Designing Bivariate Auto-Regressive Timeseries with Controlled Granger Causality

Shohei Hidaka * and **Takuma Torii**

Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi 923-1292, Ishikawa, Japan;
tak.torii@jaist.ac.jp
* Correspondence: shhidaka@jaist.ac.jp

Abstract: In this manuscript, we analyze a bivariate vector auto-regressive (VAR) model in order to draw the design principle of a timeseries with a controlled statistical inter-relationship. We show how to generate bivariate timeseries with given covariance and Granger causality (or, equivalently, transfer entropy), and show the trade-off relationship between these two types of statistical interaction. In principle, covariance and Granger causality are independently controllable, but the feasible ranges of their values which allow the VAR to be proper and have a stationary distribution are constrained by each other. Thus, our analysis identifies the essential tri-lemma structure among the stability and properness of VAR, the controllability of covariance, and that of Granger causality.

Keywords: Granger causality; transfer entropy; vector auto-regressive model; Lyapunov equation

1. Introduction

1.1. Background and Motivation

In the field of cognitive psychology, the human perception of the life-likeness (called *animacy perception*) of one or multiple moving geometric patterns has been studied for decades [1–5]. There are multiple findings on the effect of "synchrony" or "temporal contingency" between multiple moving points on animacy perception. Findings from one line of research [2] have suggested that a higher degree of "temporal contingency" of the moving objects is related to a higher likelihood of animacy perception. Findings from the other line of research [6] have suggested that the highest "temporal contingency", presented in the form of perfect synchronization, would decrease the likelihood of animacy perception.

These two lines of research have together suggested the existence of multiple types of "temporal contingency". Nevertheless, this past research does not appear to clarify these types. Further, confusion surrounding these two distinct types of effects have led to two lines of apparently conflicting effects of "temporal contingency".

With this potential conflict in the literature on animacy perception in mind, we explore a theoretical framework which can generate timeseries of multiple random variables with multiple distinct types of statistical dependency. One such system, which is sufficiently simple and readily manipulable, is vector auto-regression (VAR). Vector auto-regression is a random process for generating multivariate timeseries for a given set of parameters. In this manuscript, we specifically consider only bivariate VAR, which is a minimal system with interaction between two moving points.

1.2. Vector Auto-Regressive Model, Granger Causality, and Transfer Entropy

Importantly, bivariate VAR, a series of paired random variables (x_t, y_t) for $t = 0, 1, \ldots$, has two types of statistical dependency—that is, the correlation and *Granger causality* of a timeseries [7], which has been identified as transfer entropy [8] of the timeseries generated by a Gaussian process by [9]. The correlation between univariate series x and y is a statistical dependency between x_t and y_t in the limit $t \to \infty$ (if it exists), while a Granger causality from y to x is that between x_t and y_{t-1} given x_{t-1} in the limit $t \to \infty$ (if it exists).

Conceptually, correlation captures a type of similarity between two timeseries, whereas Granger causality captures the "reactiveness" of one timeseries to another.

Thus, given these differences in theory, our goal is to propose a theoretical method to generate a bivariate random timeseries with a desired correlation and two types. The Granger causality of VAR has also been considered in other fields. In econometrics, the field in which it was originally proposed [7], Granger causality has been used as a measure of interaction between a pair of economic timeseries [10–12]. It has also been used in general behavioral sciences [13,14], particularly in computational neuroscience [15–18]. Such an in-principle data generation technique would be vital to the testing of any hypothesis on the empirical nature of timeseries (e.g., animacy perception) in the empirical sciences using VAR and Granger causality, as mentioned above. To our knowledge, however, there has been no mathematical analysis of the theoretical limitation of such a data generation technique for a given statistics.

Thus, we first need to explore the mathematical relationship among the parameters in a VAR with the correlation and Granger causality in a timeseries generated from it. In this paper, we therefore explore a theoretical structure of bivariate VAR from the designer's perspective, and analyze a mathematical limit to the extent which we can simultaneously control correlation and Granger causality of a bivariate timeseries.

This paper is written with the following structure. In Section 2, the VAR model is defined, from which a set of basic statistical properties of the VAR model are derived, such as Granger causality (Section 3) with a set of parameters. In Section 4, the existence of the stationary distribution of the VAR is analyzed. This is a foundation which sets the limit of a controllable set of parameters. In Section 5, we give a method to derive the parameters of VAR for a given set of statistics in bivariate timeseries. In Section 6, the mathematical analysis provided in this paper is summarized and a remark on the design principle of bivariate timeseries generated by VAR is added.

2. Vector Auto-Regression (VAR)

In theory, Granger causality (GC) is the transfer entropy of random variables in a bivariate vector auto-regression (VAR) model up to a constant factor of 2, if the VAR model has a stationary distribution [9]. Thus, it is straightforward to start with the bivariate VAR and derive its transfer entropy. In this way, we can derive a rich mathematical relationship between GC and the properties of VAR, rather than just a statistics of the bivariate timeseries.

Definition 1. *For some real vector $\mu \in \mathbb{R}^2$ and some positive-definite matrix $\Sigma \in \mathbb{R}^{2 \times 2}$, suppose the random variable ϵ_t for every integer $t = 0, 1, \ldots$ is drawn from the bivariate normal distribution $N(\epsilon_t | \mu, \Sigma)$ with mean μ and variance Σ. Define the initial vector by $v_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, and for $t \geq 0$ and a given coefficient matrix with real entries $A := \begin{pmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{pmatrix} \in \mathbb{R}^{2 \times 2}$, define the random variable v_t by*

$$v_{t+1} := A v_t + \epsilon_t. \tag{1}$$

Then, bivariate vector auto-regression is defined by the semi-infinite series of random variables $V = (v_0, v_1, v_2, \ldots)$.

