Mathematical Biology

A group theoretic approach to model comparison with simplicial representations

Sean T. Vittadello[1] · **Michael P. H. Stumpf[1]**

© The Author(s) 2022

Abstract

The complexity of biological systems, and the increasingly large amount of associated experimental data, necessitates that we develop mathematical models to further our understanding of these systems. Because biological systems are generally not well understood, most mathematical models of these systems are based on experimental data, resulting in a seemingly heterogeneous collection of models that ostensibly represent the same system. To understand the system we therefore need to understand how the different models are related to each other, with a view to obtaining a unified mathematical description. This goal is complicated by the fact that a number of distinct mathematical formalisms may be employed to represent the same system, making direct comparison of the models very difficult. A methodology for comparing mathematical models based on their underlying conceptual structure is therefore required. In previous work we developed an appropriate framework for model comparison where we represent models, specifically the conceptual structure of the models, as labelled simplicial complexes and compare them with the two general methodologies of comparison by distance and comparison by equivalence. In this article we continue the development of our model comparison methodology in two directions. First, we present a rigorous and automatable methodology for the core process of comparison by equivalence, namely determining the vertices in a simplicial representation, corresponding to model components, that are conceptually related and the identification of these vertices via simplicial operations. Our methodology is based on considerations of vertex symmetry in the simplicial representation, for which we develop the required mathematical theory of group actions on simplicial complexes. This methodology greatly simplifies and expedites the process of determining model equivalence. Second, we provide an alternative mathematical framework for our model-comparison

✉ Sean T. Vittadello
   sean.vittadello@unimelb.edu.au

[1]   School of Mathematics and Statistics and School of BioSciences, The University of Melbourne, Parkville, VIC 3010, Australia

🖄 Springer

methodology by representing models as groups, which allows for the direct application of group-theoretic techniques within our model-comparison methodology.

1 Introduction

Describing biological systems mathematically remains a challenge. We can collect data at unprecedented scales (Howe et al. 2008; Marx 2013; International Society for Biocuration 2018; Mahmud et al. 2021), but we still lack the models required to integrate the often heterogeneous data into a single analytical and predictive framework. Mathematical models can achieve this, but also much more. They are of fundamental importance in developing new knowledge of these complex systems for multiple reasons: analysis and interpretation of large data sets; simulations of the biological system, or a subsystem thereof, to develop new hypotheses; and, guiding new experimental design (Tomlin and Axelrod 2007; Sneddon et al. 2010; Gunawardena 2014; Wolkenhauer 2014; Torres and Santos 2015; Pezzulo and Levin 2016; Transtrum and Qiu 2016; Banwarth-Kuhn and Sindi 2020; King et al. 2021).

Formally, a *model* is an abstraction of an observable phenomenon (Rosenblueth and Wiener 1945). Models may assume many different forms, and here we consider two general classes, conceptual models and mathematical models (Torres and Santos 2015). A *conceptual model* is a qualitative representation of the reference system under investigation, consisting of the relevant concepts which correspond to observable objects or phenomena, and the interconnections between these concepts. A conceptual model is obtained by integrating our knowledge of a system into a descriptive framework, with a range of possible levels of conceptual detail depending on our knowledge of the system. For example, in a developmental-patterning system two possible concepts are a *morphogen* and *diffusion of the morphogen*, which are dependent concepts in the system since the morphogen undergoes diffusion. Some developmental-patterning systems may have a second morphogen, which is unrelated to the first morphogen since they are independent concepts. A *mathematical model* is a quantitative representation of a conceptual model, and hence a representation of the reference system, in the formalism of mathematical concepts. Mathematical models are based on conceptual models through specification of the relevant mathematical concepts and their relationships. Mathematical models allow for the calculation of quantifiable concepts of the system, for example the spatial distribution of the morphogen concentration as a function of time.

Comparison of mathematical models based on their underlying conceptual structure is of fundamental importance for understanding the corresponding biological systems. Many aspects of complex biological systems are not well understood, so the development of mathematical models of such systems necessarily follows an inductive

process employing experimental data which provides limited constraint for the range of possible conceptual models, resulting in non-uniqueness of the mathematical models (Karplus 1977; Babtie et al. 2014). While mathematical models make conceptual models more specific and precise, such specificity may provide room for many alternative mathematical representations. Indeed, there are many possible formulations of a quantitative mathematical model from a qualitative conceptual model, depending on the particular quantitative description required, and alternative mathematical models can appear very distinct and unrelated since the underlying concepts may be obfuscated by the mathematical formalism or may be interpretable in different ways. Further, while the mathematical model is intended to be in direct correspondence with the conceptual model, and therefore the reference system, the particular formalism of the mathematical model introduces many implicit connections between the mathematical concepts in the mathematical model. Some of these interconnections may not be intended, and may be inconsistent with the underlying conceptual model; therefore the mathematical model may not correspond directly to the conceptual model. It is therefore essential to consider carefully whether the underlying conceptual model of the mathematical model represents the reference system. Indeed, since the mathematical formalism may introduce conceptual interconnections unrelated to the original system, care must be taken to ensure that the model is not a weak or incorrect representation of the system of interest.

Because of the possible multiplicity, different mathematical formalisms, and misrepresentation of mathematical models associated with a biological system, a general methodology for model comparison is required in order to further understanding of the system underlying the different types of models (Gay et al. 2010; Henkel et al. 2018). Model comparison is therefore of fundamental importance in understanding the similarities and differences between models and between data sets, and consequently in understanding biological systems. While there are many approaches to compare the quantitative outputs of mathematical models, see for example Tapinos and Mendes (2013) and Cabbia et al. (2020), such comparisons provide little insight into the underlying conceptual structure of the mathematical models and hence of the corresponding biological systems. Neither can we rely on model selection (Kirk et al. 2013) or robustness analysis, as this also relies on a comparison of the outputs of different models. These approaches cannot be used *a priori* to determine conceptual similarities, equivalences, or, indeed, differences between mathematical models.

Model comparison can simplify the landscape of mathematical modelling by organising models into categories of models that are closely related. For example, models in the same category may be incremental variations of each other, but share some fundamental qualitative characteristic. We are interested in this fundamental level of agreement: the conceptual basis of a model. Or alternatively, the set of design principles (Barnes et al. 2011; Brophy and Voigt 2014) shared by all models in the same category. The mathematical models are merely tools for quantitative computations, and the mathematical formalism can in fact be a hindrance to the model comparison at the conceptual level.

In Vittadello and Stumpf (2021) we present a framework for model comparison where models are represented as labelled simplicial complexes. Models are then compared in terms of their simplicial representations using two general methodologies:

the first is comparison by distance, where the difference between models is measured in terms of the differences between the corresponding simplicial representations; the second is comparison by equivalence, where we identify any equivalences between the components of the simplicial representations of different models, and then employ particular operations on the complexes to transform one into the other — such equivalence, when it exists, reveals common underlying characteristics of the models. Our methodology for model comparison allows for any model to be represented and compared within our framework, irrespective of formalism, granularity, and level of abstraction. To construct the simplicial representation of a model, we must first define the components of the model that we would like to represent, and for this we can fall back on relevant domain expertise. Depending on the level of detail required, the simplicial representation of a model may not be unique. One main outcome of our model-comparison methodology is for models of developmental-pattern formation (Green and Sharpe 2015; Scholes et al. 2019), where we demonstrate that the Turing-pattern activator-inhibitor model is equivalent to the positional-information annihilation model from a significant conceptual perspective.

In this article we continue the development of our methodology for model comparison (Vittadello and Stumpf 2021) in two respects. First, we develop a rigorous and automatable methodology for the main process when comparing models with respect to equivalence: the determination of the vertices that are conceptually related in a simplicial representation and the subsequent identification of these vertices via simplicial operations. We say that vertices are conceptually related when their labels, corresponding to model components, are conceptually related. Without automatability, establishing which vertices are conceptually related in a simplicial representation is challenging if not unfeasible, even for relatively small models. This methodology is based on considerations of vertex symmetry in the simplicial representation, and we establish the necessary mathematical theory for group actions on simplicial complexes. The methodology provides a major step toward simplifying and facilitating the process of determining model equivalence, which in turn can assist in the unification of models for a biological system and the discovery of design principles for synthetic biology. Second, we present a concise description of an alternative mathematical framework for our model-comparison methodology where we represent models as groups. This allows for the direct application of group-theoretic techniques within our model-comparison methodology, in addition to the current algebraic-topological techniques associated with simplicial complexes.

The remainder of this article is organised as follows. In Sect. 2 we provide an overview of the required background material: Sect. 2.1 outlines the required background in algebraic topology; we then describe our methodology for model comparison by equivalence in Sect. 2.2. We present and discuss our new results in Sect. 3. In particular, in Sect. 3.1 we provide a descriptive overview of our approach for identifying equivalent components in simplicial representations through group actions, and then develop this mathematical theory in Sect. 3.2; we present a step-by-step methodology for determining model equivalence in Sect. 3.3, which employs our automatable theory from Sect. 3.2; we then apply our methodology to a collection of Turing-pattern and positional-information models (see the Appendix in Sect. 5 for the model details) in Sect. 3.4; in Sect. 3.5 we give an incisive description of an alternative mathematical

framework for our model-comparison methodology by representing models as groups, which we call *G*-representations. Finally, in Sect. 4, we conclude with a discussion of the utility of our new theory for comparing models.

2 Background

In this section we provide a brief and informal description of the necessary background in simplicial algebraic topology; further details are available from standard references (Spanier 1966; Rotman 1988; Munkres 2018). We also discuss our methodology for model comparison by equivalence as described in Vittadello and Stumpf (2021), and developed further in this article.

2.1 Algebraic topology

Geometrically a *p-simplex* is a generalisation of a filled triangle to an arbitrary dimension *p*, whereby a point is a 0-simplex, a line segment is a 1-simplex, a filled triangle is a 2-simplex, a filled tetrahedron is a 3-simplex, and so forth. We may regard 0-simplices as vertices and 1-simplices as edges, so that we consider a simple graph as a collection of 0-simplices and 1-simplices. A *k-face* of a simplex is a *k*-dimensional subsimplex, and the 0-faces (or vertices) of a simplex *span* the simplex. If the simplex τ is a face of the simplex σ then we write $\tau \leq \sigma$. A face τ of a simplex σ is *proper* if $\tau \neq \sigma$.

A *simplicial complex* is a generalisation of a simple graph, whereby simplicies of dimension higher than one represent higher-dimensional interactions between the vertices. Specifically, a simplicial complex *K* consists of a set of simplices such that: if σ is a simplex in *K* then every face of σ is also in *K*; and, the nonempty intersection of any two simplices in *K* is a simplex in *K*. We denote the set of vertices of *K* as Vert(*K*). A *simplicial subcomplex L* of a simplicial complex *K* is a collection $L \subseteq K$ that is also a simplicial complex. A simplicial subcomplex *L* of a simplicial complex *K* is a *full subcomplex* if every simplex of *K* that has all its vertices in *L* is also a simplex of *L*. The *k-skeleton* of a simplicial complex *K* is the subcomplex $K^{(k)}$ consisting of the simplices in *K* with dimension at most *k*. In particular, the 0-skeleton $K^{(0)}$ is the set of vertices and the 1-skeleton $K^{(1)}$ is the *underlying graph* of the simplicial complex *K*.

In some contexts it is useful to define the *empty simplex*, denoted $\emptyset$, as a (-1)-simplex, and then the corresponding simplicial complex is the *empty simplicial complex*, denoted $\{\emptyset\}$, which has dimension -1. In this work we exclude the empty simplex, as it is not required. We include, however, the *void simplicial complex*, which is the simplicial complex consisting of the empty collection of sets. We denote the void simplicial complex by $\emptyset$, which has the assigned dimension of $-\infty$. Further details on these degenerate cases are found in (Kozlov 2008, Page 8, Remark 2.3).

We shall move freely between the interpretations of simplicial complexes as both geometric and abstract objects, since they are equivalent and both contain the same combinatorial information; and we have no need for the topological aspects of geo-

metric simplicial complexes. Notationally, if $v_1, v_2, \ldots, v_n$ are vertices that span a simplex K then we denote K as $\{v_1, v_2, \ldots, v_n\}$ when we want to emphasise the spanning vertices. Note that vertices are therefore denoted as both v_i and $\{v_i\}$.

2.2 Model comparison by equivalence

We refer to a specific conceptual detail of a given model as a *component* of the model. Every model consists of a finite number of components and the conceptual interconnections between these components. For example, a Turing-pattern system of reaction-diffusion equations is a model of a developmental-patterning process, and includes components such as morphogens, diffusion, reactions, boundary conditions, and a morphogen gradient, along with the interconnections between particular components such as a morphogen and its diffusion.

To compare a collection of models we first determine the set of all model components, at the required level of conceptual detail, associated with the models in the collection. The model components are the scientific concepts represented within the model, interpreted with knowledge of the relevant scientific domains.

Definition 2.1 (Model components, generated model) Let $\mathcal{C}$ be the set of all required components from the collection of models under consideration for comparison. We say that each model in the collection is *generated* by a subset of components from $\mathcal{C}$.

We represent each model in the collection for comparison as a labelled simplicial complex, which we call a *simplicial representation*, where each model component is represented as a labelled 0-simplex. The label associated with a 0-simplex corresponds to the relevant concept in the model, and for simplicity and ease of automation could have a numerical label:

Notation 2.2 (Representations of model components) Let $\mathrm{Ord} \colon \mathcal{C} \to \{1, 2, \ldots, |\mathcal{C}|\}$ be a bijection that specifies an arbitrary order for the categorical data elements in $\mathcal{C}$. We represent each model component from $\mathcal{C}$ as a 0-simplex that is labelled with the name of the model component or the corresponding positive integer under Ord.

Definition 2.3 (Simplicial representation) Let $\mathcal{C}$ be the set of all required components from the collection of models under consideration for comparison. To each model generated by a subset of components from $\mathcal{C}$ we associate a labelled simplicial complex, which we call a *simplicial representation*, where the labelled 0-simplices correspond to components of the model, and the one- and higher-dimensional simplices correspond to the conceptual interconnections between the model components as determined by the model.

The ordering function Ord provides for efficient labelling of simplices. For example, a 2-simplex can be labelled as $\{1, 2, 3\}$, where the 0-simplices are $\{1\}$, $\{2\}$, and $\{3\}$, and the 1-simplices are $\{1, 2\}$, $\{1, 3\}$, and $\{2, 3\}$.

For a particular level of component detail of a given model we associate a simplicial representation consisting of labelled 0-simplices, which represent the model components, along with the one- and higher-dimensional simplices that represent

the conceptual interconnections between the components: the dyadic interactions are 1-simplices, the triadic interactions are 2-simplices, and so forth as required. The labelling of the 0-simplices induces a labelling of the higher-dimensional simplices through their spanning 0-simplices. Since the models for comparison are generated by the same set of components $\mathcal{C}$, a particular labelled simplex representing specific components and interconnections can be identified within different simplicial representations associated with the models.

