrossRef](#)]
24. Jacod, J.; Shiryaev, A.N. *Limit Theorems for Stochastic Processes*; Springer: Berlin, Germany, 1987.
25. Linkov, Y.N.; Shevlyakov, Y.A. Large deviation theorems in the hypotheses testing problems for processes with independent increments. *Theory Stoch. Process* **1998**, *4*, 198–210.
26. Liese, F. Hellinger integrals, error probabilities and contiguity of Gaussian processes with independent increments and Poisson processes. *J. Inf. Process. Cybern.* **1985**, *21*, 297–313.
27. Kabanov, Y.M.; Liptser, R.S.; Shiryaev, A.N. On the variation distance for probability measures defined on a filtered space. *Probab. Theory Relat. Fields* **1986**, *71*, 19–35. [[CrossRef](#)]
28. Liese, F. Hellinger integrals of diffusion processes. *Statistics* **1986**, *17*, 63–78. [[CrossRef](#)]
29. Vajda, I. Distances and discrimination rates for stochastic processes. *Stoch. Process. Appl.* **1990**, *35*, 47–57. [[CrossRef](#)]
30. Stummer, W. The Novikov and entropy conditions of multidimensional diffusion processes with singular drift. *Probab. Theory Relat. Fields* **1993**, *97*, 515–542. [[CrossRef](#)]
31. Stummer, W. On a statistical information measure of diffusion processes. *Stat. Decis.* **1999**, *17*, 359–376. [[CrossRef](#)]
32. Stummer, W. On a statistical information measure for a generalized Samuelson-Black-Scholes model. *Stat. Decis.* **2001**, *19*, 289–314. [[CrossRef](#)]
33. Bartoszyński, R. Branching processes and the theory of epidemics. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. IV*; Le Cam, L.M., Neyman, J., Eds; University of California Press: Berkeley, CA, USA, 1967; pp. 259–269.
34. Ludwig, D. Qualitative behaviour of stochastic epidemics. *Math. Biosci.* **1975**, *23*, 47–73. [[CrossRef](#)]
35. Becker, N.G. Estimation for an epidemic model. *Biometrics* **1976**, *32*, 769–777. [[CrossRef](#)]
36. Becker, N.G. Estimation for discrete time branching processes with applications to epidemics. *Biometrics* **1977**, *33*, 515–522. [[CrossRef](#)]
37. Metz, J.A.J. The epidemic in a closed population with all susceptibles equally vulnerable; some results for large susceptible populations and small initial infections. *Acta Biotheor.* **1978**, *27*, 75–123. [[CrossRef](#)] [[PubMed](#)]
38. Heyde, C.C. On assessing the potential severity of an outbreak of a rare infectious disease. *Austral. J. Stat.* **1979**, *21*, 282–292. [[CrossRef](#)]
39. Von Bahr, B.; Martin-Löf, A. Threshold limit theorems for some epidemic processes. *Adv. Appl. Prob.* **1980**, *12*, 319–349. [[CrossRef](#)]
40. Ball, F. The threshold behaviour of epidemic models. *J. Appl. Prob.* **1983**, *20*, 227–241. [[CrossRef](#)]
41. Jacob, C. Branching processes: Their role in epidemics. *Int. J. Environ. Res. Public Health* **2010**, *7*, 1186–1204. [[CrossRef](#)]

42. Barbour, A.D.; Reinert, G. Approximating the epidemic curve. *Electron. J. Probab.* **2013**, *18*, 1–30. [[CrossRef](#)]
43. Britton, T.; Pardoux, E. Stochastic epidemics in a homogeneous community. In *Stochastic Epidemic Models*; Britton, T., Pardoux, E., Eds.; Springer: Cham, Switzerland, 2019; pp. 1–120.
44. Dion, J.P.; Gauthier, G.; Latour, A. Branching processes with immigration and integer-valued time series. *Serdica Math. J.* **1995**, *21*, 123–136.