In general, one can generate a timeseries $v_0, v_1, \ldots$ by fixing a set of the VAR parameters, the coefficient matrix A, and the base covariance matrix Σ, where the base mean vector μ is omitted as its effect is lost in the limit $t \to \infty$ when the VAR is stationary. The stationary correlation (covariance) $\hat{\Sigma}$ and Granger causality G_0 and G_1 defined later are the statistics of the timeseries generated by a VAR model (Figure 1). In what follows, we first explore the forward relationship of how the statistics $\hat{\Sigma}$ and G_0, G_1 are given by the VAR parameters (A, Σ). We then consider the backward relationship in which the VAR

parameters (A, Σ) suffice to generate a timeseries with given desired timeseries statistics $(G_0, G_1, \hat{\Sigma})$.
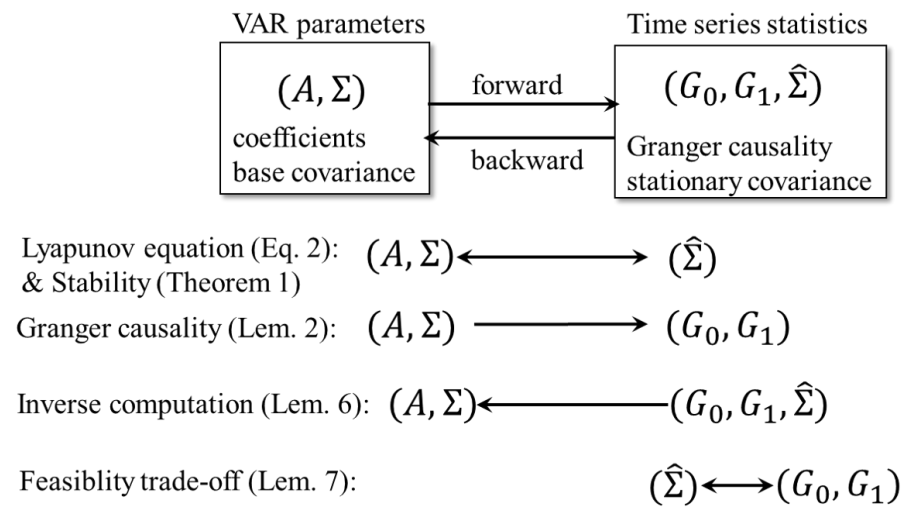

Figure 1. Schematic diagram of the organization of this paper.

2.1. Marginal Distribution of the VAR at Each Step

Lemma 1 (Marginal distribution of the VAR random variables at each step). *The VAR model with the initial vector $v_0 \in \mathbb{R}^2$ and the coefficient matrix $A \in \mathbb{R}^{2 \times 2}$ has the bivariate normal distribution*

$$N(v_t | \mu_t, \Sigma_t)$$

as its marginal distribution of the random variable v_t at each step $t = 0, 1, \ldots$, where

$$\mu_t := A^t v_0, \; \Sigma_t = \sum_{s=0}^{t} A^s \Sigma (A^s)^\top.$$

Proof. By Definition 1, Lemma 1 holds for $t = 0$. For $t + 1 > 0$, we prove Lemma 1 by assuming that it holds up to $t \geq 0$. By this assumption held for t, we have the distribution of $v_t \in \mathbb{R}^2$ as the bivariate normal distribution

$$N(v_t | \mu_t, \Sigma_t) = (2\pi)^{-1} |\Sigma_t|^{-\frac{1}{2}} e^{-\frac{1}{2}(v_t - \mu_t)^\top \Sigma_t^{-1} (v_t - \mu_t)}$$

with its mean μ_t and its covariance matrix Σ_t. Then, the random variable Av_t is distributed by the normal distribution

$$\begin{aligned} N(Av_t | A\mu_t, A\Sigma_t A^\top) &= (2\pi)^{-1} |A\Sigma_t A^\top|^{-\frac{1}{2}} \\ &\times e^{-\frac{1}{2}(A(v_t - \mu_t))^\top (A\Sigma_t A^\top)^{-1}(A(v_t - \mu_t))} \end{aligned} \tag{2}$$

with its mean $A\mu_t$ and the covariance matrix $A\Sigma_t A^\top$. The random variable ϵ_t is distributed by the following normal distribution:

$$N(\epsilon_t | 0, \Sigma) = (2\pi)^{-1} |\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}\epsilon_t^\top \Sigma^{-1} \epsilon_t}.$$

Thus, by VAR, Equation (1), we have the random variable

$$v_{t+1} := Av_t + \epsilon_t$$

which has a distribution calculated by the following integral:

$$P(v_{t+1}) = \int_{\epsilon_t \in \mathbb{R}^2} N(v_{t+1} - \epsilon_t | A\mu_t, A\Sigma_t A^\top) N(\epsilon_t | 0, \Sigma) \mathrm{d}\epsilon_t.$$

Calculating this, we have

$$P(v_{t+1}) = N\left(v_{t+1} | A\mu_t, A\Sigma_t A^\top + \Sigma\right). \tag{3}$$

Thus defining by $\mu_{t+1} := A\mu_t$ and $\Sigma_{t+1} = A\Sigma_t A^\top + \Sigma$, Lemma 1 holds for $t+1$. By expanding this, we have the Lemma 1 for any integer $t \geq 0$. $\square$

2.2. Stability of VAR: Lyapunov Equation

By Lemma 1, the mean and covariance matrix of the random variable at the tth step are

$$\mu_t = A^t v_0 \text{ and } \Sigma_t = \sum_{s=0}^{t} A^s \Sigma (A^s)^\top.$$

From this, we have the stationary distribution

$$\lim_{t \to \infty} N(v_t | \mu_t, \Sigma_t),$$

if and only if the absolute values of all the eigenvalues $\lambda_0, \lambda_1 \in \mathbb{C}$ of the coefficient matrix A are less than 1. If there is such a stationary distribution, we call the VAR *stable*, and its stationary distribution is the bivariate normal distribution

$$N(\hat{v} | \mathbf{0}_2, \hat{\Sigma}),$$

where the stationary mean vector $\hat{v} \in \mathbb{R}^2$ and stationary covariance matrix $\hat{\Sigma} \in \mathbb{R}^{2 \times 2}$ are defined as follows. If the VAR is stable, we have the following Lyapunov equation of the stationary covariance matrix $\hat{\Sigma} \in \mathbb{R}^{2 \times 2}$:

$$\hat{\Sigma} = \Sigma + A\hat{\Sigma}A^\top. \tag{4}$$

The Lyapunov equation is solved analytically by

$$\mathrm{vec}(\hat{\Sigma}) = (I_4 - A \otimes A)^{-1} \mathrm{vec}(\Sigma), \tag{5}$$

where $I_d \in \mathbb{R}^{d \times d}$ is the dth order identity matrix, $\otimes$ denotes the Kronecker product, and $\mathrm{vec}(X)$ for any matrix $X = (x_{i,j})_{i=1,\dots,n, j=1,\dots,m}$ is the vectorization operator $\mathrm{vec}() : \mathbb{R}^{n \times m} \to \mathbb{R}^{nm \times 1}$ defined by

$$\mathrm{vec}(X) := (x_{1,1}, x_{2,1}, \dots, x_{n,1}, \dots, x_{1,m}, x_{2,m}, \dots, x_{n,m})^\top.$$

The Lyapunov Equation (4) has the solution for $\hat{\Sigma}$ if the VAR is stable, but not vice versa. This is shown by Lemma 4.