Our definition of the equivalence of two models is in terms of a formal equivalence between the two corresponding simplicial representations. The formal equivalence is based on partial operations on the relevant set of simplicial representations, which we make explicit here.

Definition 2.4 (Partial operation on a set) Let $\mathcal{X}$ be a set. A *partial operation* on $\mathcal{X}$ is a partial unary function $f \colon \mathcal{X} \rightharpoonup \mathcal{X}$. Denoting the domain of definition of f as $\mathcal{D} \subseteq \mathcal{X}$, we say that f is *invertible* if there exists a partial operation $g \colon \mathcal{X} \rightharpoonup \mathcal{X}$ such that $(g \circ f)|_{\mathcal{D}} = \mathrm{id}_{\mathcal{D}}$ and $(f \circ g)|_{f(\mathcal{D})} = \mathrm{id}_{f(\mathcal{D})}$.

Note that the domain of definition of a partial operation may have cardinality one.

Definition 2.5 (Equivalence of simplicial representations) Let $\mathcal{C}$ be the set of all components that appear in the collection of models under consideration for comparison, and let $\mathcal{S}$ be a collection of simplicial representations corresponding to models generated by subsets from $\mathcal{C}$. Further, let $\mathcal{O}$ be a set of partial operations on $\mathcal{S}$. Two simplicial representations $K, L \in \mathcal{S}$ are *equivalent* with respect to $\mathcal{O}$ if and only if there exists a (possibly empty) sequence of invertible partial operations $(f_i)_{i=0}^{n}$ in $\mathcal{O}$ such that $f_n \circ \cdots \circ f_1 \circ f_0(K) = L$.

Remark 2.6 Note that the equivalence of simplicial representations is always relative to the set $\mathcal{O}$ of partial operations on $\mathcal{S}$. The choice of the set $\mathcal{O}$ determines which model components will be considered as conceptually equivalent. We employ partial operations on $\mathcal{S}$ that we call *admissible*, meaning that they only relate model components that we consider to be conceptually equivalent. We discuss this further below when we define the operations that we employ.

Remark 2.7 It follows from our notion of equivalence of simplicial representations in Definition 2.5 that the set $\{(K, L) \in \mathcal{S} \times \mathcal{S} \mid K \text{ and } L \text{ are equivalent}\}$ is an equivalence relation on $\mathcal{S}$.

Remark 2.8 Note that the equivalence of two simplicial representations $K, L \in \mathcal{S}$ can also be defined as follows: K and L are *equivalent* if and only if there exist two (possibly empty) sequences of invertible partial operations $(f_i)_{i=0}^{m}$ and $(g_i)_{i=0}^{n}$ on $\mathcal{S}$ such that $f_m \circ \cdots \circ f_1 \circ f_0(K) = g_n \circ \cdots \circ g_1 \circ g_0(L)$. In this work we use the fact that this form of equivalence of simplicial representations implies the form in Definition 2.5.

We previously defined five partial operations that we use to establish equivalence of simplicial representations, namely: (Operation 1) Adjacent-vertex identification, (Operation 2) Nonadjacent-vertex identification, (Operation 3) Vertex split, (Operation

4) Inclusion, and (Operation 5) Vertex substitution. These five partial operations are each induced by a map between simplicial complexes, and are all invertible for suitable domains of definition: an adjacent-vertex identification is mutually inverse with a corresponding vertex split; a nonadjacent-vertex identification is mutually inverse with an inclusion; and, a vertex substitution is mutually inverse with another vertex substitution. From the perspective of Remark 2.8 we will only explicitly consider Operations 1, 2, and 5, while Operations 3 and 4 will be implicit.

For the equivalence of simplicial representations to be conceptually meaningful we need to ensure that the employed partial operations are themselves conceptually meaningful, or *admissible*, meaning that the operations preserve the conceptual representation of the general physical system. Importantly, an established equivalence between models is based on the components of the models that we choose to identify as equivalent, and therefore the corresponding partial operations that we consider admissible. The equivalence of two models is therefore determined by the formal requirements in terms of partial operations, along with tight domain-specific constraints.

We now state Operations 1, 2, and 5. Further details are in Vittadello and Stumpf (2021). Recall that two distinct vertices in a simplicial complex are *adjacent* if they belong to the same simplex, and for a vertex u in a simplicial complex K we denote the set of all vertices adjacent to u as $V_K(u)$, noting that $u \notin V_K(u)$. Further, let $\mathcal{C}$ be the set of all components of models under consideration, and let K and L be two simplicial representations with labels from $\mathcal{C}$.

Definition 2.9 (Operation 1: Adjacent-vertex identification) Let u and v be a pair of adjacent vertices in K such that the following all hold:

- $V_K(u) \setminus \{v\} = V_K(v) \setminus \{u\}$.
- For any nonempty subset $W \subseteq V_K(u) \setminus \{v\}$, the vertices $W \cup \{u\}$ span a simplex in K if and only if the vertices $W \cup \{v\}$ span a simplex in K.

A simplicial map $\pi_1 \colon K \to L$ is an *adjacent-vertex identification* if π_1 is surjective, and is injective and label preserving on every vertex except at the pair of vertices u and v which are mapped to a single vertex $c \in L^{(0)}$.

Definition 2.10 (Operation 2: Nonadjacent-vertex identification) Let u and v be a pair of nonadjacent vertices in K such that the following all hold:

- $V_K(u) = V_K(v)$.
- For any nonempty subset $W \subseteq V_K(u)$, the vertices $W \cup \{u\}$ span a simplex in K if and only if the vertices $W \cup \{v\}$ span a simplex in K.

A simplicial map $\pi_2 \colon K \to L$ is a *nonadjacent-vertex identification* if π_2 is surjective, and is injective and label preserving on every vertex except at the pair of vertices u and v that are mapped to a single vertex $c \in L^{(0)}$.

Definition 2.11 (Operation 5: Vertex substitution) A simplicial map $\pi_5 \colon K \to L$ is a *vertex substitution* if π_5 is bijective and preserves all labels except for one whereby the labelled vertex $u \in K^{(0)}$ is mapped to the labelled vertex $c \in L^{(0)}$.

We may extend each partial operation to effect, for example, multiple vertex identifications or substitutions by extending the assumptions in Operations 1, 2, and 5.

Remark 2.12 We note here that our model-comparison methodology is designed to be flexible, so that the partial operations can be chosen based on specific requirements. Different operations may produce different model equivalences, however operations that correspond to very weak conceptual similarity may yield conceptual equivalences that are not scientifically meaningful.

The specific partial operations on simplicial representations that we employ (vertex identifications and substitution) are chosen to represent our notion of conceptual similarity for the models and underlying biological systems that we consider here. Of primary interest for our work are the vertex-identification operations, which identify vertices in a simplicial representation corresponding to conceptually-related model components; the identification of such vertices in a simplicial representation in a rigorous and automatable manner is the subject of Sect. 3. Additional partial operations may be useful for grouping models based on equivalence, however the operations that we consider provide much flexibility while preserving conceptual equivalence.

3 Results and discussion

Here we have two main objectives. The first is to develop an automatable methodology for identifying equivalent components in models. The second is to provide a representation of models as groups, as an alternative to simplicial representations, with which models can be compared in a manner equivalent to that with simplicial representations. We begin with an overview of our approach for formalising the identification of equivalent model components with group actions.

3.1 Overview

The most difficult aspect of establishing the equivalence of two simplicial representations is in determining whether two adjacent or two nonadjacent vertices in one simplicial representation can be identified conceptually with a vertex in the other simplicial representation, either via a vertex identification or vertex split operation. Identifiable vertices in a simplicial representation must be in positions that are conceptually equivalent, and in particular in symmetric positions within the corresponding unlabelled simplicial complex. We can therefore employ group actions on simplicial complexes to greatly simplify and expedite the process of determining the existence of identifiable vertices, while also providing for the process to be automated. The existence of a relevant nontrivial group action then induces a vertex-identification operation.

Given a simplicial representation, our goal is to determine whether any pairs of vertices are in symmetrical positions in the simplicial complex, and so are candidates for a possible vertex-identification operation. We therefore need to determine whether the simplicial complex has any symmetries of a certain class. By a symmetry of a simplicial complex we mean a permutation on the set of vertices of the complex that

acts simplicially, so sends simplices to simplices, and that sends the whole complex to itself. We are interested in a class of symmetries, which we may call *exchange symmetries*, of the complex that exchange two vertices and leave the other vertices fixed, since these symmetries reveal the vertices that are candidates for a possible vertex-identification operation. From a set of exchange symmetries of the complex we then obtain a new simplicial complex in which the exchangeable vertices are identified by a vertex-identification operation; this process can be iterated until a simplicial complex is obtained for which no exchange symmetries exist, and only relabelling of the vertices need be considered by a vertex-substitution operation.

3.2 Identification of equivalent model components through group actions

We begin by establishing the relevant symmetries of simplicial complexes, which are described by a group action. Recall the definition of a permutation of a set:

Definition 3.1 (Permutation, Transposition) A *permutation* of a set X is a bijection from X onto itself, and the set of all permutations of X is the symmetric group on X which we denote by $\mathrm{Sym}(X)$. A *transposition* is a permutation that exchanges two elements of X and leaves the other elements fixed.

Notation 3.2 We write permutations in both function notation and cycle notation, as is convenient, and we compose permutations from right to left. In cycle notation a transposition is a 2-cycle, written $(x\ y)$ for $x, y \in X$.

Remark 3.3 Herein our convention is to view permutations in the active sense.

Formally we refer to a symmetry of a simplicial complex as an automorphism. Recall the following standard definition:

Definition 3.4 (Automorphism of a simplicial complex, Automorphism group) An *automorphism*, or symmetry, of a simplicial complex K is a permutation on the set of vertices $\mathrm{Vert}(K)$ that acts simplicially. The *automorphism group* of K is the set of all automorphisms of K, denoted $\mathrm{Aut}(K)$, which is isomorphic to a subgroup of $\mathrm{Sym}\big(\mathrm{Vert}(K)\big)$.

Herein we denote the set of positive integers from 1 to n as $[n]$.

Notation 3.5 Let π be an automorphism of a simplicial complex K, and denote $\mathrm{Vert}(K) = \{v_i\}_{i=1}^{n}$. If $\sigma := \{v_i\}_{i \in I}$ is a simplex in K for some subset $I \subseteq [n]$, then we denote the action of π on σ by $\pi(\sigma) = \{\pi(v_i)\}_{i \in I}$. We will usually denote vertices as singleton sets of positive integers, in which case $\mathrm{Vert}(K) = \big\{\{a_i\}\big\}_{i=1}^{n}$ where each a_i is a positive integer, and we then denote the action of π on σ by $\pi(\sigma) = \big\{\{\pi(a_i)\}\big\}_{i \in I}$.

The automorphisms of a simplicial complex are described by the action of a group on the complex. Our definition of a (left) group action on a simplicial complex is standard (Bredon 1972, Chapter III, Page 115).

Definition 3.6 (Group action) Let G be a group with identity e, and let X be a set. An *action* of the group G on X is a map $\Theta \colon G \times X \to X$ such that:

1. $\Theta(gh, x) = \Theta(g, \Theta(h, x))$ for all $g, h \in G$ and $x \in X$;
2. $\Theta(e, x) = x$ for all $x \in X$.

If K is a simplicial complex then, setting $X = K$, we also assume that:

3. For each $g \in G$ the map $\theta_g \colon K \to K$, given by $\theta_g(\sigma) = \Theta(g, \sigma)$ for $\sigma \in K$, acts simplicially on K.

We say that each $g \in G$ acts simplicially on K, and each such action of a group G is called a *G-action*.

Remark 3.7 Equivalently, an action of the group G on the set X is a group homomorphism $\Phi \colon G \to \mathrm{Aut}(X)$. In particular, recalling that an automorphism of a simplicial complex K is simplicial, a *simplicial action* of the group G on the simplicial complex K is a group homomorphism $\Phi \colon G \to \mathrm{Aut}(K)$.

Notation 3.8 For an action of the group G on the set X, and for $g \in G$ and $x \in X$, we denote $\Theta(g, x) = \Phi(g)(x)$ by $g \cdot x$.

While conditions of *regularity* (Bredon 1972, Chapter III, Page 116, Definition 1.2) are generally assumed for G-actions, such conditions are neither required nor desirable for our purposes since we are interested in identifying labelled simplices that, upon permutation of the spanning vertices, produce an invariant simplicial representation. Indeed, a *regular* simplicial action of the group G on the simplicial complex K is a simplicial action that satisfies the following condition for the action of each subgroup of G (Bredon 1972, Chapter III, Page 115, Statement (B)): Condition (A) — if $g_1, g_2, \ldots, g_m \in G$, and $\{v_1, v_2, \ldots, v_m\}$ and $\{g_1 \cdot v_1, g_2 \cdot v_2, \ldots, g_m \cdot v_m\}$ are both simplices of K, then there exists an element $g \in G$ such that $g \cdot v_i = g_i \cdot v_i$ for all $i \in [m]$. Condition (A) implies the following Condition (B) (Bredon 1972, Chapter III, Page 116, Statement (A$'$) and the following two paragraphs): Condition (B) — if $g \in G$, v is a vertex in K, and v and $g \cdot v$ belong to the same simplex, then $v = g \cdot v$. Condition (B) illustrates that regular actions fix vertices that are sent to the same simplex. To the contrary, we require nonregular actions that allow for the permutation of the vertices of a given simplex, as illustrated in the following simple example.

Example 3.9 Let K be the 1-simplex $\big\{ \{1\}, \{2\}, \{1, 2\} \big\}$, regarded as a simplicial representation. Since the vertices $\{1\}$ and $\{2\}$ are adjacent and in symmetrical positions in the complex, they are conceptually equivalent so may be identified by an adjacent-vertex identification. While we can directly observe from K that the two vertices are conceptually equivalent, this is very difficult for more complicated simplicial representations, and the use of group actions simplifies the process. In this case, let $G = \big\langle (1\ 2) \big\rangle$ be the permutation group on the set $\{1, 2\}$ generated by the transposition $(1\ 2)$, that is, $G = \big\{ e, (1\ 2) \big\}$. Then G acts simplicially on K by permuting the vertices of K (we discuss such actions in detail below). Now, $\{1, 2\}$ and $\big\{ e \cdot 1, (12) \cdot 2 \big\} = \{1\}$ are both simplices of K, however there is no $g \in G$ such that $g \cdot \{1, 2\} = \{1\}$. So G does not satisfy Condition (A), and hence is not a regular action. Such a group action is, however, a symmetry of interest in our work here, since the permutation of the vertices of the simplex K results in an invariant complex.

To reveal any conceptually-equivalent vertices in a simplicial representation we require a specific class of simplicial automorphisms, which we call exchange automorphisms (see (Papadopoulos 2012, Section 3.2, Page 322, Definition 3.20)).