45. Grunwald, G.K.; Hyndman, R.J.; Tedesco, L.; Tweedie, R.L. Non-Gaussian conditional linear AR(1) models. *Aust. N. Z. J. Stat.* **2000**, *42*, 479–495. [[CrossRef](#)]
46. Kedem, B.; Fokianos, K. *An Regression Models for Time Series Analysis*; Wiley: Hoboken, NJ, USA, 2002.
47. Held, L.; Höhle, M.; Hofmann, M. A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Stat. Model.* **2005**, *5*, 187–199. [[CrossRef](#)]
48. Weiss, C.H. *An Introduction to Discrete-Valued Time Series*; Wiley: Hoboken, NJ, USA, 2018.
49. Feigin, P.D.; Passy, U. The geometric programming dual to the extinction probability problem in simple branching processes. *Ann. Probab.* **1981**, *9*, 498–503. [[CrossRef](#)]
50. Mordecki, E. Asymptotic mixed normality and Hellinger processes. *Stoch. Stoch. Rep.* **1994**, *48*, 129–143. [[CrossRef](#)]
51. Sriram, T.N.; Vidyashankar, A.N. Minimum Hellinger distance estimation for supercritical Galton-Watson processes. *Stat. Probab. Lett.* **2000**, *50*, 331–342. [[CrossRef](#)]
52. Guttorp, P. *Statistical Inference for Branching Processes*; Wiley: New York, NY, USA, 1991.
53. Linkov, Y.N.; Lunyova, L.A. Large deviation theorems in the hypothesis testing problems for the Galton-Watson processes with immigration. *Theory Stoch. Process* **1996**, *2*, 120–132; Erratum in *Theory Stoch. Process* **1997**, *3*, 270–285.
54. Heathcote, C.R. A branching process allowing immigration. *J. R. Stat. Soc. B* **1965**, *27*, 138–143; Erratum in: Heathcote, C.R. Corrections and comments on the paper “A branching process allowing immigration”. *J. R. Stat. Soc. B* **1966**, *28*, 213–217. [[CrossRef](#)]
55. Athreya, K.B.; Ney, P.E. *Branching Processes*; Springer: New York, NY, USA, 1972.
56. Jagers, P. *Branching Processes with Biological Applications*; Wiley: London, UK, 1975.
57. Asmussen, S.; Hering, H. *Branching Processes*; Birkhäuser: Boston, MA, USA, 1983.
58. Haccou, P.; Jagers, P.; Vatutin, V.A. *Branching Processes: Variation, Growth, and Extinction of Populations*; Cambridge University Press: Cambridge, UK, 2005.
59. Heyde, C.C.; Seneta, E. Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Probab.* **1972**, *9*, 235–256. [[CrossRef](#)]
60. Basawa, I.V.; Rao, B.L.S. *Statistical Inference of Stochastic Processes*; Academic Press: London, UK, 1980.
61. Basawa, I.V.; Scott, D.J. *Asymptotic Optimal Inference for Non-Ergodic Models*; Springer: New York, NY, USA, 1983.
62. Sankaranarayanan, G. *Branching Processes and Its Estimation Theory*; Wiley: New Delhi, India, 1989.
63. Wei, C.Z.; Winnicki, J. Estimation of the means in the branching process with immigration. *Ann. Stat.* **1990**, *18*, 1757–1773. [[CrossRef](#)]
64. Winnicki, J. Estimation of the variances in the branching process with immigration. *Probab. Theory Relat. Fields* **1991**, *88*, 77–106. [[CrossRef](#)]
65. Yanev, N.M. Statistical inference for branching processes. In *Records and Branching Processes*; Ahsanullah, M., Yanev, G.P., Eds.; Nova Science Publishes: New York, NY, USA, 2008; pp. 147–172.