3. Transfer Entropy and Granger Causality

In [9], the transfer entropy of an appropriate triplet of variables in the VAR model is shown to be equivalent to Granger causality up to the constant factor 2. Following this guide, we define this quantity as the Granger causality of the VAR model, below.

Although this relationship has been known in a more general form [9], we re-derive it for bivariate VAR in order to later analyze the structure of VAR and GCs in depth—for example, its upper and lower bounds (Lemma 3), stability (Section 4), and design principle (Section 5).

Definition 2. *If VAR with its random variables $v_t = (x_t, y_t)^\top \in \mathbb{R}^2$ for $t = 0, 1, \ldots$ is stable, transfer entropy from y to x is defined by*

$$T_{y \to x} := \lim_{t \to \infty} (H(x_{t+1}|x_t) - H(x_{t+1}|x_t, y_t)),$$

and the transfer entropy from x to y is defined by

$$T_{x \to y} := \lim_{t \to \infty} (H(y_{t+1}|y_t) - H(y_{t+1}|y_t, x_t)),$$

where the differential entropy of random variable x with its probability density function P is

$$H(x) := - \int_{x \in \Omega} P(x) \log P(x) \mathrm{d}x,$$

and the conditional entropy is

$$H(x|y) := H(x, y) - H(y).$$

In particular, we call two times of transfer entropy Granger causality denoted by

$$G_0 = 2T_{y \to x} \text{ and } G_1 = 2T_{x \to y}. \tag{6}$$

Specifically, GCs are specifically written by the terms of the VAR parameters in the following lemma.

Lemma 2 (Granger causality). *If a stable VAR has its covariance matrix, coefficient matrix, and stationary matrix*

$$\Sigma = \begin{pmatrix} \sigma_{0,0} & \sigma_{0,1} \\ \sigma_{1,0} & \sigma_{1,1} \end{pmatrix}, \ A = \begin{pmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{pmatrix}, \ \hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_{0,0} & \hat{\sigma}_{0,1} \\ \hat{\sigma}_{1,0} & \hat{\sigma}_{1,1} \end{pmatrix},$$

each Granger causality of this VAR for $i = 0, 1$ is

$$G_i = \log \left(1 + \frac{a_{i,1-i}^2 \det(\hat{\Sigma})}{\hat{\sigma}_{i,i} \sigma_{i,i}} \right). \tag{7}$$

Proof. In general, the differential entropy of multivariate normal distribution $N(v|\mu, \Sigma)$ is

$$H(v) = \frac{1}{2} \log |2\pi e \Sigma|,$$

where $e \approx 2.71$ is Napier's constant. For the joint probability distribution of $v_t = (x_t, y_t)^\top$

$$P(v_{t+1}|v_t) = N(v_{t+1}|Av_t, \Sigma),$$

the two marginal probability distributions of x_t, y_t are

$$P(x_{t+1}|v_t) = N(x_{t+1}|(1,0)Av_t, \sigma_{0,0}) \text{ and } P(y_{t+1}|v_t) = N(y_{t+1}|(0,1)Av_t, \sigma_{1,1}).$$

Thus, the conditional entropy of x_{t+1} and y_{t+1} given $v_t = (x_t, y_t)^\top$ are

$$H(x_{t+1}|x_t, y_t) = \frac{1}{2} \log |2\pi e \sigma_{0,0}| \text{ and } H(y_{t+1}|x_t, y_t) = \frac{1}{2} \log |2\pi e \sigma_{1,1}|.$$

With the conditional probability distribution and marginal probability distribution

$$P(v_{t+1}|v_t) = N(v_{t+1}|Av_t, \Sigma) \text{ and } P(v_t) = N(v_t|A^t v_0, \Sigma_t),$$

the joint probability distribution of v_t and v_{t+1} is their product

$$P(v_{t+1}, v_t) = N(v_{t+1}|Av_t, \Sigma)N(v_t|A^t v_0, \Sigma_t).$$

Specifically, this quad-variate normal distribution is

$$P(v_{t+1}, v_t) = e^{-\frac{1}{2}(v_{t+1}-Av_t)^\top \Sigma^{-1}(v_{t+1}-Av_t) - \frac{1}{2}(v_t - A^t v_0)^\top \Sigma_t^{-1}(v_t - A^t v_0)} (2\pi)^{-2} |\Sigma|^{-\frac{1}{2}} |\Sigma_t|^{-\frac{1}{2}}.$$

Applying the identities

$$v_{t+1} - Av_t = v_{t+1} - A^{t+1}v_0 - A(v_t - A^t v_0),$$

$$\Sigma_t' := \begin{pmatrix} \Sigma + A\Sigma_t A^\top & A\Sigma_t \\ \Sigma_t A^\top & \Sigma_t \end{pmatrix} = \begin{pmatrix} \Sigma^{-1} & -\Sigma^{-1}A \\ -A^\top \Sigma^{-1} & \Sigma_t^{-1} + A\Sigma^{-1}A^\top \end{pmatrix}^{-1},$$

and $|\Sigma||\Sigma_t| = |\Sigma_t'|$ to $P(v_{t+1}, v_t)$, we have

$$P(v_{t+1}, v_t) = N(v_t'|\mu_t', \Sigma_t'),$$

where

$$v_t' := \begin{pmatrix} v_{t+1} \\ v_t \end{pmatrix}, \quad \mu_t' := \begin{pmatrix} A^{t+1}v_0 \\ A^t v_0 \end{pmatrix}, \quad \Sigma_t' := \begin{pmatrix} \Sigma + A\Sigma_t A^\top & A\Sigma_t \\ \Sigma_t A^\top & \Sigma_t \end{pmatrix}.$$

From this joint probability distribution $P(v_{t+1}, v_t)$, we drive the marginal distributions

$$P(x_{t+1}, x_t) = N(x_{t+1}, x_t|\mu_{t,0}, \Sigma_{t,0}), \quad P(y_{t+1}, y_t) = N(y_{t+1}, y_t|\mu_{t,1}, \Sigma_{t,1}),$$

where for $i = 0, 1$ the mean vectors and covariance matrices are defined as follows:

$$\mu_{t,i} := \begin{pmatrix} e_i^\top A^{t+1}v_0 \\ e_i^\top A^t v_0 \end{pmatrix} = (I_2 \otimes e_i)\mu_t',$$

$$\Sigma_{t,i} := \begin{pmatrix} e_i^\top (\Sigma + A\Sigma_t A^\top)e_i & e_i^\top A\Sigma_t e_i \\ e_i^\top \Sigma_t A^\top e_i & e_i^\top \Sigma_t e_i \end{pmatrix} = (I_2 \otimes e_i)^\top \Sigma_t'(I_2 \otimes e_i),$$

with the unit vectors $e_0 := (1,0)^\top$, $e_1 := (0,1)^\top$.