Definition 3.10 (Exchange automorphism, Exchangeable vertices) Let K be a simplicial complex, and let $u, v \in \mathrm{Vert}(K)$. An *exchange automorphism* for the vertices u and v is an automorphism $\phi \in \mathrm{Aut}(K)$ such that $\phi(u) = v$, $\phi(v) = u$, and $\phi(w) = w$ for all $w \in \mathrm{Vert}(K) \setminus \{u, v\}$. If an exchange automorphism exists for vertices u and v then we say that u and v are *exchangeable* in K.

Note that if K is a simplicial complex and $u, v \in \mathrm{Vert}(K)$ are exchangeable then the associated exchange automorphism on K is unique.

Remark 3.11 Let K be a simplicial complex. The binary relation $\{ (u, v) \mid u, v \in \mathrm{Vert}(K) \text{ are exchangeable} \}$ over the set of exchangeable vertices in K is an equivalence relation. Reflexivity follows from the identity automorphism on K, and symmetry follows from the definition of exchangeable vertices. For transitivity we observe that if $\phi \in \mathrm{Aut}(K)$ is an exchange automorphism for the vertices u and v, and $\psi \in \mathrm{Aut}(K)$ is an exchange automorphism for the vertices v and w, then $\phi \circ \psi \circ \phi \in \mathrm{Aut}(K)$ is an exchange automorphism for the vertices u and w.

We employ group actions of a particular class:

Definition 3.12 (Group action by exchangeable vertices) A *group action by exchangeable vertices* on a simplicial complex K consists of an action of a group G on K such that $G = \langle M \rangle$ where:

- $M \subseteq \mathrm{Sym}\big(\mathrm{Vert}(K)\big)$ is a set of transpositions $(u\ v)$ where u and v are exchangeable in K; and,
- the action of each $(u\ v) \in M$ is the exchange automorphism on K that exchanges u and v.

Once identifiable vertices are found in a simplicial representation through a group action, we can obtain a new simplicial complex, called the orbit space, with these vertices identified as single vertices. The definition of the orbit space of an action of a group on a simplicial complex follows (Bredon 1972, Chapter III, Page 117, Paragraph 2).

Definition 3.13 (Orbit, Orbit space) Let K be a simplicial complex, and suppose that the group G acts on K. The *orbit*, or *G-orbit*, of an element $\sigma \in K$ is the set $G \cdot \sigma := \{ g \cdot \sigma \mid g \in G \}$. The *orbit space*, or *quotient*, of an action of a group G on K, denoted K/G, is the set consisting of the orbits $\mathrm{Vert}(K/G) := \{ G \cdot v \mid v \in \mathrm{Vert}(K) \}$, along with the finite subsets $\{G \cdot v_i\}_{i=1}^n$ of $\mathrm{Vert}(K/G)$ for which $\{u_i\}_{i=1}^n$ spans a simplex in K, where each u_i is a representative of $G \cdot v_i$, and the existence of such a simplex in K is not required for all systems of representatives of the orbits $G \cdot v_i$.

Example 3.14 Let K be the simplicial 2-complex $\big\{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\big\}$, which is geometrically a filled triangle. Let $G := \big\langle (1\ 2) \big\rangle \cong \mathbb{Z}/2\mathbb{Z}$ be the cyclic subgroup of S_3, the symmetric group of degree 3, generated by the transposition
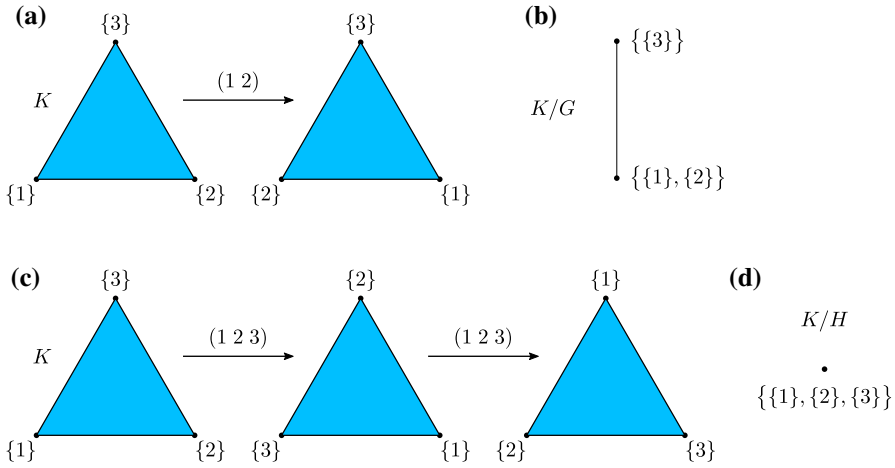
Fig. 1 Geometric illustration of the action of the group G on the simplicial complex K: (a) The action of the transposition (1 2) on K; (b) The orbit space K/G is the 1-simplex with vertices $\{\{3\}\}$ and $\{\{1\}, \{2\}\}$. Geometric illustration of the action of the group H on the simplicial complex K: (c) The action of the permutation (1 2 3) on K; (d) The orbit space K/H is the 0-simplex $\{\{1\}, \{2\}, \{3\}\}$

(1 2). Then $g \cdot K = K$ for all $g \in G$. The action of G on K is illustrated geometrically in Fig. 1a, where (1 2) is a reflection. We have $G \cdot \{1\} = G \cdot \{2\} = \{\{1\}, \{2\}\}$ and $G \cdot \{3\} = \{\{3\}\}$, hence $\mathrm{Vert}(K/G) = \{G \cdot \{1\}, G \cdot \{3\}\}$. Since $\{1, 3\}$, or alternatively $\{2, 3\}$, is a simplex in K, $\{G \cdot \{1\}, G \cdot \{3\}\}$ is a simplex in K/G. Therefore $K/G = \{G \cdot \{1\}, G \cdot \{3\}, \{G \cdot \{1\}, G \cdot \{3\}\}\}$ is a 1-simplex, as illustrated in Fig. 1b, resulting from the identification of the symmetric vertices $\{1\}$ and $\{2\}$ in K.

Now let $H := \langle (1\ 2\ 3) \rangle \cong \mathbb{Z}/3\mathbb{Z}$ be the cyclic subgroup of S_3 generated by the cyclic permutation (1 2 3). Then $h \cdot K = K$ for all $h \in H$. The action of H on K is illustrated geometrically in Fig. 1c, where (1 2 3) is a counter-clockwise rotation. We have $H \cdot \{1\} = H \cdot \{2\} = H \cdot \{3\} = \{\{1\}, \{2\}, \{3\}\}$, and it follows that $K/H = \{H \cdot \{1\}\}$ is a 0-simplex, as illustrated in Fig. 1d, resulting from the identification of the three vertices in K.

It is a standard observation that the orbit space K/G is a simplicial complex (Bredon 1972, Chapter III, Page 117, Paragraph 2), however we provide a proof for completeness.

Proposition 3.15 *Let K be a simplicial complex, and let G be a group action on K. Then the orbit space K/G is a simplicial complex with vertex set* $\mathrm{Vert}(K/G)$.

Proof By definition, K/G consists of the set of vertices $\mathrm{Vert}(K/G) := \{ G \cdot v \mid v \in \mathrm{Vert}(K) \}$ along with the subsets $\{G \cdot v_i\}_{i \in I}$ of $\mathrm{Vert}(K/G)$, for index set I, for which there exists $(u_i)_{i \in I} \in \prod_{i \in I} G \cdot v_i$ such that $\{u_i\}_{i \in I}$ spans a simplex in K.

To show that K/G is a simplicial complex we need to show that for each subset $\{G \cdot v_i\}_{i \in I}$ of $\mathrm{Vert}(K/G)$ that spans a simplex in K/G, every nonempty subset of $\{G \cdot v_i\}_{i \in I}$ also spans a simplex in K/G. So let $\{G \cdot v_j\}_{j \in J}$ be a subset of $\{G \cdot v_i\}_{i \in I} \in$

K/G, where $J \subseteq I$ is nonempty. Since $\{G \cdot v_i\}_{i \in I} \in K/G$, there exists a simplex $\{u_i\}_{i \in I} \in K$ with $(u_i)_{i \in I} \in \prod_{i \in I} G \cdot v_i$. Then $\{u_j\}_{j \in J}$ is a subsimplex of $\{u_i\}_{i \in I}$ in K, with $(u_j)_{j \in J} \in \prod_{j \in J} G \cdot v_j$, hence $\{G \cdot v_j\}_{j \in J}$ is a simplex in K/G, as required.   □

We now show that there is a canonical map from a simplicial complex onto the orbit space corresponding to a G-action. This simplicial map induces a vertex-identification operation between the two simplicial complexes.

Proposition 3.16 *Let K be a simplicial complex, and let G be a group action on K. Then the canonical vertex map $v \mapsto G \cdot v$ extends to a surjective simplicial map $p \colon K \to K/G$.*

Proof Let $\psi \colon \mathrm{Vert}(K) \to \mathrm{Vert}(K/G)$ be the canonical vertex map such that $v \mapsto G \cdot v$. Define the map $p \colon K \to K/G$ by $p(\{v_i\}_{i \in I}) := \{\psi(v_i)\}_{i \in I}$ for a simplex $\{v_i\}_{i \in I} \in K$, noting that the $\psi(v_i)$ may be equal for distinct indices in I. Then p is simplicial since, letting $J \subseteq I$ be a subset of smallest cardinality such that $\{\psi(v_j)\}_{j \in J} = \{\psi(v_i)\}_{i \in I}$, we have that $\{v_j\}_{j \in J}$ is a simplex in K and hence, by the definition of K/G, $\{\psi(v_j)\}_{j \in J}$ is a simplex in K/G. Surjectivity of p follows from the surjectivity of ψ, and p extends ψ since $p|_{\mathrm{Vert}(K)} = \psi$.   □

Notation 3.17 We refer to the canonical map $p \colon K \to K/G$ in Proposition 3.16 as the *projection map*.

When we identify conceptually-equivalent vertices in a simplicial representation K to give an orbit space, we would expect that the orbit space is isomorphic to a simplicial subcomplex of K as a consequence of the underlying symmetry identified with the group action. We now confirm that this is indeed the case, beginning with two lemmas.

Lemma 3.18 *Let G be a group acting on a simplicial complex K by exchangeable vertices, and let $u \in \mathrm{Vert}(K)$. If $v, w \in G \cdot u$ then the transposition $(v\ w)$ is in G. In particular, v and w are exchangeable in K.*

Proof Since $v, w \in G \cdot u$ there exists $g \in G$ such that $g \cdot v = w$, where $g = t_n t_{n-1} \cdots t_2 t_1$ is a product of transpositions $t_i \in G$ for $i \in [n]$. Without loss of generality we may assume that there is no proper subsequence of the sequence of transpositions $(t_i)_{i=1}^n$ whose product sends v to w. In particular: v is an element of t_1 only; w is an element of t_n only; and the transpositions are mutually disjoint unless they are successive in the sequence, whereby their intersection is a singleton. We then have $t_1 = (v\ x_1)$, $t_{i+1} = (x_i\ x_{i+1})$ for $i \in [n-2]$, and $t_n = (x_{n-1}\ w)$, for some $\{x_i\}_{i=1}^{n-1} \subseteq \mathrm{Vert}(K)$. It follows that $(v\ w) = t_1 t_2 \cdots t_{n-1} t_n t_{n-1} \cdots t_2 t_1 \in G$.   □

Lemma 3.19 *Let G be a group acting on a simplicial complex K by exchangeable vertices. Suppose that $\{G \cdot v_i\}_{i \in I}$ spans a simplex in K/G, for some index set I, and $\{u_i\}_{i \in I}$ spans a simplex in K with $(u_i)_{i \in I} \in \prod_{i \in I} G \cdot v_i$. Then for each $(w_i)_{i \in I} \in \prod_{i \in I} G \cdot v_i$ we have that $\{w_i\}_{i \in I}$ spans a simplex in K, and there is a $g \in G$ such that, for each $i \in I$, $g \cdot u_i = w_i$, $g \cdot w_i = u_i$, and $g \cdot v = v$ for all $v \in \mathrm{Vert}(K) \setminus (\{u_i\}_{i \in I} \cup \{w_i\}_{i \in I})$.*

Proof Suppose $(w_i)_{i \in I} \in \prod_{i \in I} G \cdot v_i$. It follows from Lemma 3.18 that for each $i \in I$ the transposition $t_i := (u_i \ w_i)$ is in G. So define $g \in G$ as the product $g := \prod_{i \in I} t_i$, where the order in which the t_i are multiplied is unimportant since these group elements mutually commute: indeed, two vertices from distinct G-orbits are not exchangeable. Then $g \in G$ is the required group element, and it follows that the action of g maps the simplex spanned by $\{u_i\}_{i \in I}$ to the simplex spanned by $\{w_i\}_{i \in I}$. $\qquad\square$

Definition 3.20 (Fundamental domain) Let K be a simplicial complex with an action of the group G on K. A *fundamental domain* for the action of G on K is a full subcomplex L of K that intersects each vertex G-orbit exactly once.

Note that a fundamental domain is not necessarily unique, and the group action transforms between the different fundamental domains of the complex. We now show that a fundamental domain exists for a group action on a simplicial complex by exchangeable vertices, and that the corresponding orbit space is isomorphic to the fundamental domain. From this result one infers that all fundamental domains are isomorphic for a given group action on a simplicial complex by exchangeable vertices.

Theorem 3.21 *Let G be a group acting on a simplicial complex K by exchangeable vertices. Then a fundamental domain exists for the action of G on K, and the orbit space K/G is isomorphic to the fundamental domain.*

Proof We first show that a fundamental domain exists for G on K. Let $\mathrm{Vert}(K/G) := \{\, G \cdot v_i \,\}_{i \in I}$ be the set of distinct vertex orbits where I is the index set, choose $(w_i)_{i \in I} \in \prod_{i \in I} G \cdot v_i$, and let L be the full subcomplex of K with vertex set $\{w_i\}_{i \in I}$. Then L is a fundamental domain.

It remains to show that K/G is isomorphic to L. Let $\lambda \colon K/G \to L$ be the map such that if σ is a simplex in K/G spanned by the vertices $\{\, G \cdot v_j \,\}_{j \in J}$, where $J \subseteq I$, then $\lambda(\sigma)$ is the simplex in L spanned by the vertices $\{w_j\}_{j \in J}$. The map λ is well defined, since if the simplex $\sigma \in K/G$ is spanned by $\{\, G \cdot v_j \,\}_{j \in J}$ then there exists a simplex in K spanned by a set of vertices $\{u_j\}_{j \in J}$ with $(u_j)_{j \in J} \in \prod_{j \in J} G \cdot v_j$, so by Lemma 3.19 the vertices $\{w_j\}_{j \in J}$ span a simplex in K, which is also in the full subcomplex L. The map λ is then a simplicial bijection, and hence a simplicial isomorphism. $\qquad\square$

Since the orbit space may have exchangeable vertices that were either not accounted for or not present in the original complex, we can obtain the orbit space of the orbit space, and so on for a finite number of steps until an orbit space is obtained which has no further exchangeable vertices. To see this more clearly, we now show how a group action on a simplicial complex by exchangeable vertices relates to the full group action on the corresponding orbit space by exchangeable vertices. Note that we say that a group action on a simplicial complex by exchangeable vertices is a *full group action* when the group action gives all possible exchange automorphisms of the complex. We first require a lemma.