66. Harris, T.E. *The Theory of Branching Processes*; Springer: Berlin, Germany, 1963.
67. Gauthier, G.; Latour, A. Convergence forte des estimateurs des paramètres d’un processus GENAR(p). *Ann. Sci. Math. Que.* **1994**, *18*, 49–71.
68. Latour, A. Existence and stochastic structure of a non-negative integer-valued autoregressive process. *J. Time Ser. Anal.* **1998**, *19*, 439–455. [[CrossRef](#)]
69. Rydberg, T.H.; Shephard, N. BIN models for trade-by-trade data. Modelling the number of trades in a fixed interval of time. In *Econometric Society World Congress; Contributed Papers No. 0740*; Econometric Society: Cambridge, UK, 2000.
70. Brandt, P.T.; Williams, J.T. A linear Poisson autoregressive model: The Poisson AR(p) model. *Polit. Anal.* **2001**, *9*, 164–184. [[CrossRef](#)]

71. Heinen, A. Modelling time series count data: An autoregressive conditional Poisson model. In *Core Discussion Paper*; MPRA Paper No. 8113; University of Louvain: Louvain, Belgium, 2003; Volume 62. Available online: <https://mpra.ub.uni-muenchen.de/8113> (accessed on 18 May 2020).
72. Held, L.; Hofmann, M.; Höhle, M.; Schmid, V. A two-component model for counts of infectious diseases. *Biostatistics* **2006**, *7*, 422–437. [[CrossRef](#)] [[PubMed](#)]
73. Finkenstädt, B.F.; Bjornstad, O.N.; Grenfell, B.T. A stochastic model for extinction and recurrence of epidemics: Estimation and inference for measles outbreak. *Biostatistics* **2002**, *3*, 493–510. [[CrossRef](#)] [[PubMed](#)]
74. Ferland, R.; Latour, A.; Oraichi, D. Integer-valued GARCH process. *J. Time Ser. Anal.* **2006**, *27*, 923–942. [[CrossRef](#)]
75. Weiß, C.H. Modelling time series of counts with overdispersion. *Stat. Methods Appl.* **2009**, *18*, 507–519. [[CrossRef](#)]
76. Weiß, C.H. The INARCH(1) model for overdispersed time series of counts. *Comm. Stat. Sim. Comp.* **2010**, *39*, 1269–1291. [[CrossRef](#)]
77. Weiß, C.H. INARCH(1) processes: Higher-order moments and jumps. *Stat. Probab. Lett.* **2010**, *80*, 1771–1780. [[CrossRef](#)]
78. Weiß, C.H.; Testik, M.C. Detection of abrupt changes in count data time series: Cumulative sum derivations for INARCH(1) models. *J. Qual. Technol.* **2012**, *44*, 249–264. [[CrossRef](#)]
79. Kaslow, R.A.; Evans, A.S. Epidemiologic concepts and methods. In *Viral Infections of Humans*; Evans, A.S., Kaslow, R.A., Eds.; Springer: New York, NY, USA, 1997; pp. 3–58.
80. Osterholm, M.T.; Hedberg, C.W. Epidemiologic principles. In *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases*, 8th ed.; Bennett, J.E., Dolin, R., Blaser, M.J., Eds.; Elsevier: Philadelphia, PA, USA, 2015; pp. 146–157.
81. Grassly, N.C.; Fraser, C. Mathematical models of infectious disease transmission. *Nat. Rev.* **2008**, *6*, 477–487. [[CrossRef](#)]
82. Keeling, M.J.; Rohani, P. *Modeling Infectious Diseases in Humans and Animals*; Princeton UP: Princeton, NJ, USA, 2008.
83. Yan, P. Distribution theory stochastic processes and infectious disease modelling. In *Mathematical Epidemiology*; Brauer, F., van den Driessche, P., Wu, J., Eds.; Springer: Berlin, Germany, 2008; pp. 229–293.