Thus, we have the joint entropy of x_t and x_{t+1}

$$H(x_{t+1}, x_t) = \frac{1}{2}\log|2\pi e \Sigma_{t,0}| \tag{8}$$

$$= \frac{1}{2}\log(2\pi e)^2 |e_0^\top (\Sigma + A\Sigma_t A^\top)e_0 e_0^\top \Sigma_t e_0 - (e_0^\top A\Sigma_t e_0)^2|, \tag{9}$$

and the marginal distribution of x_t

$$H(x_t) = \frac{1}{2}\log 2\pi e|e_0^\top \Sigma_t e_0|. \tag{10}$$

Using these, we have the conditional entropy

$$H(x_{t+1}|x_t) = \frac{1}{2}\log 2\pi \frac{e_0^\top (\Sigma + A\Sigma_t A^\top)e_0 e_0^\top \Sigma_t e_0 - (e_0^\top A\Sigma_t e_0)^2}{|e_0^\top \Sigma_t e_0|}. \tag{11}$$

By the stability of VAR, the Lyapunov Equation (4) holds, and this conditional entropy in the limit $t \to \infty$ is

$$\lim_{t \to \infty} H(x_{t+1}|x_t) = \frac{1}{2}\log 2\pi \frac{(e_0^\top \hat{\Sigma} e_0)^2 - (e_0^\top A\hat{\Sigma} e_0)^2}{|e_0^\top \hat{\Sigma} e_0|}. \tag{12}$$

Applying Definition 2 and denoting by entries in the stationary covariance matrix $\hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_{0,0} & \hat{\sigma}_{0,1} \\ \hat{\sigma}_{1,0} & \hat{\sigma}_{1,1} \end{pmatrix}$, we have

$$G_0 = 2T_{y \to x} = \log \frac{(\hat{\sigma}_{0,0})^2 - (a_{0,0}\hat{\sigma}_{0,0} + a_{0,1}\hat{\sigma}_{1,0})^2}{\hat{\sigma}_{0,0}\sigma_{0,0}}. \tag{13}$$

Similarly, we have

$$G_1 = 2T_{x \to y} = \log \frac{(\hat{\sigma}_{1,1})^2 - (a_{1,0}\hat{\sigma}_{0,1} + a_{1,1}\hat{\sigma}_{1,1})^2}{\hat{\sigma}_{1,1}\sigma_{1,1}}. \tag{14}$$

Let us define for $i = 0, 1$

$$\delta_i := (\hat{\sigma}_{i,i})^2 - (a_{i,i}\hat{\sigma}_{i,i} + a_{i,1-i}\hat{\sigma}_{1-i,i})^2 \text{ and } \delta_i' := \hat{\sigma}_{i,i}\sigma_{i,i}. \tag{15}$$

By the Lyapunov equation, $\sigma_{i,i} = \hat{\sigma}_{i,i} - e_i^\top A\hat{\Sigma}A^\top e_i$. Applying this to δ_i', we have

$$\delta_i = \hat{\sigma}_{i,i}^2 - e_i^\top A \begin{pmatrix} \hat{\sigma}_{i,i}^2 & \hat{\sigma}_{i,i}\hat{\sigma}_{1-i,i} \\ \hat{\sigma}_{i,i}\hat{\sigma}_{1-i,i} & \hat{\sigma}_{1-i,i}^2 \end{pmatrix} A^\top e_i, \tag{16}$$

$$\delta_i' = \hat{\sigma}_{i,i}^2 - e_i^\top A \begin{pmatrix} \hat{\sigma}_{i,i}\hat{\sigma}_{0,0} & \hat{\sigma}_{i,i}\hat{\sigma}_{0,1} \\ \hat{\sigma}_{i,i}\hat{\sigma}_{1,0} & \hat{\sigma}_{i,i}\hat{\sigma}_{1,1} \end{pmatrix} A^\top e_i. \tag{17}$$

As $\delta_i - \delta_i' = a_{i,1-i}^2\det(\hat{\Sigma})$ and $G_0 = \log\left(1 + \frac{\delta_i - \delta_i'}{\delta_i'}\right)$, we have

$$G_i = \log\left(1 + \frac{a_{i,1-i}^2\det(\hat{\Sigma})}{\hat{\sigma}_{i,i}\sigma_{i,i}}\right).$$

□

The Granger causality has its lower and upper bounds in theory. Although these bounds may be further narrowed by considering the stability of the VAR, what follows below are the theoretical bounds regardless of the stability of the VAR.

Lemma 3 (The upper and lower bound for Granger causality). *For each $i = 0, 1$, Granger causality G_i has the following bounds:*

$$0 \leq G_i \leq \log \gamma_i, \tag{18}$$

where $\gamma_i := \frac{\hat{\sigma}_{i,i}}{\sigma_{i,i}} \geq 1$ due to the Lyapunov Equation (4). The lower bound $G_i = 0$ is given only if

$$a_{i,1-i}^2\det(\hat{\Sigma}) = 0. \tag{19}$$

The upper bound $G_i = \log \gamma_i$ is given only if

$$a_{i,i}\hat{\sigma}_{i,i} + a_{i,1-i}\hat{\sigma}_{1-i,i} = 0. \tag{20}$$

Proof. As the stationary covariance matrix is (semi-)positive definite, $\det(\hat{\Sigma}) \geq 0$. Thus, the lower bound of Granger causality is $G_i \geq \log(1) = 0$ and this bound is only reacheable when $a_{i,1-i}^2\det(\hat{\Sigma}) = 0$.