Lemma 3.22 *Let G be a group action on a simplicial complex K by exchangeable vertices, and let t be a transposition, not necessarily in G, from the full group action on K by exchangeable vertices that acts to exchange the vertices $u, w \in \mathrm{Vert}(K)$.*

Further, let H be the full group action on K/G by exchangeable vertices, and let $\{G \cdot v_i\}_{i \in I}$ be the set of distinct orbits in K/G which partition $\mathrm{Vert}(K)$. Then there exists a transposition $\bar{t} \in H$ that acts to exchange the vertices $G \cdot v_j$, $G \cdot v_k \in \mathrm{Vert}(K/G)$ where $j, k \in I$, $u \in G \cdot v_j$, and $w \in G \cdot v_k$. Further, $\bar{t}$ is the identity transposition if and only if $t \in G$.

Proof Suppose that $t \in \mathrm{Sym}\big(\mathrm{Vert}(K)\big)$ transposes the vertices $u, w \in \mathrm{Vert}(K)$, and that $(u, w) \in G \cdot v_j \times G \cdot v_k$ for some $j, k \in I$. The permutation π_t on the vertices $\{G \cdot v_i\}_{i \in I}$ of K/G which transposes $G \cdot v_j$ and $G \cdot v_k$ and fixes all other vertices induces a map $\phi_t \colon K/G \to K/G$ which sends a simplex $\tau \in K/G$ spanned by $\{G \cdot v_i\}_{i \in J}$ for some $J \subseteq I$ to the simplex $\phi_t(\tau)$ spanned by $\big\{\pi_t(G \cdot v_i)\big\}_{i \in J}$ where $G \cdot v_j$ and $G \cdot v_k$ are exchanged, if required.

To show that ϕ_t is well defined we need to establish that $\phi_t(\tau)$ is a simplex in K/G. Since $\{G \cdot v_i\}_{i \in J}$ spans the simplex $\tau \in K/G$, there exists $(u_i)_{i \in J} \in \prod_{i \in J} G \cdot v_i$ such that $\{u_i\}_{i \in J}$ spans a simplex $\sigma \in K$ by the definition of K/G. By Lemma 3.19 we may assume without loss of generality that $u_j = u$ if $j \in J$ and $u_k = w$ if $k \in J$. Then the image of σ under the action of t is the simplex $t \cdot \sigma \in K$ spanned by $\{t \cdot u_i\}_{i \in J}$ where u and w are exchanged, if required. Now, if $i \neq j$ or k then $t \cdot u_i = u_i \in G \cdot v_i = \pi_t(G \cdot v_i)$, if $i = j$ then $t \cdot u_i = t \cdot u_j = u_k \in \pi_t(G \cdot v_j)$, and if $i = k$ then $t \cdot u_i = t \cdot u_k = u_j \in \pi_t(G \cdot v_k)$. It follows that the simplex $t \cdot \sigma \in K$ satisfies $(t \cdot u_i)_{i \in J} \in \prod_{i \in J} \pi_t(G \cdot v_i)$, hence $\phi_t(\tau)$ is a simplex in K/G. So ϕ_t is a well-defined simplicial map.

Since ϕ_t is self-inverse it is bijective and therefore an automorphism. It follows that ϕ_t is an exchange automorphism that exchanges the vertices $G \cdot v_j$ and $G \cdot v_k$, so the transposition $\bar{t} := (G \cdot v_j \ \ G \cdot v_k)$ is in H.

Finally, $t \in G$ if and only if u and w are in the same G-orbit if and only if $G \cdot v_j = G \cdot v_k$ if and only if $\bar{t}$ is the identity transposition, where the first equivalence uses Lemma 3.18. $\qquad\square$

Proposition 3.23 *Let G and F be two group actions on a simplicial complex K by exchangeable vertices, and let H be the full group action on K/G by exchangeable vertices. Then there is a group homomorphism $\alpha \colon F \to H$ such that $\alpha(t_n \cdots t_2 t_1) = \bar{t}_n \cdots \bar{t}_2 \bar{t}_1$ for each element $t_n \cdots t_2 t_1 \in F$, where the t_i for $i \in [n]$ are generating transpositions in F, and $\ker(\alpha) = F \cap G$.*

Proof We first show that $\alpha \colon F \to H$ is a well-defined map. Let $f \in F$, and suppose that $f = t_m \cdots t_2 t_1$ where the t_i, for $i \in [m]$, are generating transpositions in F. By Lemma 3.22, each $\bar{t}_i$ for $i \in [m]$ is a transposition in H, so $\bar{f} := \bar{t}_m \cdots \bar{t}_2 \bar{t}_1 \in H$. Now, if we also have that $f = s_n \cdots s_2 s_1$ where the s_i, for $i \in [n]$, are generating transpositions in F, then the corresponding permutations of $\mathrm{Vert}(K/G)$, namely $\bar{t}_m \cdots \bar{t}_2 \bar{t}_1$ and $\bar{s}_n \cdots \bar{s}_2 \bar{s}_1$, are equal. It follows that α is well defined.

To show that α is a homomorphism, let $t_m \cdots t_2 t_1$ and $s_n \cdots s_2 s_1$ be two permutations in F, in terms of the generating transpositions. Then $\alpha\big((t_m \cdots t_2 t_1)(s_n \cdots s_2 s_1)\big) = (\bar{t}_m \cdots \bar{t}_2 \bar{t}_1)(\bar{s}_n \cdots \bar{s}_2 \bar{s}_1) = \alpha(t_m \cdots t_2 t_1)\alpha(s_n \cdots s_2 s_1)$, and the result follows.

It remains to show that the kernel of α is equal to $F \cap G$. For this, let $\{G \cdot v_i\}_{i \in I}$ be the set of distinct vertices in K/G, which partition $\mathrm{Vert}(K)$. Suppose that $f \in \ker(\alpha)$, so that $\alpha(f) = \bar{f}$ is the identity in H. Then $\bar{f}$ fixes each vertex $G \cdot v_i$ of K/G, so the

definition of $\overline{f}$ implies that $f \cdot (G \cdot v_i) \subseteq G \cdot v_i$ for $i \in I$. We can write f as a product of disjoint permutations, $f = \prod_{i \in I} p_i$, where each such permutation p_i permutes only elements in $G \cdot v_i$ with the same action as f on $G \cdot v_i$, and p_i fixes all elements in $\text{Vert}(K) \setminus G \cdot v_i$. Then each p_i can be expressed as a product of transpositions in G, hence $f \in G$. It follows that $\ker(\alpha) \subseteq F \cap G$. Conversely, if $f \in F \cap G$ then, since $f \in G, \alpha(f) = \overline{f} \in H$ fixes the G-orbits of $\text{Vert}(K)$, which are the vertices of K/G, so $\alpha(f)$ is the identity in H, hence $f \in \ker(\alpha)$. Therefore $F \cap G \subseteq \ker(\alpha)$. □

In Sect. 3.3 we employ the theory developed here to provide a methodology for determining model equivalence.

3.3 Methodology for determining model equivalence

To compare the simplicial representations corresponding to two models, with respect to equivalence, we can compare the two corresponding orbit spaces that have no further exchangeable vertices, which we refer to as the *final* orbit spaces. These orbit spaces, which may be smaller than the original simplicial complexes, allow for the identification of subsets of equivalent components within each model, and in turn the equivalence of components between the two models.

Employing the theory in Sect. 3.2 for identifying equivalent components within simplicial representations, we now provide a general methodology for determining model equivalence in terms of Operations 1, 2, and 5. There is not necessarily a unique order with which Operations 1, 2, and 5 should be applied to a simplicial representation. Given two simplicial representations, however, the easiest approach is to first reduce the simplicial representations by applying vertex identifications (Operations 1 and 2) when possible, and finally relabelling one of the complexes by a vertex substitution (Operation 5). We will assume that all possible vertex identifications for a given simplicial representation are applied simultaneously. We could apply a proper subset of the possible vertex identifications to the simplicial representation, and then apply the remaining identifications to the corresponding orbit space (see Proposition 3.23), however applying them all at once minimises the number of required simplicial maps.

The following steps describe the process of finding model equivalences using simplicial representations. For the purpose of automation on computer we can relabel the vertices of the simplicial representations, which correspond to model components, with positive integers so that the simplices of the simplicial representations are subsets of these positive integers. Assume that K and L are two simplicial representations corresponding to two models.

Step 1. **Find all pairs of exchangeable vertices of** K **and** L: We describe the process for K, and the process for L is similar. Two vertices u and v in $\text{Vert}(K)$ are *exchangeable* if exchanging the labels of the two vertices leaves K unchanged (Definition 3.10). All exchangeable vertices of K can be found by exchanging all pairs of vertices of K in succession, however this process can be made more efficient in a number of ways: for example, exchangeable vertices must have the same number of adjacencies, so two vertices with different numbers of adjacencies are not exchangeable. Given a pair of exchangeable vertices u and v in $\text{Vert}(K)$ there is a corresponding transposition $(u\ v) \in \text{Sym}\left(\text{Vert}(K)\right)$,

and the set of all such transpositions corresponding to exchangeable vertices in K generates a subgroup G_1 of Sym $\big(\mathrm{Vert}(K)\big)$. We then have an action of G_1 on K by automorphisms, from which we construct the orbit space K/G_1 with vertices the G_1-orbits of the vertices in K (Definition 3.13). If K has no exchangeable vertices then G_1 is the trivial subgroup which acts as the identity automorphism of K, so we can then take K/G_1 to be K since these complexes are isomorphic.

A transposition $(u\ v) \in G_1$ acts as an exchange automorphism of K that exchanges the vertices $u,\ v \in K$, and induces a corresponding adjacent-/nonadjacent-vertex identification operation in terms of a surjective simplicial map $p\colon K \to K/\langle(u\ v)\rangle$ which sends both u and v to a single vertex in $K/\langle(u\ v)\rangle$ and fixes all other vertices in K. Note that p is the projection map (Proposition 3.16). Applying all permutations in G_1 simultaneously corresponds to G_1 acting on K, and therefore induces a finite number of vertex-identification operations in terms of the projection map $p_1\colon K \to K/G_1$.

Step 2. **Find the sequences of orbit spaces corresponding to K and L**: Again, we describe the process for K, and it is similar for L. Finding the exchangeable vertices of K/G_1 follows a process analogous to that for K in Step 1, except the vertices of K/G_1 are subsets of vertices of K. Given a pair of exchangeable vertices u and v in $\mathrm{Vert}(K/G_1)$ there is a corresponding transposition $(u\ v) \in \mathrm{Sym}\big(\mathrm{Vert}(K/G_1)\big)$, and the set of all such transpositions corresponding to exchangeable vertices in K/G_1 generates a subgroup G_2 of Sym $\big(\mathrm{Vert}(K/G_1)\big)$. We then have an action of G_2 on K/G_1 by automorphisms, from which we construct the orbit space $(K/G_1)/G_2$ with vertices the G_2-orbits of the vertices in K/G_1. The action of G_2 on K/G_1 induces a finite number of vertex-identification operations in terms of the projection map $p_2\colon K/G_1 \to (K/G_1)/G_2$.

Continuing in this manner until we obtain an orbit space with no exchangeable vertices, namely the final orbit space, results in the following sequence of simplicial complexes and projection maps p_i representing vertex identifications,

$$
\begin{aligned}
K &\xrightarrow{\ p_1\ } K/G_1 \\
&\xrightarrow{\ p_2\ } (K/G_1)/G_2 \xrightarrow{\ p_3\ } \cdots \xrightarrow{\ p_m\ } \Big(\cdots\big((K/G_1)/G_2\big)\cdots\Big)/G_m,
\end{aligned} \quad (1)
$$

where the final orbit space in the sequence, which we denote by $\widehat{K}$, has no exchangeable vertices. Similarly, for the simplicial representation L we have a sequence of subgroups H_i generated by transpositions giving the following sequence of simplicial complexes and simplicial maps q_i representing vertex identifications,

$$
\begin{aligned}
L &\xrightarrow{\ q_1\ } L/H_1 \\
&\xrightarrow{\ q_2\ } (L/H_1)/H_2 \xrightarrow{\ q_3\ } \cdots \xrightarrow{\ q_n\ } \Big(\cdots\big((L/H_1)/H_2\big)\cdots\Big)/H_n,
\end{aligned} \quad (2)
$$

where the final orbit space in the sequence, which we denote by $\widehat{L}$, has no exchangeable vertices.

Note that if any complex from the sequence in Eq. (1) is isomorphic to a complex from the sequence in Eq. (2), as unlabelled simplicial complexes, then $\widehat{K}$ and $\widehat{L}$ are isomorphic as unlabelled simplicial complexes. Therefore, it suffices to determine whether or not $\widehat{K}$ and $\widehat{L}$ are isomorphic, which are the simplest complexes in their respective sequences.

With regard to the sequences in Eqs. (1) and (2) note that the number of vertices in an orbit space is reduced, relative to the previous complex in the sequence, by the number of distinct and nontrivial transpositions in the associated group action. This observation provides a measure of the reduction in the number of possible isomorphisms between the orbit spaces of each simplicial complex.

Step 3. Determine all isomorphisms of $\widehat{K}$ and $\widehat{L}$: An isomorphism between $\widehat{K}$ and $\widehat{L}$ is a bijective simplicial map from $\widehat{K}$ onto $\widehat{L}$, that is, a bijective vertex map from $\mathrm{Vert}(\widehat{K})$ onto $\mathrm{Vert}(\widehat{L})$ that sends a simplex in $\widehat{K}$ to a simplex in $\widehat{L}$. We begin by finding all bijective vertex maps from $\mathrm{Vert}(\widehat{K})$ onto $\mathrm{Vert}(\widehat{L})$, noting that there may be none. If there are multiple bijective vertex maps then we may reduce the number of these maps by considering properties that must be preserved by a simplicial isomorphism, such as the dimension of simplices and the number of edge adjacencies of vertices. Further, if $\widehat{K}$ and $\widehat{L}$ each have a vertex label containing the same model component then we would generally ensure that the corresponding vertices must be identified by any bijective vertex map, based on the requirement for conceptual equivalence. We could omit the preservation of model components for the bijective vertex maps, however this will likely only increase the number of bijective vertex maps to consider without yielding any new meaningful equivalences between the model concepts. It should not be the case that a vertex label in $\widehat{K}$ has model components in common with two distinct vertices in $\widehat{L}$, as this would indicate a conceptual inconsistency with at least one of the simplicial representations. Given a bijective vertex map, which we can describe by a correspondence between the positive-integer labels of the vertices in $\widehat{K}$ and $\widehat{L}$, we then try to extend it to a bijection from $\widehat{K}$ onto $\widehat{L}$ by relabelling $\widehat{K}$ with the corresponding vertex labels of $\widehat{L}$ in accordance with the bijective vertex maps. We then check whether the relabelled complex $\widehat{K}$ is the same as $\widehat{L}$, and if so then the bijective vertex map extends canonically to a simplicial isomorphism, so that $\widehat{K}$ and $\widehat{L}$ are isomorphic. If no simplicial isomorphism exists between $\widehat{K}$ and $\widehat{L}$ then they are not isomorphic as complexes.