84. Yan, P.; Chowell, G. *Quantitative Methods for Investigating Infectious Disease Outbreaks*; Springer: Cham, Switzerland, 2019.
85. Britton, T. Stochastic epidemic models: A survey. *Math. Biosci.* **2010**, *225*, 24–35. [[CrossRef](#)]
86. Diekmann, O.; Heesterbeek, H.; Britton, T. *Mathematical Tools for Understanding Infectious Disease Dynamics*; Princeton University Press: Princeton, NJ, USA, 2013.
87. Cummings, D.A.T.; Lessler, J. Infectious disease dynamics. In *Infectious Disease Epidemiology: Theory and Practice*; Nelson, K.E., Masters Williams, C., Eds.; Jones & Bartlett Learning: Burlington, MA, USA, 2014; pp. 131–166.
88. Just, W.; Callender, H.; Drew LaMar, M.; Toporikova, N. Transmission of infectious diseases: Data, models and simulations. In *Algebraic and Discrete Mathematical Methods of Modern Biology*; Robeva, R.S., Ed.; Elsevier: London, UK, 2015; pp. 193–215.
89. Britton, T.; Giardina, F. Introduction to statistical inference for infectious diseases. *J. Soc. Franc. Stat.* **2016**, *157*, 53–70.
90. Fine, P.E.M. The interval between successive cases of an infectious disease. *Am. J. Epidemiol.* **2003**, *158*, 1039–1047. [[CrossRef](#)] [[PubMed](#)]
91. Svensson, A. A note on generation times in epidemic models. *Math. Biosci.* **2007**, *208*, 300–311. [[CrossRef](#)] [[PubMed](#)]
92. Svensson, A. The influence of assumptions on generation time distributions in epidemic models. *Math. Biosci.* **2015**, *270*, 81–89. [[CrossRef](#)] [[PubMed](#)]
93. Wallinga, J.; Lipsitch, M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. R. Soc. B* **2007**, *274*, 599–604. [[CrossRef](#)]
94. Forsberg White, L.; Pagano, M. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. *Stat. Med.* **2008**, *27*, 2999–3016. [[CrossRef](#)]
95. Nishiura, H. Time variations in the generation time of an infectious disease: implications for sampling to appropriately quantify transmission potential. *Math. Biosci.* **2010**, *7*, 851–869.

96. Scalia Tomba, G.; Svensson, A.; Asikainen, T.; Giesecke, J. Some model based considerations on observing generation times for communicable diseases. *Math. Biosci.* **2010**, *223*, 24–31. [[CrossRef](#)]
97. Trichereau, J.; Verret, C.; Mayet, A.; Manet, G. Estimation of the reproductive number for A(H1N1) pdm09 influenza among the French armed forces, September 2009–March 2010. *J. Infect.* **2012**, *64*, 628–630. [[CrossRef](#)]
98. Vink, M.A.; Bootsma, M.C.J.; Wallinga, J. Serial intervals of respiratory infectious diseases: a systematic review and analysis. *Am. J. Epidemiol.* **2014**, *180*, 865–875. [[CrossRef](#)]
99. Champredon, D.; Dushoff, J. Intrinsic and realized generation intervals in infectious-disease transmission. *Proc. R. Soc. B* **2015**, *282*, 20152026. [[CrossRef](#)]
100. An der Heiden, M.; Hamouda, O. Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland—Nowcasting. *Epid. Bull.* **2020**, *17*, 10–16. (In German)
101. Ferretti, L.; Wymant, C.; Kendall, M.; Zhao, L.; Nurtay, A.; Abeler-Dörner, L.; Parker, M.; Bonsall, D.; Fraser, C. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **2020**, *368*, eabb6936. [[CrossRef](#)] [[PubMed](#)]