Modifying (13) and (14), for $i = 0, 1$ we have

$$(a_{i,i}\hat{\sigma}_{i,i} + a_{i,1-i}\hat{\sigma}_{1-i,i})^2 = \hat{\sigma}_{i,i}\left(\hat{\sigma}_{i,i} - \sigma_{i,i}e^{G_i}\right). \tag{21}$$

As $(a_{i,i}\hat{\sigma}_{i,i} + a_{i,1-i}\hat{\sigma}_{1-i,i})^2 \geq 0$ and $\hat{\sigma}_{i,i} > 0$,

$$G_i \leq \log \gamma_i.$$

This upper bound holds only if $a_{i,i}\hat{\sigma}_{i,i} + a_{i,1-i}\hat{\sigma}_{1-i,i} = 0$. □

The upper bound Lemma 3 can also be obtained by the following information-theoretic identity:

$$\lim_{t \to \infty} I(x_{t-1}; x_t) + I(x_t; x_{t-1}|x_{t-1}) = \lim_{t \to \infty} I(x_t; x_{t-1}, x_{t-1}),$$

where $\lim_{t \to \infty} I(x_t; x_{t-1}|x_{t-1}) = \frac{1}{2}G_0$ is the transfer entropy, $\lim_{t \to \infty} I(x_t; x_{t-1}, x_{t-1}) = \frac{1}{2}\log\frac{\hat{\sigma}_{0,0}}{\sigma_{0,0}}$, and

$$\lim_{t \to \infty} I(x_{t-1}; x_t) = \frac{1}{2}\log \frac{\hat{\sigma}_{0,0}^2}{\begin{vmatrix} \hat{\Sigma} & \hat{\Sigma}A^\top \\ A\hat{\Sigma} & \hat{\Sigma} \end{vmatrix}} = \frac{1}{2}(\log \gamma_0 - G_0).$$

4. Stability and Constraints of VAR

In this study, we primarily consider the class of stable VAR models with a proper set of parameters. In this class, the statistical nature of any VAR is characterized by the base covariance matrix $\Sigma \in \mathbb{R}^{2\times2}$, coefficient matrix $A \in \mathbb{R}^{2\times2}$, and stationary covariance matrix $\hat{\Sigma} \in \mathbb{R}^{2\times2}$. Let us denote the set of (strictly) positive definite matrices by

$$\mathbb{R}_+^{2\times2} := \left\{ M \in \mathbb{R}^{2\times2} \middle| \det(M) > 0 \text{ and } \mathrm{tr}(M) > 0 \right\},$$

and the set of coefficient matrices of stable VAR models

$$\mathbb{R}_*^{2\times2} := \left\{ M \in \mathbb{R}^{2\times2} \middle| -1 + |\mathrm{tr}(M)| < \det(M) < 1 \right\}.$$

We will briefly show that the stable set $\mathbb{R}_*^{2\times2}$ includes all and only coefficient matrices of stable bivariate VAR models.

With this notation of the set of matrices, the two conditions that any proper VAR model needs to satisfy are as follows.

Stability Any stable VAR model has both of the eigenvalues λ_0, λ_1 of its coefficient matrix A meeting $|\lambda_0|, |\lambda_1| < 1$.

Properness To have a proper (non-degenerated) bivariate normal distribution in a VAR model, its base covariance matrix Σ and stationary covariance matrix $\hat{\Sigma}$ need to satisfy $\Sigma, \hat{\Sigma} \in \mathbb{R}_+^{2\times2}$. The set of positive-definite matrices is equivalently written with the entries of the following matrix:

$$\mathbb{R}_+^{2\times2} = \left\{ C \in \mathbb{R}^{2\times2} \mid C_{0,0} > 0, C_{1,1} > 0, \text{ and } C_{0,0}C_{1,1} - C_{0,1}C_{1,0} > 0 \right\}. \quad (22)$$

4.1. Stability of VAR

As stated previously in Section 2.2, the stability of VAR is primarily characterized by the eigenvalues of the coefficient matrix A. However, this condition is equivalent to $A \in \mathbb{R}_*^{2\times2}$, as shown by the following lemma.

Lemma 4. *A given bivariate VAR model with its coefficient matrix $A \in \mathbb{R}^{2 \times 2}$ is stable if and only if*

$$|\text{tr}(A)| - 1 < \det(A) < 1. \tag{23}$$

Proof. Let λ be an eigenvalue of the coefficient matrix A. Such an eigenvalue then satisfies

$$f(\lambda) = |A - \lambda I_2| = \lambda^2 - \text{tr}(A)\lambda + \det(A) = 0. \tag{24}$$

If a VAR is stable, this eigenvalue needs to satisfy $|\lambda| < 1$. As (24) is rewritten by

$$f(\lambda) = \left(\lambda - \frac{1}{2}\text{tr}(A)\right)^2 - \frac{1}{4}\left(\text{tr}(A)^2 - 4\det(A)\right), \tag{25}$$

we analyze this condition on (24) for the following two cases with λ being real or non-real:

1. If λ is real, this stability condition is equivalent to

$$\text{tr}(A)^2 \le 4\det(A), f(1) > 0, \ f(-1) > 0, \ |\text{tr}(A)| < 2. \tag{26}$$

2. If λ is not real, this stability condition is equivalent with

$$\text{tr}(A)^2 < 4\det(A), \ |\lambda|^2 < 1. \tag{27}$$

If λ of (24) is non-real (Case 2), λ (and its conjugate) is

$$\lambda = \frac{1}{2}\text{tr}(A) \pm \frac{j}{2}\sqrt{|\text{tr}(A)^2 - 4\det(A)|}, \tag{28}$$

with the imaginary unit denoted by j.

With the inequality (27), the stability condition in this case is

$$\left(\frac{\text{tr}(A)}{2}\right)^2 < |\lambda|^2 = \det(A) < 1. \tag{29}$$

If λ of (24) is real, $\text{tr}(A)^2 - 4\det(A) \ge 0$ and

$$f(1) = 1 - \text{tr}(A) + \det(A) > 0 \tag{30}$$
$$f(-1) = 1 + \text{tr}(A) + \det(A) > 0 \tag{31}$$
$$|\text{tr}(A)| < 2. \tag{32}$$

Combining (30) and (31), we have $|\text{tr}(A)| - 1 < \det(A)$. This inequality with (26),

$$C_0 < \det(A) \le \left(\frac{\text{tr}(A)}{2}\right)^2 < C_1, \tag{33}$$

where $C_0 := |\text{tr}(A)| - 1$ and $C_1 := \min\left(1, \left(\frac{1+\det(A)}{2}\right)^2\right)$. Find for an arbitrary $A \in \mathbb{R}^{2 \times 2}$ we have the following two inequalities:

$$|\text{tr}(A)| - 1 \le \left(\frac{1}{2}\text{tr}(A)\right)^2, \tag{34}$$

and

$$\det(A) \le \left(\frac{1 + \det(A)}{2}\right)^2. \tag{35}$$

The inequality (34) holds equality for and only for $\text{tr}(A) = 2$, and the inequality (34) holds equality for and only for $\det(A) = 1$. As both of these equality conditions do not hold under (29), (29) is equivalent to

$$C_0 < \left(\frac{\text{tr}(A)}{2}\right)^2 < \det(A) < C_1. \tag{36}$$

Integrating the two inequalities (33) for real λ and (36) for non-real λ, the VAR with the coefficient matrix A is stable if

$$C_0 < \det(A) < C_1 \tag{37}$$

and

$$C_0 < \left(\frac{\text{tr}(A)}{2}\right)^2 < C_1. \tag{38}$$

As the inequality (37) implies $0 < \frac{1+\det(A)}{2} < 1$ and $\text{tr}(A) < 2$, (38) is equivalent to

$$\left(\frac{\text{tr}(A)}{2}\right)^2 < \left(\frac{1+\det(A)}{2}\right)^2. \tag{39}$$