Step 4. Equivalence of K and L: If K and L are equivalent then $\widehat{K}$ and $\widehat{L}$ are isomorphic; therefore, if $\widehat{K}$ and $\widehat{L}$ are not isomorphic then K and L are not equivalent. If $\widehat{K}$ and $\widehat{L}$ are isomorphic then for each simplicial isomorphism we can consider, for each pair of vertices identified between $\widehat{K}$ and $\widehat{L}$ via the isomorphism, whether the set of model components associated with the vertex in $\widehat{K}$ is conceptually related to the set of model components associated with the vertex in $\widehat{L}$. Since there may be more than one isomorphism between $\widehat{K}$ and $\widehat{L}$ we consider the vertex identifications for each such isomorphism.

Importantly, an isomorphism between $\widehat{K}$ and $\widehat{L}$ can reveal novel conceptual relations between K and L, and therefore the corresponding models.

If there is an isomorphism between $\widehat{K}$ and $\widehat{L}$ that provides a conceptually meaningful relationship between the concepts of the two models then we can use a vertex-substitution operation induced by the isomorphism to transform $\widehat{K}$ into $\widehat{L}$, and then we conclude that K and L, and hence the corresponding models, are equivalent. Otherwise, if no isomorphism between $\widehat{K}$ and $\widehat{L}$ provides a conceptually meaningful relationship between the concepts of the two models then we conclude that K and L, and hence the corresponding models, are not equivalent.

Step 5. **Restricting conceptual equivalence:** A simplicial isomorphism between the final orbit spaces $\widehat{K}$ and $\widehat{L}$ provides for the maximum conceptual equivalence between the components of the two models. It may happen, however, that we regard a certain equivalence for a model, corresponding to a vertex identification, as invalid. In this case, we don't consider $\widehat{K}$ and $\widehat{L}$ as equivalent. There may, however, be equivalent simplicial subcomplexes of K and L with respect to a restricted set of vertex identifications corresponding to the valid conceptual equivalences of model components. To determine whether equivalent subcomplexes exist in this restricted sense, we can work backwards from $\widehat{K}$ and $\widehat{L}$ to remove all of the invalid vertex identifications, while retaining the isomorphic correspondence. If we can remove corresponding vertex identifications stepwise from both $\widehat{K}$ and $\widehat{L}$, where each vertex identification is either adjacent for both complexes or nonadjacent for both complexes, and arrive at isomorphic complexes with only valid equivalences between model components then K and L are equivalent, and otherwise not.

Remark 3.24 Here we briefly discuss some computational aspects of our model comparison methodology.

1. Orbit spaces: Constructing the orbit space of a simplicial complex K requires first determining the exchange automorphisms, which only involves exchanging pairs of vertices in K and checking that K is unchanged. Once all relevant exchange automorphisms are found, the orbit space is constructed from the vertex orbits and the known simplices in K.

2. Isomorphism of simplicial complexes: We need to determine whether the final orbit spaces are isomorphic. In general, checking isomorphism of simplicial complexes is computationally intensive, however our simplicial complexes are labelled (with sets of model components, or corresponding positive integers) and we are usually interested in isomorphisms that preserve model components. This restriction reduces the number of possible isomorphisms to consider.

3. Automation: Our methodology for model comparison by equivalence consists of three main parts. Given a model, the first part requires the construction of the simplicial representation. This requires a translation of the model into its underlying concepts and their interrelations. In general this is very difficult to automate, as it requires the ability to identify the model concepts and interrelations for a model with any formalism, as well as equivalences between model concepts. The latter is also dependent on context and individual requirements.

The second part of our methodology, which we discuss in this article, is completely automatable up to and including the determination of whether the final orbit spaces are isomorphic. This provides for possible equivalences between model concepts. The second part is the most challenging of all three parts, since establishing whether vertices in a simplicial representation are in symmetrical positions is difficult for even small models.

The third part requires determining whether the possible equivalences between model concepts are valid, which returns to part one, and is difficult to automate in general.

3.4 Example

In this example we apply our model-comparison methodology to the two main categories of models for developmental pattern formation, namely Turing-pattern (TP) models and positional-information (PI) models. We have described models from these two categories previously (Vittadello and Stumpf 2021): four TP models, namely activator-inhibitor, substrate depletion, inhibition of an inhibition, and modulation; five PI models, namely linear gradient, synthesis-diffusion-degradation (SDD), opposing gradients, annihilation, and induction-contraction (active modulation). We also discuss these models in the Appendix (Sect. 5). Interestingly, we found that the TP activator-inhibitor model is equivalent to the PI annihilation model from a significant conceptual perspective. This finding was obtained by visual inspection of the two simplicial representations, quite by chance as it is generally difficult to compare simplicial representations visually. Employing the automatable aspect of our methodology for model equivalence as described in this article, however, provides simple and rigorous comparison of simplicial representations, and we now compare all nine models for equivalence.

Of the nine models for developmental pattern formation we consider here, consisting of four TP models and five PI models, three of the models have final orbit spaces that are not isomorphic to the final orbits spaces of any of the other eight models, so each of these three models is not equivalent to any of the other eight models. These three models are PI induction-contraction, TP inhibition of an inhibition, and TP modulation. We therefore consider the remaining six models. Note that the labelling for these simplicial representations, both as positive integers and model components, is described in Vittadello and Stumpf (2021). For each of these simplicial representations the 0- and 1-simplices are specified by the model, and the higher-dimensional simplices are obtained by forming cliques, where possible, incrementally in dimensions two and higher. For simplicity, in Figs. 2 and 3 we show only the 1-skeletons of the simplicial representations.

The three PI models linear gradient, synthesis-diffusion-degradation, and opposing gradients, all have final orbit spaces that are a 0-simplex, so consist of a single vertex (Fig. 2). The unlabelled final orbit spaces are therefore isomorphic, so we can compare the corresponding sets of model components to determine model equivalence.

The orbit space for the PI linear gradient model has the label $\big\{\{1, 2, 40, 43\}, \{6, 7\}\big\}$, and the orbit space for the PI SDD model has the label $\big\{\{1, 2, 40\}, \{3, 6\}\big\}$. Equivalence
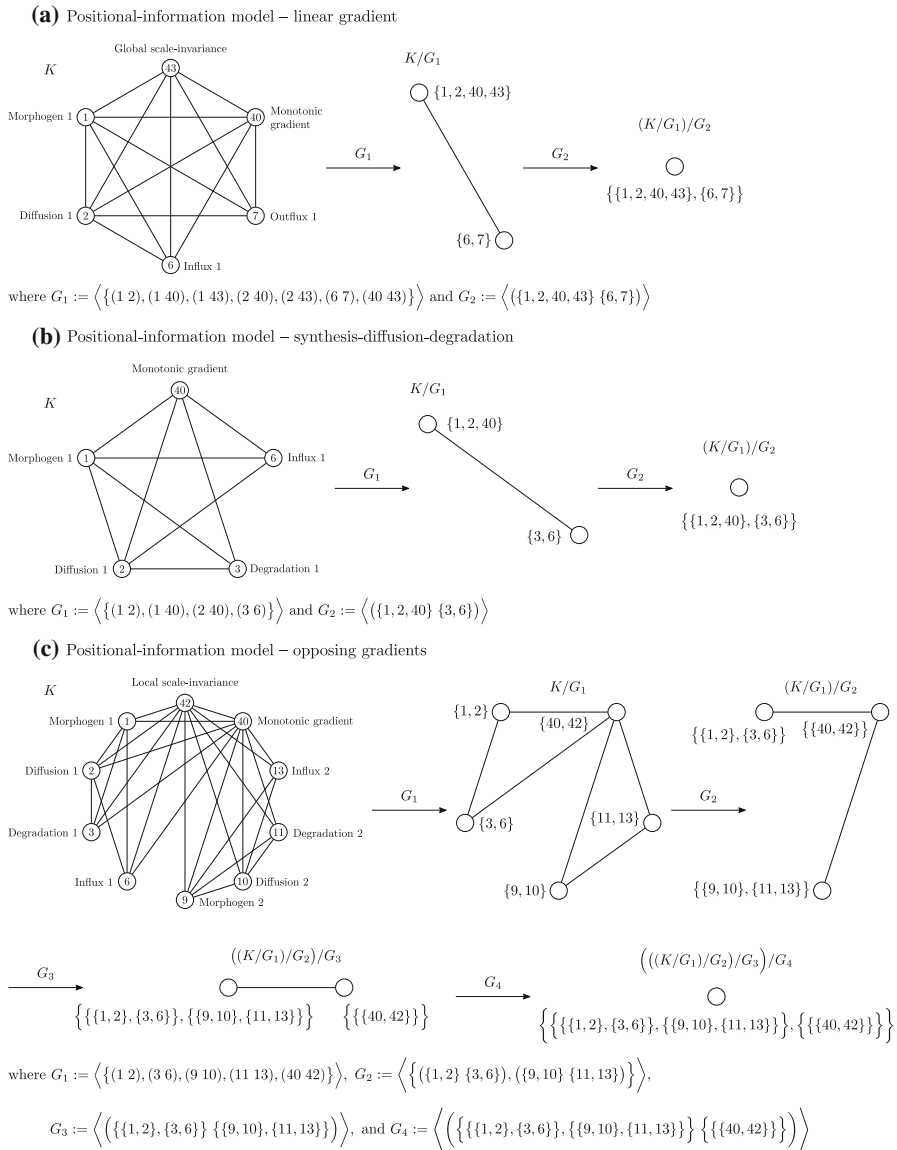
(a) Positional-information model − linear gradient

where $G_1 := \left\langle \{(1\ 2),(1\ 40),(1\ 43),(2\ 40),(2\ 43),(6\ 7),(40\ 43)\} \right\rangle$ and $G_2 := \left\langle (\{1,2,40,43\}\ \{6,7\}) \right\rangle$

(b) Positional-information model − synthesis-diffusion-degradation

where $G_1 := \left\langle \{(1\ 2),(1\ 40),(2\ 40),(3\ 6)\} \right\rangle$ and $G_2 := \left\langle (\{1,2,40\}\ \{3,6\}) \right\rangle$

(c) Positional-information model − opposing gradients

where $G_1 := \left\langle \{(1\ 2),(3\ 6),(9\ 10),(11\ 13),(40\ 42)\} \right\rangle$, $G_2 := \left\langle \{(\{1,2\}\ \{3,6\}),(\{9,10\}\ \{11,13\})\} \right\rangle$,

$G_3 := \left\langle (\{\{1,2\},\{3,6\}\}\ \{\{9,10\},\{11,13\}\}) \right\rangle$, and $G_4 := \left\langle \left(\{\{\{1,2\},\{3,6\}\},\{\{9,10\},\{11,13\}\}\}\ \{\{40,42\}\} \right) \right\rangle$

Fig. 2 1-skeletons of the simplicial representations and orbit spaces for the **(a)** PI linear gradient, **(b)** PI synthesis-diffusion-degradation, and **(c)** PI opposing gradients models

of these two models requires that we consider the model components in $\{1, 2, 40, 43\}$ and $\{6, 7\}$ as equivalent to the model components in $\{1, 2, 40\}$ and $\{3, 6\}$, respectively. That is, we need the following: 7 is equivalent to 3; since 1, 2, 40, and 43 are conceptually equivalent, any one of the vertex identifications where 1 and 43 are identified as 1, or 2 and 43 are identified as 2, or 40 and 43 are identified as 40. The equivalence of 7 (Outflux 1) and 3 (Degradation 1) may be reasonable, since they both represent removal

of Morphogen 1 from the system. The identification of 43 (Global scale-invariance) with one of 1 (Morphogen 1), 2 (Diffusion 1), or 40 (Monotonic gradient), however, assumes that we are not considering the property of scale-invariance of the morphogen gradient for any model for comparison, which would depend on the required level of conceptual detail. For our purposes, scale-invariance of the morphogen gradient is important for establishing developmental patterning, so is a necessary conceptual detail in the models, and we therefore conclude that the PI linear gradient model is not equivalent to the PI SDD model.

While the PI linear gradient model and PI SDD model each have a single morphogen, the PI opposing gradients model has two morphogens which we regard as an important conceptual detail, so we do not consider the PI opposing gradients model as equivalent to the PI linear gradient model or PI SDD model. Indeed, the label for the orbit space of the PI opposing gradients model shows that the model concepts associated with 1 (Morphogen 1), that is $\{\{1, 2\}, \{3, 6\}\}$, are identified with the model concepts associated with 9 (Morphogen 2), that is $\{\{9, 10\}, \{11, 13\}\}$, to effectively reduce to a single morphogen.

In Fig. 3 we consider the PI annihilation model, the TP activator-inhibitor model, and the TP substrate depletion model. These three models have isomorphic unlabelled final orbit spaces, so we can compare the corresponding sets of model components to determine model equivalence. Note that there are two possible isomorphisms between any two of these three orbit spaces when unlabelled, however only one of these isomorphisms preserves the labelling when the same label is in both complexes. First we consider the equivalence of the PI annihilation model and the TP activator-inhibitor model, which we established in Vittadello and Stumpf (2021), however now we employ the automatable approach using orbit spaces. We regard the model components 6, 13, 26, and 40 as conceptually equivalent to the model components {5, 27}, 12, {28, 29}, and 41, respectively. In particular, we consider vertex identifications where 5 and 27 are identified as 6, and 28 and 29 are identified as 26. We conclude that the PI annihilation model and the TP activator-inhibitor model are equivalent from a significant conceptual perspective. For further discussion of this example, including a more detailed consideration of the required partial operations, see (Vittadello and Stumpf 2021, Sect. 3.2.4).