102. Ganyani, T.; Kremer, C.; Chen, D.; Torneri, A.; Faes, C.; Wallinga, J.; Hens, N. Estimating the generation interval for COVID-19 based on symptom onset data. *medRxiv Prepr.* **2020**. [[CrossRef](#)]
103. Li, M.; Liu, K.; Song, Y.; Wang, M.; Wu, J. Serial interval and generation interval for respectively the imported and local infectors estimated using reported contact-tracing data of COVID-19 in China. *medRxiv Prepr.* **2020**. [[CrossRef](#)]
104. Nishiura, H.; Linton, N.M.; Akhmetzhanov, A.R. Serial interval of novel coronavirus (COVID-19) infections. *medRxiv Prepr.* **2020**. [[CrossRef](#)]
105. Park, M.; Cook, A.R.; Lim, J.J.; Sun, X.; Dickens, B.L. A systematic review of COVID-19 epidemiology based on current evidence. *J. Clin. Med.* **2020**, *9*, 967. [[CrossRef](#)]
106. Spouge, J.L. An accurate approximation for the expected site frequency spectrum in a Galton-Watson process under an infinite sites mutation model. *Theor. Popul. Biol.* **2019**, *127*, 7–15. [[CrossRef](#)]
107. Taneyhill, D.E.; Dunn, A.M.; Hatcher, M.J. The Galton-Watson branching process as a quantitative tool in parasitology. *Parasitol. Today* **1999**, *15*, 159–165. [[CrossRef](#)]
108. Parnes, D. Analyzing the contagion effect of foreclosures as a branching process: A close look at the years that follow the Great Recession. *J. Account. Financ.* **2017**, *17*, 9–34.
109. Le Cam, L. *Asymptotic Methods in Statistical Decision Theory*; Springer: New York, NY, USA, 1986.
110. Heyde, C.C.; Johnstone, I.M. On asymptotic posterior normality for stochastic processes. *J. R. Stat. Soc. B* **1979**, *41*, 184–189. [[CrossRef](#)]
111. Johnson, R.A.; Susarla, V.; van Ryzin, J. Bayesian non-parametric estimation for age-dependent branching processes. *Stoch. Proc. Appl.* **1979**, *9*, 307–318. [[CrossRef](#)]
112. Scott, D. On posterior asymptotic normality and asymptotic normality of estimators for the Galton-Watson process. *J. R. Stat. Soc. B* **1987**, *49*, 209–214. [[CrossRef](#)]
113. Yanev, N.M.; Tsokos, C.P. Decision-theoretic estimation of the offspring mean in mortal branching processes. *Comm. Stat. Stoch. Models* **1999**, *15*, 889–902. [[CrossRef](#)]
114. Mendoza, M.; Gutierrez-Pena, E. Bayesian conjugate analysis of the Galton-Watson process. *Test* **2000**, *9*, 149–171. [[CrossRef](#)]
115. Feicht, R.; Stummer, W. An explicit nonstationary stochastic growth model. In *Economic Growth and Development (Frontiers of Economics and Globalization, Vol. 11)*; De La Grandville, O., Ed.; Emerald Group Publishing Limited: Bingley, UK, 2011; pp. 141–202.
116. Dorn, F.; Fuest, C.; Göttert, M.; Krolage, C.; Lautenbacher, S.; Link, S.; Peichl, A.; Reif, M.; Sauer, S.; Stöckli, M.; et al. Die volkswirtschaftlichen Kosten des Corona-Shutdown für Deutschland: Eine Szenarienrechnung. *ifo Schnelldienst* **2020**, *73*, 29–35. (In German)
117. Dorn, F.; Khailaie, S.; Stöckli, M.; Binder, S.; Lange, B.; Peichl, A.; Vanella, P.; Wollmershäuser, T.; Fuest, C.; Meyer-Hermann, M. Das gemeinsame Interesse von Gesundheit und Wirtschaft: Eine Szenarienrechnung zur Eindämmung der Corona-Pandemie. *ifo Schnellid. Dig.* **2020**, *6*, 1–9.