As the upper bound for $\det(A)$ in (37) can be implied by $\det(A) < 1$, it is equivalent to

$$|\text{tr}(A)| - 1 < \det(A) < 1. \tag{40}$$

Thus, the pair of inequalities (37) and (38) for A is equivalent to the single inequality (40) for A. $\square$

4.2. Stability and Existence of the Solution for the Lyapunov Equation

Intuitively, it would be reasonable if there was a stationary covariance matrix $\hat{\Sigma} \in \mathbb{R}_+^{2\times2}$ satisfying the Lyapunov Equation (4), if the coefficient matrix is $A \in \mathbb{R}_*^{2\times2}$. However, this is not trivial, as the opposite may not be always true: the existence of $\hat{\Sigma} \in \mathbb{R}_+^{2\times2}$ does not imply $A \in \mathbb{R}_*^{2\times2}$. This relationship between A and $\hat{\Sigma}$ is stated by the following Theorem 1.

Theorem 1. *There is a stationary covariance matrix $\hat{\Sigma} \in \mathbb{R}_+^{2\times2}$ satisfying the Lyapunov Equation (4), if the coefficient matrix is $A \in \mathbb{R}_*^{2\times2}$. However, the existence of $\hat{\Sigma} \in \mathbb{R}_+^{2\times2}$ does not imply $A \in \mathbb{R}_*^{2\times2}$.*

Proof. Find the identity

$$
\begin{aligned}
\det(I_4 - A \otimes A) &= \det\big((I_2 - a_{0,0}A)(I_2 - a_{1,1}A) - a_{0,1}a_{1,0}A^2\big) \\
&= \det\big(I_2 - A\text{tr}(A) + A^2\det(A)\big) \\
&= (1 - \det(A))^2\big((1 - a_{0,1} - a_{0,0}^2)(1 - a_{0,1} - a_{1,1}^2) - a_{0,1}a_{1,0}\text{tr}(A)\big) \\
&= (1 - \det(A))^2\big((1 + \det(A))^2 - \text{tr}(A)^2\big).
\end{aligned}
\tag{41}
$$

By Lemma 4 and (41), we have $\det(I_4 - A \otimes A) > 0$. Thus, Lyapunov Equation (5) has the solution for $\hat{\Sigma}$, as the matrix $(I_4 - A \otimes A)$ is invertible. The converse of this theorem does not hold, as we construct a counter-example of the coefficient matrix A such that $\det(I_4 - A \otimes A) < 0$, with which there is a $\hat{\Sigma} \in \mathbb{R}_+^{2\times2}$, but such a VAR is not stable. $\square$

5. Design of Bivariate Timeseries Given GCs

The goal of this study was to derive a design principle of bivariate timeseries generated by a VAR model with the desired correlation and two types of Granger causality. In this section, we explore the inter-dependent relationships among the variables in the VAR. This analysis revealed a trade-off limitation in designing these variables of timeseries.

Specifically, a timeseries with a certain range of desired Granger causality cannot be realized by a stable VAR, in which no stationary covariance is defined in theory.

The set of parameters in any stable VAR model includes

- The coefficient matrix A;
- The base covariance matrix Σ;
- The stationary covariance matrix $\hat{\Sigma}$; and
- The two types of Granger causality G_0, G_1.

There are equality constraints on these variables:

- The variables $A, \Sigma, \hat{\Sigma}$ need to satisfy the Lyapunov Equation (4).
- Granger causality G_i ($i = 0, 1$) is the function of $a_{i,1-i}$, $\sigma_{i,i}$, and $\hat{\Sigma}$ (Lemma 2).

Besides, it is important to know the feasibility of a set of parameters in VAR, which constrains the range of these variables:

- Stability: $A \in \mathbb{R}_*^{2\times2}$ (Section 4);
- Properness: $\Sigma, \hat{\Sigma} \in \mathbb{R}_+^{2\times2}$ (Section 4) and $\sigma_{i,i} \leq \hat{\sigma}_{i,i}$ due to the existence of a solution for the Lyapunov equation; and
- The bound for each Granger causality: $G_i \in [0, \log \gamma_i]$ (Lemma 3).

The Lyapunov Equation (4) on the matrices can be decomposed into the three equations on the scalar variables as follows. For a coefficient matrix $A = \begin{pmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{pmatrix}$, let us define two vectors by

$$a_0 := \begin{pmatrix} a_{0,0} \\ a_{0,1} \end{pmatrix}, \; a_1 := \begin{pmatrix} a_{1,0} \\ a_{1,1} \end{pmatrix}.$$

The Lyapunov equation is then equivalently written with these vectors a_0, a_1 by the set of the three equations

$$\hat{\sigma}_{0,0} - \sigma_{0,0} = a_0^\top \hat{\Sigma} a_0 \tag{42}$$

$$\hat{\sigma}_{1,1} - \sigma_{1,1} = a_1^\top \hat{\Sigma} a_1 \tag{43}$$

$$\hat{\sigma}_{0,1} - \sigma_{0,1} = a_0^\top \hat{\Sigma} a_1. \tag{44}$$

Equations (42) and (43) above imply that each of the vectors a_0 and a_1 are on an ellipsis on each of their planes. This gives the lower bound for $\hat{\sigma}_{i,i} \geq \sigma_{i,i}$ (i.e., one condition of the properness above), as $x^\top \hat{\Sigma} x \geq 0$ for any $x \in \mathbb{R}^2$ with a positive-definite matrix $\hat{\Sigma}$.