Now consider the equivalence of the TP activator-inhibitor model and the TP substrate depletion model, which would require that the model components 9 (Morphogen 2), 10 (Diffusion 2), 11 (Degradation 2), 12 (Basal production 2), and {28, 29} (Activation of Morphogen 2 by Morphogen 1, Inhibition of Morphogen 1 by Morphogen 2) are conceptually equivalent to the model components 18 (Substrate 1), 19 (Diffusion of Substrate 1), 20 (Degradation of Substrate 1), and 21 (Basal production of Substrate 1), 34 (Depletion of Substrate 1 by Morphogen 1), respectively. We consider that the agent Morphogen 2 in the TP activator-inhibitor model is conceptually equivalent to the agent Substrate 1 in the TP substrate depletion model. Further, the two reactions corresponding to 28 and 29 form an inhibitory cycle in the TP activator-inhibitor model, which we consider to be conceptually equivalent to the reaction corresponding to 20 in the TP substrate depletion model (compare with the discussion in Vittadello and Stumpf (2021) regarding the equivalence of the PI annihilation model and the TP activator-inhibitor model). We conclude that the TP activator-inhibitor model and the

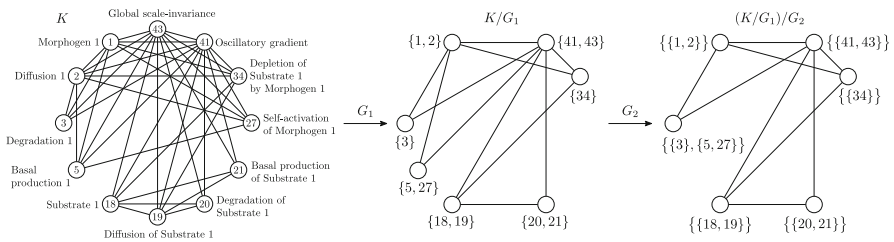**(a)** Positional-information model – annihilation
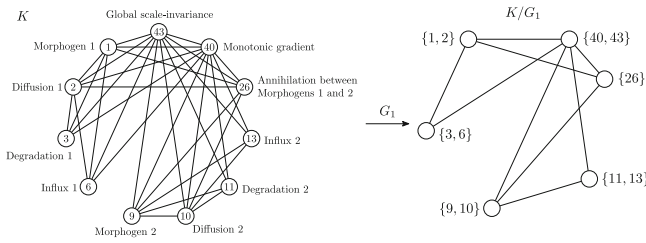
where $G_1 := \left\langle \{(1\ 2), (3\ 6), (9\ 10), (11\ 13), (40\ 43)\} \right\rangle$

(b) Turing-pattern model – activator-inhibitor
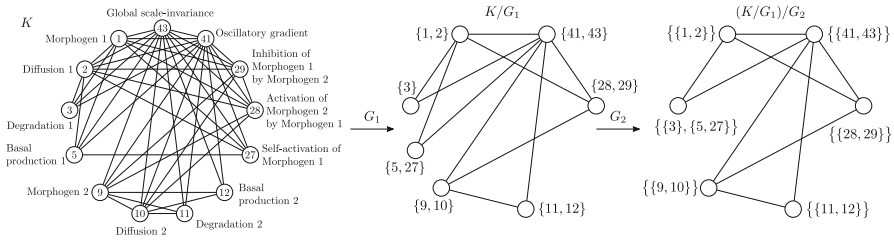
where $G_1 := \left\langle \{(1\ 2), (5\ 27), (9\ 10), (11\ 12), (28\ 29), (41\ 43)\} \right\rangle$ and $G_2 := \left\langle (\{3\}\ \{5, 27\}) \right\rangle$

(c) Turing-pattern model – substrate depletion

where $G_1 := \left\langle \{(1\ 2), (5\ 27), (18\ 19), (20\ 21), (41\ 43)\} \right\rangle$ and $G_2 := \left\langle (\{3\}\ \{5, 27\}) \right\rangle$

Fig. 3 1-skeletons of the simplicial representations and orbit spaces for the **(a)** PI annihilation, **(b)** TP activator-inhibitor, and **(c)** TP substrate depletion models

TP substrate depletion model are equivalent from a significant conceptual perspective, which is not surprising since the TP substrate depletion model is based very closely on the TP activator-inhibitor model. Since the relation of model equivalence is transitive, it also follows that the PI annihilation model is equivalent to the TP substrate depletion model.

3.5 *G*-representations

In this subsection we provide an overview of an alternative mathematical framework for our model-equivalence methodology by associating a group to a simplicial representation. While the group-theoretic framework is closely related to the simplicial-

complex framework for model equivalence, the group-theoretic approach provides an alternative mathematical perspective that allows for the application of group-theoretic techniques for model comparison. While the mathematical details of the two frameworks differ, the underlying methodology is the same, so we provide only the minimal details required to develop the group-theoretic framework for model equivalence.

We first describe a standard construction of a group from a given set.

Proposition 3.25 *Let S be a set and let* 2^S *be the power set of S. With the closed binary operation of symmetric difference,* $\triangle\colon 2^S \times 2^S \to 2^S$, *the pair* $(2^S, \triangle)$ *is a group.*

Proof Under the operation $\triangle$, 2^S is closed, the identity is the empty set $\emptyset$, and each element is self-inverse. It remains to show associativity of the operation, so let A, B, and $C \in 2^S$. We show that $(A\triangle B)\triangle C = A\triangle(B\triangle C)$. Setting some notation, for a set $D \in 2^S$ we denote by $\mathbb{1}_D\colon S \to \{0, 1\}$ the characteristic function of D such that, for $x \in S$, $x \mapsto 1$ if and only if $x \in D$. Further, denote by $\oplus$ the associative operation such that for any two characteristic functions of sets $D, E \in 2^S$ we set $(\mathbb{1}_D \oplus \mathbb{1}_E)(x) := \mathbb{1}_D(x) + \mathbb{1}_E(x) \pmod 2$. Associativity of $\triangle$ then follows from

$$\mathbb{1}_{(A\triangle B)\triangle C} = \mathbb{1}_{A\triangle B} \oplus \mathbb{1}_C = (\mathbb{1}_A \oplus \mathbb{1}_B) \oplus \mathbb{1}_C$$
$$= \mathbb{1}_A \oplus (\mathbb{1}_B \oplus \mathbb{1}_C) = \mathbb{1}_A \oplus \mathbb{1}_{B\triangle C} = \mathbb{1}_{A\triangle(B\triangle C)}.$$

$\square$

To a simplicial representation we now associate a group as in Proposition 3.25, which we call a G-representation to indicate that the representation has the structure of a group. While we could construct the group without any reference to the simplicial representation, the group encodes the same model structure in terms of model components and their interconnections, so it is natural to indicate the relationship between the group and the corresponding simplicial representation.

Definition 3.26 (G-representation) Let K be a simplicial representation of a model. The G-*representation* of the model is the group $G_K := (2^K, \triangle)$.

For a simplicial representation K, the elements of G_K are subsets of simplicies from K. These subsets are of three types: a set containing a single simplex from K is identified as a simplex; a set containing one or more simplices from K is a simplicial subcomplex of K if it is closed under taking faces, in particular a set containing just a vertex is both a simplex and a simplicial complex; a set containing two or more simplices from K that is not a simplicial subcomplex of K is a set of simplices.

Notation 3.27 Let K be a simplicial representation of a model, with corresponding G-representation G_K. The subset $\mathrm{Vert}(G_K) := \Big\{ \{\{v\}\} \mid v \in \mathrm{Vert}(K) \Big\}$ of G_K is referred to as the vertices of G_K. Further, we denote by $\rho_K\colon 2^{\mathrm{Vert}(G_K)} \setminus \emptyset \to K$ the map such that $\rho(S) = \big\{ \cup_{i\in I} \{s_i\} \big\}$ for $S = \Big\{ \{\{s_i\}\} \Big\}_{i\in I} \in 2^{\mathrm{Vert}(G_K)} \setminus \emptyset$, which sends a set of vertices in G_K to the simplex spanned by the corresponding vertices in K.

Since simplicial representations are labelled simplicial complexes, the elements of G-representations are also labelled, and there is a canonical bijection $K \mapsto G_K$ from the set of simplicial representations of models onto the set of G-representations of models.

Example 3.28 Let K be the 1-simplex with vertices labelled 1 and 2, that is, $K = \big\{\{1\}, \{2\}, \{1, 2\}\big\}$. Then $G_K = \Big\{\emptyset, \{\{1\}\}, \{\{2\}\}, \{\{1, 2\}\}, \{\{1\}, \{2\}\}, \{\{1\}, \{1, 2\}\}, \{\{2\}, \{1, 2\}\}, \{\{1\}, \{2\}, \{1, 2\}\}\Big\}$. Note that $\{\{1\}\}$ and $\{\{2\}\}$ correspond to both simplices and simplicial subcomplexes of K, $\{\{1, 2\}\}$ corresponds to a simplex of K, $\{\{1\}, \{2\}\}$ and $\{\{1\}, \{2\}, \{1, 2\}\} = K$ correspond to simplicial subcomplexes of K, while each of the elements $\{\{1\}, \{1, 2\}\}$ and $\{\{2\}, \{1, 2\}\}$ are not closed under taking faces so are sets of simplices from K.

While the group G_K contains subsets of K that are not subcomplexes, we can describe G_K in terms of the simplices in K. Recall that for a prime p, a p-group is a group in which every element has order a power of p. Since every nontrivial element of G_K has order 2, G_K is therefore an elementary commutative 2-group. Further, since G_K is a finite p-group it follows from the Burnside Basis Theorem (Rotman 1995, Chapter 5, Page 124, Theorem 5.50) that any two minimal generating sets of G_K have equal cardinality, here equal to $|K|$, and are therefore generating sets of smallest cardinality.

Proposition 3.29 *Let K be a simplicial representation of a model. Then $\mathcal{K} := \big\{\{\sigma\} \mid \sigma \in K \big\}$ is a minimal set of generators for G_K.*

Proof Let $S \in G_K$ be nontrivial. Then $S \subseteq \mathcal{K}$ by the definition of G_K, so by associativity of the symmetric difference we have $S = \triangle_{\sigma \in S}\{\sigma\}$, and it follows that $\mathcal{K}$ generates G_K as a group. Further, $\mathcal{K}$ is minimal since, for each $\{\sigma\} \in \mathcal{K}$, $\{\sigma\} \notin \langle \mathcal{K} \setminus \{\sigma\} \rangle \subseteq G_K$. $\square$

Since G_K is an elementary commutative 2-group it can be simply expressed as an internal direct sum of cyclic subgroups of subcomplexes of K, or isomorphically as an external direct sum of copies of $\mathbb{Z}/2\mathbb{Z}$.

Proposition 3.30 *Let K be a simplicial representation of a model. Then $G_K \cong \bigoplus_{\sigma \in K} \langle \{\sigma\} \rangle \cong \bigoplus_{\sigma \in K} \mathbb{Z}/2\mathbb{Z}$.*

Proof By the fundamental theorem of finite commutative groups (see (Rotman 1995, Chapter 6, Page 128, Theorem 6.5)), G_K is a direct sum of primary cyclic groups. Since $\mathcal{K}$ is a minimal set of generators for G_K by Proposition 3.29, we therefore have $G_K \cong \bigoplus_{\sigma \in K} \langle \{\sigma\} \rangle$. Finally, each $\{\sigma\} \in \mathcal{K}$ has order 2 in G_K, so $\langle \{\sigma\} \rangle \cong \mathbb{Z}/2\mathbb{Z}$. $\square$

Viewing G_K as the direct sum $\bigoplus_{\sigma \in K} \langle \{\sigma\} \rangle$ highlights the structure of the group, and allows for easier characterisation of, for example, homomorphisms between such groups.

Following from the definition of distance between two simplicial representations K and L of two models (Vittadello and Stumpf 2021), we can define the distance between

the two G-representations G_K and G_L as the cardinality of the symmetric difference of K and L. Indeed, this distance between G-representations can be formalised as a metric space, similar to the metric space of simplicial representations, and is also isometric to the latter.

The group G_K associated with a simplicial representation K is itself a representation of the model corresponding to K. The group elements of G_K have labels that are induced by the labelled simplices contained in the group elements, which in turn have labels induced by the component labels of the model. The labels of the elements of G_K must be accounted for by group homomorphisms when required. We therefore introduce our concept of a *label-preserving group homomorphism*.

Definition 3.31 (Label-preserving group homomorphism) Let G_K and G_L be two G-representations associated with two simplicial representations K and L of two models. A group homomorphism $\phi \colon G_K \to G_L$ is *label preserving* if whenever the group elements $g \in G_K$ and $h \in G_L$ have the same label then $\phi(g) = h$.

Our definition of the equivalence of two G-representations of models is as follows.

Definition 3.32 (Equivalence of G-representations) Let $\mathcal{C}$ be the set of all components that appear in the collection of models under consideration for comparison, and let $\mathcal{G}$ be a collection of G-representations G_K associated with simplicial representations which correspond to models generated by subsets from $\mathcal{C}$. Two G-representations G_K, $G_L \in \mathcal{G}$ are *equivalent* if and only if there exist two (possibly empty) sequences of invertible partial operations $(f_i)_{i=0}^m$ and $(g_i)_{i=0}^n$ on $\mathcal{G}$ such that $f_m \circ \cdots \circ f_1 \circ f_0(G_K) = g_n \circ \cdots \circ g_1 \circ g_0(G_L)$.

Note that the set $\{ (G_K, G_L) \in \mathcal{G} \times \mathcal{G} \mid G_K \text{ and } G_L \text{ are equivalent} \}$ is an equivalence relation on $\mathcal{G}$. For the equivalence of G-representations to be conceptually meaningful, just as for the equivalence of simplicial representations, we need to ensure that the employed partial operations are themselves conceptually meaningful. We consider specific operations on G-representations that are analogous to those for simplicial representations. For the G-representations, however, the operations are now induced by label-preserving group homomorphisms.

For brevity we only outline the operations on G-representations, as they correspond closely with the operations for simplicial operations. Note that for G-representations the action of an adjacent/nonadjacent-vertex identification does not send two vertices to the same vertex as for simplicial representations, but rather removes all group elements associated with one of the two vertices. This action is essentially the same as that for simplicial representations, and allows for a description in terms of group homomorphisms. Further, we omit definitions for the vertex split operation and the inclusion operation, both of which can be defined in terms of canonical injections between the direct-sum decompositions of appropriate G-representations, since we explicitly describe the equivalence of G-representations with the operations of adjacent/nonadjacent-vertex identifications and vertex substitution, while the vertex split and inclusion operations are employed implicitly. Similar to the operations for simplicial representations, all five of these partial operations on a set of G-representations are invertible for suitable domains of definition, and the equivalence of G-representations must be based on the admissibility of the partial operations.

For all operations, let $\mathcal{C}$ be the set of all components of models under consideration, and let G_K and G_L be G-representations associated with the two simplicial representations K and L with labels from $\mathcal{C}$.