118. Kißlinger, A.-L.; Stummer, W. A new toolkit for robust distributional change detection. *Appl. Stoch. Models Bus. Ind.* **2018**, *34*, 682–699. [[CrossRef](#)]
119. Dehning, J.; Zierenberg, J.; Spitzner, F.P.; Wibral, M.; Neto, J.P.; Wilczek, M.; Priesemann, V. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **2020**, *369*, eabb9789. [[CrossRef](#)]

120. Friesen, M. Statistical surveillance. Optimality and methods. *Int. Stat. Review* **2003**, *71*, 403–434. [[CrossRef](#)]
121. Friesen, M.; Andersson, E.; Schiöler, L. Robust outbreak surveillance of epidemics in Sweden. *Stat. Med.* **2009**, *28*, 476–493. [[CrossRef](#)]
122. Brauner, J.M.; Mindermann, S.; Sharma, M.; Stephenson, A.B.; Gavenciak, T.; Johnston, D.; Salvatier, J.; Leech, G.; Besiroglu, T.; Altman, G.; et al. The effectiveness and perceived burden of nonpharmaceutical interventions against COVID-19 transmission: A modelling study with 41 countries. *medRxiv Prepr.* **2020**. [[CrossRef](#)]
123. Österreicher, F.; Vajda, I. Statistical information and discrimination. *IEEE Trans. Inform. Theory* **1993**, *39*, 1036–1039. [[CrossRef](#)]
124. De Groot, M.H. Uncertainty, information and sequential experiments. *Ann. Math. Statist.* **1962**, *33*, 404–419. [[CrossRef](#)]
125. Krafft, O.; Plachky, D. Bounds for the power of likelihood ratio tests and their asymptotic properties. *Ann. Math. Stat.* **1970**, *41*, 1646–1654. [[CrossRef](#)]
126. Basawa, I.V.; Scott, D.J. Efficient tests for branching processes. *Biometrika* **1976**, *63*, 531–536. [[CrossRef](#)]
127. Feigin, P.D. The efficiency criteria problem for stochastic processes. *Stoch. Proc. Appl.* **1978**, *6*, 115–127. [[CrossRef](#)]
128. Sweeting, T.J. On efficient tests for branching processes. *Biometrika* **1978**, *65*, 123–127. [[CrossRef](#)]
129. Linkov, Y.N. *Lectures in Mathematical Statistics, Parts 1 and 2*; American Mathematical Society: Providence, RI, USA, 2005.
130. Feller, W. Diffusion processes in genetics. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*; Neyman, J., Ed; University of California Press: Berkeley, CA, USA, 1951; pp. 227–246.
131. Jirina, M. On Feller's branching diffusion process. *Časopis Pěst. Mat.* **1969**, *94*, 84–89.
132. Lamperti, J. Limiting distributions for branching processes. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. II, Part 2*; Le Cam, L.M., Neyman, J., Eds; University of California Press: Berkeley, CA, USA, 1967; pp. 225–241.
133. Lamperti, J. The limit of a sequence of branching processes. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **1967**, *7*, 271–288. [[CrossRef](#)]
134. Lindvall, T. Convergence of critical Galton-Watson branching processes. *J. Appl. Prob.* **1972**, *9*, 445–450. [[CrossRef](#)]
135. Lindvall, T. Limit theorems for some functionals of certain Galton-Watson branching processes. *Adv. Appl. Prob.* **1974**, *6*, 309–321. [[CrossRef](#)]
136. Grimvall, A. On the convergence of sequences of branching processes. *Ann. Probab.* **1974**, *2*, 1027–1045. [[CrossRef](#)]
137. Borovkov, K.A. On the convergence of branching processes to a diffusion process. *Theor. Probab. Appl.* **1986**, *30*, 496–506. [[CrossRef](#)]