Fixing G_0 and G_1 imposes each of the two vectors a_0 and a_1 on the two parallel lines by

$$(a_i^\top \hat{\sigma}_i)^2 = \tau_i^2, \tag{45}$$

where

$$\hat{\sigma}_i := (\hat{\sigma}_{i,0}, \hat{\sigma}_{i,1})^\top, \; \tau_i^2 := \hat{\sigma}_{i,i}^2 \left(1 - \gamma_i^{-1} e^{G_i} \right).$$

Thus, the solution of a_i which satisfies the Lyapunov equation and the fixed Granger causality is the four intersections of the ellipsis and the two parallel lines (Figure 2). This ellipsis is obtained by scaling and shearing transformation to the standard circle $a_{0,0}^2 + a_{0,1}^2 = 1$. This observation gives the angular parametrization of the solution vector $(a_{0,0}, a_{0,1})^\top$, which is explicitly stated by Lemma 5 in the next section.
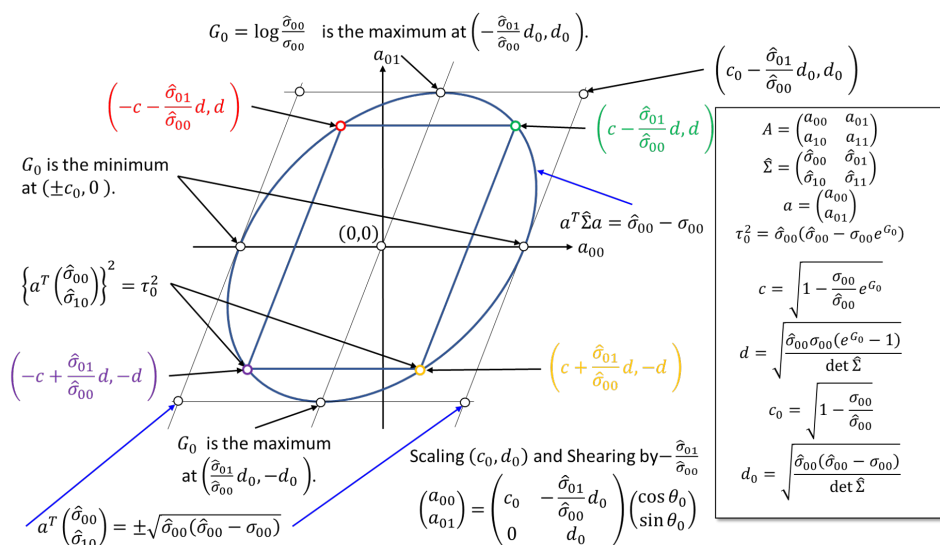
Figure 2. The ellipsis (42) and two parallel lines (45) (forming the parallelogram touching the ellipsis) on the plane $(a_{0,0}, a_{0,1}) \in \mathbb{R}^2$. The solution (a_0, a_1) is four intersections of these two (depicted by the colored points). Granger causality takes its maximum with the largest $|a_{0,1}|$ on the ellipsis and its minimum with $|a_{0,1}| = 0$.

5.1. Solution A of the Lyapunov Equality Given $\hat{\Sigma}$, G_0, and G_1

In what follows, we start with the derivation of the coefficient matrix A as a root of the equality constraint by the Lyapunov Equation (4) and the Granger causality, for a fixed proper $\hat{\Sigma}$, $\sigma_{i,i}$ and G_i for each $i = 0, 1$. The following Lemma 5 gives a necessary condition for the coefficient matrix $A \in \mathbb{R}^{2 \times 2}$ to satisfy the equality conditions above. Note, however, that such a solution A in this equation does not guarantee the stability of the corresponding VAR (i.e., $A \in \mathbb{R}_*^{2 \times 2}$). This sufficiency is explored in Section 5.2.

Lemma 5. *For a given set of parameters, a positive-definite matrix* $\hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_{0,0} & \hat{\sigma}_{0,1} \\ \hat{\sigma}_{1,0} & \hat{\sigma}_{1,1} \end{pmatrix} \in \mathbb{R}_+^{2 \times 2}$,
$\sigma_{i,i} \in (0, \hat{\sigma}_{i,i}), G_i \in [0, \log \gamma_i]$ *for each* $i = 0, 1$, *suppose that a coefficient matrix*
$A = \begin{pmatrix} a_{0,0} & a_{0,1} \\ a_{1,0} & a_{1,1} \end{pmatrix} \in \mathbb{R}^{2 \times 2}$, *satisfies the set of the equations*

$$
\begin{cases}
a_0^\top \hat{\Sigma} a_0 = \hat{\sigma}_{0,0} - \sigma_{0,0} \\
a_1^\top \hat{\Sigma} a_1 = \hat{\sigma}_{1,1} - \sigma_{1,1} \\
(\hat{\sigma}_0^\top a_0)^2 = \tau_0^2 \\
(\hat{\sigma}_1^\top a_1)^2 = \tau_1^2
\end{cases}, \tag{46}
$$

where for $i = 0, 1$

$$
\hat{\sigma}_i := (\hat{\sigma}_{i,0}, \hat{\sigma}_{i,1})^\top, \ \tau_i^2 := \hat{\sigma}_{i,i}^2 \left(1 - \gamma_i^{-1} e^{G_i} \right).
$$

Any coefficient matrix A of a root of this Equation (46) is in the form

$$
A = \left(S_0 \begin{pmatrix} \cos \theta_0 \\ \sin \theta_0 \end{pmatrix} \quad P_2 S_1 \begin{pmatrix} \cos \theta_1 \\ \sin \theta_1 \end{pmatrix} \right)^\top, \tag{47}
$$

where each pair of the angles $\theta_0 \in [0, 2\pi)$ *and* $\theta_1 \in [0, 2\pi)$ *takes one of the two or four pairs satisfying for each* $i = 0, 1$

$$
\sin^2 \theta_i = \frac{e^{G_i} - 1}{\gamma_i - 1} \tag{48}
$$

and

$$P_2 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_i := \sqrt{1 - \gamma_i^{-1}} \begin{pmatrix} 1 & -\frac{\hat{\sigma}_{i,1-i}}{\hat{\sigma}_{i,i}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{\hat{\sigma}_{i,i}}{\sqrt{\det(\hat{\Sigma})}} \end{pmatrix}.$$

Proof. Find that the following pair of equations in (46) is symmetric under exchange of $i = 0, 1$:

$$\begin{cases} a_i^\top \hat{\Sigma} a_i = \hat{\sigma}_{i,i} - \sigma_{i,i} \\ (\hat{\sigma}_i^\top a_i)^2 = \tau_i^2 \end{cases}. \tag{49}$$

Thus, we solve this for $i = 0$ below, and it holds for $i = 1$.