> **Adjacent-vertex identification:** Suppose $\{\{u, v\}\} \in G_K$, where $u, v \in \mathrm{Vert}(K)$, and that assumptions hold for u and v corresponding to those in Definition 2.9. A label-preserving group homomorphism $\phi_1 \colon G_K \to G_L$ is an *adjacent-vertex identification* if $G_L = \bigoplus_{\{\sigma \in K \mid v \notin \sigma\}} \langle\{\sigma\}\rangle$ and $\phi_1 \colon \bigoplus_{\sigma \in K} \langle\{\sigma\}\rangle \to \bigoplus_{\{\sigma \in K \mid v \notin \sigma\}} \langle\{\sigma\}\rangle$ is the canonical projection.
>
> **Nonadjacent-vertex identification:** Suppose $\{\{u\}\}, \{\{v\}\} \in G_K, \{\{u, v\}\} \notin G_K$, and that assumptions hold for u and v corresponding to those in Definition 2.10. A label-preserving group homomorphism $\phi_2 \colon G_K \to G_L$ is a *nonadjacent-vertex identification* if $G_L = \bigoplus_{\{\sigma \in K \mid v \notin \sigma\}} \langle\{\sigma\}\rangle$ and $\phi_2 \colon \bigoplus_{\sigma \in K} \langle\{\sigma\}\rangle \to \bigoplus_{\{\sigma \in K \mid v \notin \sigma\}} \langle\{\sigma\}\rangle$ is the canonical projection.
>
> **Vertex substitution:** Suppose $u \in \mathrm{Vert}(K)$, $v \in \mathrm{Vert}(L)$, and $\mathrm{Vert}(K) \setminus \{u\} = \mathrm{Vert}(L) \setminus \{v\}$. Let $\pi \colon \mathrm{Vert}(K) \to \mathrm{Vert}(L)$ be a map that is injective except that $\pi(u) = v$. A label-preserving group isomorphism $\phi_5 \colon G_K \to G_L$ is a *vertex substitution* if, for each $\sigma \in K$, $\phi_5(\{\sigma\}) = \{\{\pi(w) \mid w \in \sigma\}\}$.

To determine whether there are conceptually equivalent vertices in a G-representation G_K, we can employ group actions on G_K that correspond to exchange automorphisms for simplicial representations. Recall that an action of a group H on a group G is a group homomorphism $\Theta \colon H \to \mathrm{Aut}(G)$. While we required orbit spaces for vertex identifications when considering simplicial representations, for G-representations we have no need for orbit spaces and can employ G-representations directly. The methodology for determining model equivalence detailed in Sect. 3.3 is easily adapted for G-representations.

4 Conclusion

In this article we develop a rigorous and automatable methodology for determining whether vertices in a simplicial representation are conceptually related, corresponding to conceptually-related model components, and then identifying these vertices using simplicial operations. This process is the main consideration of comparison by equivalence as defined in Vittadello and Stumpf (2021), and exploits the symmetry associated with conceptually equivalent vertices in a simplicial representation by constructing group actions on the simplicial representations. We develop the required mathematical theory for group actions on simplicial complexes, and demonstrate how this methodology greatly simplifies and facilitates the process of determining model equivalence by reducing the initial simplicial representations to simplicial complexes with no conceptually equivalent vertices. Our approach to model comparison by equivalence provides a rigorous and efficient framework for comparing the underlying conceptual structure of mathematical models of a particular biological system, as well as mathematical models of different biological systems.

The important point to note is that our approach, like our previous one, differs from the overwhelming majority of model comparison approaches by being independent of

outputs: it is solely based on the underlying conceptual structure of the model. Our definition of model structure is also more inclusive than, for example, motif-based definitions (Alon 2007), as these are known to fall short of constraining dynamics (Ingram et al. 2006). By including model properties such as boundary conditions and symmetry relationships, for example, as components of the simplicial complexes, our approach here is capable of capturing even subtle differences in the dynamics. Our methodology is applicable to all models irrespective of mathematical formalism, and allows for comparison from different conceptual perspectives.

Model comparison is essential for managing the seemingly heterogeneous and possibly numerous models associated with a system. We conclude this analysis with a discussion of an alternative mathematical framework for our model comparison methodology by representing models as groups, which provides for the direct application of techniques from group theory within our model-comparison methodology. That we can represent our model comparison methodology with two distinct mathematical frameworks shows the flexibility of our approach, which is important for the utilisation of a broad range of mathematical techniques when working on the difficult problem of model comparison.

Modelling in biology cannot rely on the same fundamental principles—such as symmetries and conservation laws—that have been so successful in physics. Instead, statistical methods (Kirk et al. 2013), robustness analysis (Bates and Cosentino 2011), or algebraic methods (Araujo and Liotta 2018) are required to narrow down the vast 'universe' of potential mathematical models in order to identify those models that can reasonably explain a given scenario. The approach developed here is capable of taking these models, including large sets of good candidates (Scholes et al. 2019), and distilling design principles from this analysis. Design principles, or the conceptual bases for mathematical models describing a given system, are arguably our best chance to gain conceptual mechanistic insights into biological systems.

Declarations

Conflict of interest We declare we have no competing interests.

5 Appendix

Here we provide the details of the mathematical models that we use in the example in Sect. 3.4. We also provide a brief discussion on representing the models as simplicial complexes.

These models are from the two main classes of models for developmental pattern formation, namely Turing-pattern (TP) models and positional-information (PI) models. We consider the patterning to occur within a two-dimensional rectangular domain, with zero-flux boundary conditions on the two sides parallel to the morphogen concentration gradient. We assume that the velocities of the cytoplasm and the growing tissue are negligible; we therefore assume no advection. Further details of the models and the associated simplicial representations, including the vertex numbers associated with the simplicial representations, are in Vittadello and Stumpf (2021). Note that the boundary conditions for the TP models may be Dirichlet, Neumann, Robin, or periodic (Dillon et al. 1994; Varea et al. 1997; Barrio et al. 1999).

5.1 Positional-information: Linear gradient

A linear concentration profile of a morphogen results when the production and degradation of the morphogen occur outside and on opposite sides of the tissue domain, and the morphogen undergoes passive diffusion along the domain from the side where it is produced to the side where it is degraded (Stumpf 1967; Wolpert 1969; Crick 1970; Čapek and Müller 2019). Mathematically, the steady-state morphogen concentration in this system satisfies Laplace's equation, which yields global scale-invariant positional information. Specifically, if the tissue length is L with initial position at $x = 0$ then the morphogen concentration $m(x, t)$ can be modelled as

$$\frac{\partial m}{\partial t} = D \frac{\partial^2 m}{\partial x^2}, \tag{3}$$

where D is the morphogen diffusivity. The steady-state solution of Eq. 3 is the linear equation $m(x) = ax + b$, for arbitrary constants $a, b \in \mathbb{R}$. The boundary conditions are influx at one end and outflux at the other end, and we may assume without loss of generality that influx occurs at $x = 0$ and outflux at $x = L$. The original Dirichlet boundary conditions specify the constant concentrations $m(0, t) = m_0 > 0$ and $m(L, t) = m_L \geq 0$, with $m_0 > m_L$, so the solution is $m(x) = ((m_L - m_0)/L)x + m_0$. Monotonically decreasing linear gradients can also be achieved with Neumann boundary conditions at one boundary and Dirichlet boundary conditions at the opposite boundary.

The simplicial representation of the linear model is shown in Fig. 2a. Note that the simplicial complex is 4-dimensional, however we only show the 1-skeleton of the simplicial complex for simplicity.

5.2 Positional-information: Synthesis-diffusion-degradation (SDD)

In the SDD model a morphogen gradient forms by morphogen production from a localised source at the boundary combined with morphogen diffusion and uniform degradation throughout the tissue (Wartlick et al. 2009; Umulis and Othmer 2013; Čapek and Müller 2019). Mathematically, the morphogen concentration $m(\boldsymbol{x}, t)$ can be modelled as

$$\frac{\partial m}{\partial t} = D\nabla^2 m - km, \tag{4}$$

where D is the morphogen diffusivity, and k is the morphogen degradation rate which is independent of morphogen concentration. The boundary conditions are Neumann for the influx and zero outflux. The steady-state morphogen gradient is an exponential function, so this system is not scale invariant unless D, k, and the influx vary with the tissue length L in a very specific manner (Umulis and Othmer 2013).

The simplicial representation of the SDD model is shown in Fig. 2b. Note that the simplicial complex is 3-dimensional, however we only show the 1-skeleton of the simplicial complex for simplicity.

5.3 Positional-information: Opposing gradients

There are two mechanisms whereby two opposing morphogen gradients provide size information for developmental patterning (McHale et al. 2006): the gene expression depends on the relative concentrations of the two morphogens. One is the annihilation model, and the other is the scaling-by-opposing-gradients model.

For the opposing gradients model the sources of each morphogen are at opposite ends of the domain. The combination of the two morphogen gradients provides effective local scaling for the boundaries of individual target genes, however it does not produce global scaling across the entire domain (Wolpert 1969; Houchmandzadeh et al. 2005; McHale et al. 2006; Čapek and Müller 2019). The degree of scaling for the opposing gradients mechanism can be increased if the morphogens are irreversibly inactivated upon binding to each other (McHale et al. 2006). Here we assume no direct interaction between the two morphogens. Mathematically, the two morphogen gradients with concentrations $m(\boldsymbol{x}, t)$ and $c(\boldsymbol{x}, t)$ can be modelled as

$$\frac{\partial m}{\partial t} = D_m \nabla^2 m - k_m m, \tag{5}$$

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c - k_c c, \tag{6}$$

where D_m and D_c are diffusivities, and k_m and k_c are degradation rates. The boundary conditions for each morphogen are Neumann at both boundaries, with flux at the source and zero flux at the opposite boundary.

The simplicial representation of the opposing gradients model is shown in Fig. 2c. Note that the simplicial complex is 4-dimensional, however we only show the 1-skeleton of the simplicial complex for simplicity.

5.4 Positional-information: Induction-contraction

The largest degree of scaling across the patterning domain is obtained when the biophysical properties of the morphogen are influenced by an accessory modulator that senses the size of the domain (Wartlick et al. 2009; Ben-Zvi and Barkai 2010; Umulis and Othmer 2013; Čapek and Müller 2019). The modulator molecules with concentration $c(\boldsymbol{x}, t)$ may influence the diffusivity, degradation rate, or influx rate, of the morphogen with concentration $m(\boldsymbol{x}, t)$. If the kinetics and source of the modulator are dependent on the morphogen concentration then the mechanism is *active modulation*. An example of active modulation is the *induction-contraction* model, which scales globally (Rahimi et al. 2016; Shilo and Barkai 2017). In this case, the morphogen induces the production of a contractor molecule, here the modulator with high diffusivity once again, which contracts the range of the morphogen gradient through a decrease in the morphogen diffusivity or an increase of the morphogen degradation rate. Note that while the amplitude and shape of the morphogen gradient is globally scale-invariant, the same is not necessarily true of the modulator since the modulator level reflects the domain size and therefore increases or decreases accordingly.

The scaling by modulation model with the induction-contraction mechanism can be represented mathematically as

$$\frac{\partial m}{\partial t} = D_m(c)\nabla^2 m - k_m(c)m, \tag{7}$$

$$\frac{\partial c}{\partial t} = D_c\nabla^2 c - k_c c + \rho(m), \tag{8}$$

where $D_m(c)$ and D_c are diffusivities, $k_m(c)$ and k_c are degradation rates, and $\rho(m)$ is the localised production source for the modulator. Note that $D_m(c)$ is a decreasing function of c, $k_m(c)$ is an increasing function of c, and $\rho(m)$ is an increasing function of m. Further note that 'degradation' in this context refers to not only physical destruction, but to any mechanism that effects the removal of the morphogen from the patterning system, including irreversible complex formation. Morphogen boundary conditions are Neumann for the influx and zero flux at the opposite end of the domain. For the modulator there are various possibilities for the boundary conditions, such as the same type of boundary conditions as the morphogen when the modulator source is outside the domain or, as we consider here, zero flux at both boundaries when the modulator source is within the domain.

5.5 Positional-information: Annihilation model

The annihilation model, and also the opposing gradients model, are mechanisms whereby two opposing morphogen gradients provide size information for developmental patterning. In the annihilation model, the target gene responds to the concentration of Morphogen 1, to which Morphogen 2 irreversibly binds and thereby inhibits the action of Morphogen 1 on activity of transcription, so that the gradient of Morphogen 2 provides size information to the concentration field of Morphogen 1 (McHale et al. 2006).

The sources of each morphogen are at opposite ends of the domain, and the morphogens interact by an annihilation reaction with rate k that results in global scale-invariant patterning (Ben-Naim and Redner 1992; McHale et al. 2006). Mathematically, the two morphogen gradients with concentrations $m(\boldsymbol{x}, t)$ and $c(\boldsymbol{x}, t)$ can be modelled as

$$\frac{\partial m}{\partial t} = D_m \nabla^2 m - k_m m - kmc, \tag{9}$$

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c - k_c c - kmc, \tag{10}$$

where D_m and D_c are diffusivities, and k_m and k_c are degradation rates. The boundary conditions for each morphogen are Neumann at both boundaries, with influx at the source and zero flux at the opposite boundary.

The simplicial representation of the annihilation mode is shown in Fig. 3a. Note that the simplicial complex is 4-dimensional, however we only show the 1-skeleton of the simplicial complex for simplicity.

5.6 Turing-pattern: Activator-inhibitor model

The activator-inhibitor system (Gierer and Meinhardt 1972; Meinhardt 2012; Landge et al. 2020) consists of two diffusible morphogens, an autocatalytic activator with concentration $m(\boldsymbol{x}, t)$ and a rapidly diffusing inhibitor with concentration $c(\boldsymbol{x}, t)$. This model can be represented mathematically as

$$\frac{\partial m}{\partial t} = D_m \nabla^2 m + \frac{\rho m^2}{c(1 + \mu_m m^2)} - k_m m + \rho_m, \tag{11}$$

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c + \rho m^2 - k_c c + \rho_c, \tag{12}$$

where D_m and D_c are diffusivities, ρ_m and ρ_c are basal production rates, k_m and k_c are degradation rates, and μ_m is a saturation constant. The parameter ρ is the *source density*, which measures the general ability of the cells to perform the autocatalytic reaction. The patterning arises through local self-enhancement of the activator, activation of the inhibitor, and long-range inhibition of the activator. We assume that the boundary conditions are zero flux at both boundaries.

The simplicial representation of the activator-inhibitor model is shown in Fig. 3b. Note that the simplicial complex is 5-dimensional, however we only show the 1-skeleton of the simplicial complex for simplicity.

5.7 Turing-pattern: Substrate depletion

While long-range inhibition of an autocatalytic activator can occur with an inhibitor, as in the activator-inhibitor model, the antagonistic effect can also arise from the depletion of a substrate that is consumed during activator production. In this system, the production of an autocatalytic activator with concentration $m(\boldsymbol{x}, t)$ results in either direct or indirect depletion of a substrate with concentration $s(\boldsymbol{x}, t)$ (Meinhardt 2012). This model can be represented mathematically as

$$\frac{\partial m}{\partial t} = D_m \nabla^2 m + \rho s m^2 - k_m m + \rho_m, \tag{13}$$

$$\frac{\partial s}{\partial t} = D_s \nabla^2 s - \rho s m^2 - k_s s + \rho_s, \tag{14}$$

where D_m and D_s are diffusivities, ρ_m and ρ_s are basal production rates, k_m and k_s are degradation rates, and ρ is the source density for the autocatalytic reaction of the activator, similar to the activator-inhibitor model. The diffusivity D_s of the substrate must be much faster than the diffusivity D_m of the activator, and it is assumed that the substrate is produced uniformly throughout the domain. We assume that the boundary conditions are zero flux at both boundaries.