138. Ethier, S.N.; Kurtz, T.G. *Markov Processes: Characterization and Convergence*; Wiley: New York, NY, USA, 1986.
139. Durrett, R. *Stochastic Calculus*; CRC Press: Boca Raton, FL, USA, 1996.
140. Kawazu, K.; Watanabe, S. Branching processes with immigration and related limit theorems. *Theor. Probab. Appl.* **1971**, *16*, 36–54. [[CrossRef](#)]
141. Wei, C.Z.; Winnicki, J. Some asymptotic results for the branching process with immigration. *Stoch. Process. Appl.* **1989**, *31*, 261–282. [[CrossRef](#)]
142. Sriram, T.N. Invalidity of bootstrap for critical branching processes with immigration. *Ann. Stat.* **1994**, *22*, 1013–1023. [[CrossRef](#)]
143. Li, Z. Branching processes with immigration and related topics. *Front. Math. China* **2006**, *1*, 73–97. [[CrossRef](#)]
144. Dawson, D.A.; Li, Z. Skew convolution semigroups and affine Markov processes. *Ann. Probab.* **2006**, *34*, 1103–1142. [[CrossRef](#)]
145. Cox, J.C.; Ingersoll, J.E., Jr.; Ross, S.A. A theory of the term structure of interest rates. *Econometrica* **1985**, *53*, 385–407. [[CrossRef](#)]
146. Cox, J.C.; Ross, S.A. The valuation of options for alternative processes. *J. Finan. Econ.* **1976**, *3*, 145–166. [[CrossRef](#)]
147. Heston, S.L. A closed-form solution for options with stochastic volatilities with applications to bond and currency options. *Rev. Finan. Stud.* **1993**, *6*, 327–343. [[CrossRef](#)]

148. Lansky, P.; Lanska, V. Diffusion approximation of the neuronal model with synaptic reversal potentials. *Biol. Cybern.* **1987**, *56*, 19–26. [[CrossRef](#)] [[PubMed](#)]
149. Giorno, V.; Lansky, P.; Nobile, A.G.; Ricciardi, L.M. Diffusion approximation and first-passage-time problem for a model neuron. *Biol. Cybern.* **1988**, *58*, 387–404. [[CrossRef](#)] [[PubMed](#)]
150. Lanska, V.; Lansky, P.; Smith, C.E. Synaptic transmission in a diffusion model for neuron activity. *J. Theor. Biol.* **1994**, *166*, 393–406. [[CrossRef](#)] [[PubMed](#)]
151. Lansky, P.; Sacerdote, L.; Tomassetti, F. On the comparison of Feller and Ornstein-Uhlenbeck models for neural activity. *Biol. Cybern.* **1995**, *73*, 457–465. [[CrossRef](#)]
152. Ditlevsen, S.; Lansky, P. Estimation of the input parameters in the Feller neuronal model. *Phys. Rev. E* **2006**, *73*, 061910. [[CrossRef](#)]
153. Höpfner, R. On a set of data for the membrane potential in a neuron. *Math. Biosci.* **2007**, *207*, 275–301. [[CrossRef](#)]
154. Lansky, P.; Ditlevsen, S. A review of the methods for signal estimation in stochastic diffusion leaky integrate-and-fire neuronal models. *Biol. Cybern.* **2008**, *99*, 253–262. [[CrossRef](#)]
155. Pedersen, A.R. Estimating the nitrous oxide emission rate from the soil surface by means of a diffusion model. *Scand. J. Stat. Theory Appl.* **2000**, *27*, 385–403. [[CrossRef](#)]
156. Aalen, O.O.; Gjessing, H.K. Survival models based on the Ornstein-Uhlenbeck process. *Lifetime Data Anal.* **2004**, *10*, 407–423. [[CrossRef](#)]
157. Kammerer, N.B. Generalized-Relative-Entropy Type Distances Between Some Branching Processes and Their Diffusion Limits. Ph.D. Thesis, University of Erlangen-Nürnberg, Erlangen, Germany, 2011.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).