The simplicial representation of the substrate depletion model is shown in Fig. 3c. Note that the simplicial complex is 5-dimensional, however we only show the 1-skeleton of the simplicial complex for simplicity.

5.8 Turing-pattern: Inhibition of an inhibition

Dynamics analogous to the activator-inhibitor model can be realised through an inhibition of an inhibition mechanism (Meinhardt 2012). In this case, two morphogens with concentrations $a(\boldsymbol{x}, t)$ and $c(\boldsymbol{x}, t)$ inhibit the production of each other, thereby forming a switching system in which one of the morphogens becomes fully activated similar to being autcatalytic. In order that pattern formation occur, a third morphogen with concentration $b(\boldsymbol{x}, t)$ acts as a long-range signal that disrupts the indirect self-enhancement of either a or c. Morphogen b is rapidly diffusing, is produced under control of a, and inhibits the inhibition of c production by a, therefore acting as inhibitor. This model can be represented mathematically as

$$\frac{\partial a}{\partial t} = D_a \nabla^2 a + \frac{\rho_a}{\kappa_a + c^2} - k_a a, \tag{15}$$

$$\frac{\partial b}{\partial t} = D_b \nabla^2 b + k_b(a - b), \tag{16}$$

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c + \frac{\rho_c}{\kappa_c + a^2/b^2} - k_c c, \tag{17}$$

where D_a, D_b, and D_c are diffusivities, ρ_a and ρ_c are production rates, k_a and k_c are degradation rates, k_b is both the rate of production and degradation of b, and κ_a and κ_c are saturation constants that limit the production rate if the concentrations of a or c become too low. We assume that the boundary conditions are zero flux at both boundaries.

5.9 Turing-pattern: Modulation

The diffusion of a morphogen may be inhibited by adsorption on negatively-charged extracellular matrix (ECM) components, resulting in modulated diffusion and a smaller effective diffusivity for the morphogen (Nesterenko et al. 2017). The corresponding modulation model (Nesterenko et al. 2017) is an extension of the activator-inhibitor model (Gierer and Meinhardt 1972; Meinhardt 2012; Landge et al. 2020). Our description of the modulation model is based on the version of the activator-inhibitor model described above. The modulation model consists of two diffusible morphogens, an autocatalytic activator and an inhibitor, along with available binding sites on the ECM onto which the activator adsorbs. The inhibitor does not bind to the ECM, and the free activator and the inhibitor have equal diffusivities D. It is assumed that autocatalysis and activation of the inhibitor by the activator occurs in both the free and adsorbed states. It is also assumed that the degradation rate of the activator is equal in both the free and adsorbed states. Free binding sites on the ECM therefore appear due to both desorption and degradation of the activator. Such modulation allows for stable dissipative morphogens gradients. Mathematically, the free activator has concentration $a(\boldsymbol{x}, t)$, the bound activator has concentration $b(\boldsymbol{x}, t)$, the inhibitor has concentration $c(\boldsymbol{x}, t)$, and the concentration of available binding sites on the ECM is $s(\boldsymbol{x}, t)$:

$$\frac{\partial a}{\partial t} = D\nabla^2 a + \frac{\rho(a+b)^2}{c(1 + \mu_a(a+b)^2)} - k_a a + \rho_a - k_1 s a + k_{-1} b, \tag{18}$$

$$\frac{\partial c}{\partial t} = D\nabla^2 c + \rho(a+b)^2 - k_c c + \rho_c, \tag{19}$$

$$\frac{\partial s}{\partial t} = -k_1 s a + (k_{-1} + k_a)b, \tag{20}$$

where D is the diffusivity, ρ_a and ρ_c are basal production rates, k_a and k_c are degradation rates, μ_a is the saturation constant, ρ is the source density, k_1 is the rate of adsorption of the activator onto the ECM, and k_{-1} is the rate of desorption of the activator from the ECM. We assume that the boundary conditions are zero flux at both boundaries.

5.10 Construction of the simplicial representations

The simplicial representations of the four Turing-pattern and five positional-information models are constructed by first determining the set of components $\mathcal{C}$ on which all of the the models under consideration are based. In this case the general components are: agents, namely morphogens, modulators, and substrates; reactions involving the agents, such as self-activation, activation, inhibition and annihilation; agent degradation; influx and outflux boundary conditions; agent diffusion; profile of the morphogen gradient; and scale invariance of the morphogen gradient. The ordered set of these components is shown in Figure 3 in Vittadello and Stumpf (2021).

The reference complex is the simplex spanned by the complete set of components for all nine models. To construct the simplicial representation of each Turing-pattern and positional-information model, we first specify the 0-simplices, which represent the model components, and the 1-simplices, which represent direct interconnections between the components. While the 0- and 1-simplices are specified by the model, to give a combinatorial graph, for these models we obtain higher-dimensional simplices by forming cliques where possible, incrementally in dimensions 2 and higher. These higher-dimensional simplices indicate higher-dimensional interactions between the corresponding model components.

References

Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8:450–461. https://doi.org/10.1038/nrg2102

Araujo RP, Liotta LA (2018) The topological requirements for robust perfect adaptation in networks of any size. Nat Commun 9:1757. https://doi.org/10.1038/s41467-018-04151-6

Babtie AC, Kirk P, Stumpf MPH (2014) Topological sensitivity analysis for systems biology. Proc Natl Acad Sci 111:18507–18512. https://doi.org/10.1073/pnas.1414026112

Banwarth-Kuhn M, Sindi S (2020) How and why to build a mathematical model: a case study using prion aggregation. J Biol Chem 295:5022–5035. https://doi.org/10.1074/jbc.REV119.009851

Barnes CP, Silk D, Sheng X, Stumpf MPH (2011) Bayesian design of synthetic biological systems. Proc Natl Acad Sci 108:15190–15195. https://doi.org/10.1073/pnas.1017972108

Barrio RA, Varea C, Aragón JL, Maini PK (1999) A two-dimensional numerical study of spatial pattern formation in interacting Turing systems. Bull Math Biol 61:483–505. https://doi.org/10.1006/bulm.1998.0093

Bates DG, Cosentino C (2011) Validation and invalidation of systems biology models using robustness analysis. IET Syst Biol 5:229–244. https://doi.org/10.1049/iet-syb.2010.0072

Ben-Naim E, Redner S (1992) Inhomogeneous two-species annihilation in the steady state. J Phys A 25:L575–L583. https://doi.org/10.1088/0305-4470/25/9/012

Ben-Zvi D, Barkai N (2010) Scaling of morphogen gradients by an expansion-repression integral feedback control. Proc Natl Acad Sci 107:6924–6929. https://doi.org/10.1073/pnas.0912734107

Bredon GE (1972) Introduction to compact transformation groups of pure and applied mathematics, vol 46. Academic Press, New York

Brophy JAN, Voigt CA (2014) Principles of genetic circuit design. Nat Methods 11:508–520. https://doi.org/10.1038/nmeth.2926

Cabbia A, Hilbers PAJ, van Riel NAW (2020) A distance-based framework for the characterization of metabolic heterogeneity in large sets of genome-scale metabolic models. Patterns 1:100080. https://doi.org/10.1016/j.patter.2020.100080

Čapek D, Müller P (2019) Positional information and tissue scaling during development and regeneration. Development 146:dev177709. https://doi.org/10.1242/dev.177709

Crick F (1970) Diffusion in embryogenesis. Nature 225:420–422. https://doi.org/10.1038/225420a0

Dillon R, Maini PK, Othmer HG (1994) Pattern formation in generalized Turing systems. J Math Biol 32:345–393. https://doi.org/10.1007/BF00160165

Gay S, Soliman S, Fages F (2010) A graphical method for reducing and relating models in systems biology. Bioinformatics 26:i575–i581. https://doi.org/10.1093/bioinformatics/btq388

Gierer A, Meinhardt H (1972) A theory of biological pattern formation. Kybernetik 12:30–39. https://doi.org/10.1007/BF00289234

Green JBA, Sharpe J (2015) Positional information and reaction-diffusion: two big ideas in developmental biology combine. Development 142:1203–1211. https://doi.org/10.1242/dev.114991

Gunawardena J (2014) Models in biology: 'accurate descriptions of our pathetic thinking'. BMC Biol 12:29. https://doi.org/10.1186/1741-7007-12-29

Henkel R, Hoehndorf R, Kacprowski T, Knüpfer C, Liebermeister W, Waltemath D (2018) Notions of similarity for systems biology models. Brief Bioinform 19:77–88. https://doi.org/10.1093/bib/bbw090

Houchmandzadeh B, Wieschaus E, Leibler S (2005) Precise domain specification in the developing Drosophila embryo. Phys Rev E 72:061920. https://doi.org/10.1103/PhysRevE.72.061920

Howe D, Costanzo M, Fey P, Gojobori T, Hannick L, Hide W et al (2008) The future of biocuration. Nature 455:47–50. https://doi.org/10.1038/455047a

Ingram PJ, Stumpf MPH, Stark J (2006) Network motifs: structure does not determine function. BMC Genom 7(1):108. https://doi.org/10.1186/1471-2164-7-108

International Society for Biocuration (2018) Biocuration: distilling data into knowledge. PLoS Biol 16:e2002846. https://doi.org/10.1371/journal.pbio.2002846

Karplus WJ (1977) The spectrum of mathematical modeling and systems simulation. Math Comput Simul 19:3–10. https://doi.org/10.1016/0378-4754(77)90034-9

King J, Eroumé KS, Truckenmüller R, Giselbrecht S, Cowan AE, Loew L et al (2021) Ten steps to investigate a cellular system with mathematical modeling. PLoS Comput Biol 17:e1008921. https://doi.org/10.1371/journal.pcbi.1008921

Kirk P, Thorne T, Stumpf MPH (2013) Model selection in systems and synthetic biology. Curr Opin Biotechnol 24:767–774. https://doi.org/10.1016/j.copbio.2013.03.012

Kozlov D (2008) Combinatorial algebraic topology of algorithms and computation in mathematics, vol 21. Springer, Berlin. https://doi.org/10.1007/978-3-540-71762-5

Landge AN, Jordan BM, Diego X, Müller P (2020) Pattern formation mechanisms of self-organizing reaction-diffusion systems. Dev Biol 460:2–11. https://doi.org/10.1016/j.ydbio.2019.10.031

Mahmud M, Kaiser MS, McGinnity TM, Hussain A (2021) Deep learning in mining biological data. Cogn Comput 13:1–33. https://doi.org/10.1007/s12559-020-09773-x

Marx V (2013) The big challenges of big data. Nature 498:255–260. https://doi.org/10.1038/498255a

McHale P, Rappel WJ, Levine H (2006) Embryonic pattern scaling achieved by oppositely directed morphogen gradients. Phys Biol 3:107–120. https://doi.org/10.1088/1478-3975/3/2/003

Meinhardt H (2012) Turing's theory of morphogenesis of 1952 and the subsequent discovery of the crucial role of local self-enhancement and long-range inhibition. Interface Focus 2:407–416. https://doi.org/10.1098/rsfs.2011.0097

Munkres J (2018) Elements of algebraic topology. CRC Press, London. https://doi.org/10.1201/9780429493911

Nesterenko AM, Kuznetsov MB, Korotkova DD, Zaraisky AG (2017) Morphogene adsorption as a turing instability regulator: theoretical analysis and possible applications in multicellular embryonic systems. PLoS ONE 12:e0171212. https://doi.org/10.1371/journal.pone.0171212

Papadopoulos A (ed) (2012) Handbook of Teichmüller theory, vol Volume. European Mathematical Society, III

Pezzulo G, Levin M (2016) Top-down models in biology: explanation and control of complex living systems above the molecular level. J Royal Society Interface. 13:20160555. https://doi.org/10.1098/rsif.2016.0555

Rahimi N, Averbukh I, Haskel-Ittah M, Degani N, Schejter ED, Barkai N et al (2016) A WntD-dependent integral feedback loop attenuates variability in Drosophila toll signaling. Dev Cell 36:401–414. https://doi.org/10.1016/j.devcel.2016.01.023

Rosenblueth A, Wiener N (1945) The role of models in science. Philosophy of Science. 12:316–321. https://doi.org/10.1086/286874

Rotman JJ (1988) An introduction to algebraic topology. Springer, New York. https://doi.org/10.1007/978-1-4612-4576-6

Rotman JJ (1995) An introduction to the theory of groups. Springer, New York. https://doi.org/10.1007/978-1-4612-4176-8

Scholes NS, Schnoerr D, Isalan M, Stumpf MPH (2019) A comprehensive network atlas reveals that turing patterns are common but not robust. Cell Syst 9:243–257. https://doi.org/10.1016/j.cels.2019.07.007

Shilo BZ, Barkai N (2017) Buffering global variability of morphogen gradients. Dev Cell 40:429–438. https://doi.org/10.1016/j.devcel.2016.12.012

Sneddon MW, Faeder JR, Emonet T (2010) Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. Nat Methods 8(2):177–183. https://doi.org/10.1038/nmeth.1546

Spanier EH (1966) Algebraic topology. Springer, New York. https://doi.org/10.1007/978-1-4684-9322-1

Stumpf HF (1967) Über den Verlauf eines schuppenorientierenden Gefälles bei Galleria mellonella. Wilhelm Roux Arch Entwickl Mech Org 158:315–330. https://doi.org/10.1007/bf00573402

Tapinos A, Mendes P (2013) A method for comparing multivariate time series with different dimensions. PLoS ONE 8:e54201. https://doi.org/10.1371/journal.pone.0054201

Tomlin CJ, Axelrod JD (2007) Biology by numbers: mathematical modelling in developmental biology. Nat Rev Genet 8:331–340. https://doi.org/10.1038/nrg2098

Torres NV, Santos G (2015) The (mathematical) modeling process in biosciences. Front Genet 6:354. https://doi.org/10.3389/fgene.2015.00354

Transtrum MK, Qiu P (2016) Bridging mechanistic and phenomenological models of complex biological systems. PLoS Comput Biol 12:e1004915. https://doi.org/10.1371/journal.pcbi.1004915

Umulis DM, Othmer HG (2013) Mechanisms of scaling in pattern formation. Development 140:4830–4843. https://doi.org/10.1242/dev.100511

Varea C, Aragón JL, Barrio RA (1997) Confined turing patterns in growing systems. Phys Rev E 56:1250–1253. https://doi.org/10.1103/PhysRevE.56.1250

Vittadello ST, Stumpf MPH (2021) Model comparison via simplicial complexes and persistent homology. R Soc Open Sci 8:211361. https://doi.org/10.1098/rsos.211361

Wartlick O, Kicheva A, González-Gaitán M (2009) Morphogen gradient formation. Cold Spring Harb Perspect Biol 1:a001255. https://doi.org/10.1101/cshperspect.a001255

Wolkenhauer O (2014) Why model? Front Physiol 5:21. https://doi.org/10.3389/fphys.2014.00021

Wolpert L (1969) Positional information and the spatial pattern of cellular differentiation. J Theor Biol 25:1–47. https://doi.org/10.1016/S0022-5193(69)80016-0