Solving the second equation of (49) for $a_{0,0}$, we have

$$a_{0,0} = \frac{\pm \tau_0 - a_{0,1} \hat{\sigma}_{0,1}}{\hat{\sigma}_{0,0}}. \tag{50}$$

Inserting this into the first equation of (49), we have

$$a_{0,1}^2 = \frac{\hat{\sigma}_{0,0}^2 (1 - \gamma_0^{-1}) - \tau_0^2}{\det(\hat{\Sigma})}. \tag{51}$$

Inserting $\tau_0^2 = \hat{\sigma}_{0,0}^2 \left(1 - \gamma_0^{-1} e^{G_0}\right)$, we have

$$a_{0,1} = \pm \sqrt{\frac{\hat{\sigma}_{0,0} \sigma_{0,0} (e^{G_0} - 1)}{\det(\hat{\Sigma})}}. \tag{52}$$

Inserting this into (50), we have at most four vectors $a_0 = (a_{0,0}, a_{0,1})^\top$ as the solution of (49) for $i = 1$:

$$a_0 = \begin{pmatrix} c_0 - \frac{\hat{\sigma}_{0,1}}{\hat{\sigma}_{0,0}} d_0 \\ d_0 \end{pmatrix}, \begin{pmatrix} c_0 + \frac{\hat{\sigma}_{0,1}}{\hat{\sigma}_{0,0}} d_0 \\ -d_0 \end{pmatrix}, \begin{pmatrix} -c_0 - \frac{\hat{\sigma}_{0,1}}{\hat{\sigma}_{0,0}} d_0 \\ d_0 \end{pmatrix}, \begin{pmatrix} -c_0 + \frac{\hat{\sigma}_{0,1}}{\hat{\sigma}_{0,0}} d_0 \\ -d_0 \end{pmatrix}, \tag{53}$$

where for $i = 0, 1$

$$c_i := \sqrt{1 - \gamma_i^{-1} e^{G_i}}, \quad d_i := \sqrt{\frac{\hat{\sigma}_{i,i} \sigma_{i,i} (e^{G_i} - 1)}{\det(\hat{\Sigma})}}.$$

By symmetry to $i = 0, 1$, there are at most four vectors as the solution of (49) for $i = 1$:

$$P_2 a_1 = \begin{pmatrix} a_{1,1} \\ a_{1,0} \end{pmatrix} = \begin{pmatrix} c_1 - \frac{\hat{\sigma}_{1,0}}{\hat{\sigma}_{1,1}} d_1 \\ d_1 \end{pmatrix}, \begin{pmatrix} c_1 + \frac{\hat{\sigma}_{1,0}}{\hat{\sigma}_{1,1}} d_1 \\ -d_1 \end{pmatrix}, \begin{pmatrix} -c_1 - \frac{\hat{\sigma}_{1,0}}{\hat{\sigma}_{1,1}} d_1 \\ d_1 \end{pmatrix}, \begin{pmatrix} -c_1 + \frac{\hat{\sigma}_{1,0}}{\hat{\sigma}_{1,1}} d_1 \\ -d_1. \end{pmatrix} \tag{54}$$

Find that these four solution vectors parameterized by

$$a_0 = S_0 \begin{pmatrix} \cos \theta_0 \\ \sin \theta_0 \end{pmatrix} \quad \text{and} \quad a_1 = P_2 S_1 \begin{pmatrix} \cos \theta_1 \\ \sin \theta_1 \end{pmatrix},$$

satisfy (49), if (θ_0, θ_1) holds (48), with the trigonometric identity $\cos^2 \theta_i + \sin^2 \theta_i = 1$ and

$$S_0^\top \hat{\Sigma} S_0 = (\hat{\sigma}_{0,0} - \sigma_{0,0}) I_2 \quad \text{and} \quad S_1^\top P_2^\top \hat{\Sigma} P_2 S_1 = (\hat{\sigma}_{1,1} - \sigma_{1,1}) I_2. \tag{55}$$

□

5.2. Sufficiency of the Solution

For a given set of parameters, Lemma 5 in the previous section gives a set of solutions of the coefficient matrix A for the Lyapunov equation. Note that not all of these solutions A are feasible, in the sense that they satisfy all constraints such as the stability of $A \in \mathbb{R}_*^{2 \times 2}$

and the properness of $\Sigma \in \mathbb{R}_+^{2\times 2}$, in which Σ can be derived from Lyapunov Equation (4) given A and $\hat{\Sigma}$. The following lemmas provide the sufficient condition for a solution A by checking the properness of Σ and the stability of A.

Lemma 6. *Suppose A is a solution of Equation (46) in Lemma 5, represented by a pair of (θ_0, θ_1). In this case, $\Sigma \in \mathbb{R}_+^{2\times 2}$, if and only if*

$$\cos\hat{\eta} - (\gamma_0\gamma_1)^{-\frac{1}{2}} \leq \hat{\gamma}_0\hat{\gamma}_1\cos(\hat{\eta} - \theta_0 - \theta_1) \leq \cos\hat{\eta} + (\gamma_0\gamma_1)^{-\frac{1}{2}}, \tag{56}$$

where $\hat{\eta} \in [0, 2\pi]$ is the angler parametrization of the correlation coefficient defined by $\cos\hat{\eta} := \frac{\hat{\sigma}_{0,1}}{\sqrt{\hat{\sigma}_{0,0}\hat{\sigma}_{1,1}}}$ and $\hat{\gamma}_i = \sqrt{1 - \gamma_i^{-1}}$.

Proof. Applying the polar representation of the a_0, a_1 in (47) in Lemma 5 to the third Equation (44) of the Lyapunov equation, we have

$$\hat{\sigma}_{0,1} - \sigma_{0,1} = \sqrt{(\hat{\sigma}_{0,0} - \sigma_{0,0})(\hat{\sigma}_{1,1} - \sigma_{1,1})}\cos(\hat{\eta} - \theta_0 - \theta_1). \tag{57}$$

By the positive definiteness of Σ, $\sigma_{0,1}^2 \leq \sigma_{0,0}\sigma_{1,1}$. This inequality applied to (57) gives the lemma. □

If we have

$$\hat{\gamma}_0\hat{\gamma}_1 \leq \cos\hat{\eta} + (\gamma_0\gamma_1)^{-\frac{1}{2}} \text{ and } \cos\hat{\eta} - (\gamma_0\gamma_1)^{-\frac{1}{2}} \leq -\hat{\gamma}_0\hat{\gamma}_1,$$

Equation (56) holds for any pair of angles (θ_0, θ_1). This condition is

$$|\cos\hat{\eta}| \leq (\gamma_0\gamma_1)^{-\frac{1}{2}} - \hat{\gamma}_0\hat{\gamma}_1,$$

or

$$|\hat{\sigma}_{0,1}| \leq \sqrt{\sigma_{0,0}\sigma_{1,1}} - \sqrt{(\hat{\sigma}_{0,0} - \sigma_{0,0})(\hat{\sigma}_{1,1} - \sigma_{1,1})}. \tag{58}$$

On the other hand, (56) holds for any $0 < \hat{\eta} < \pi$ (equivalently $-1 < \cos\hat{\eta} < 1$) if

$$\frac{\sqrt{\gamma_0\gamma_1} - 1}{\sqrt{(\gamma_0 - 1)(\gamma_1 - 1)}} \leq \cos(\theta_0 + \theta_1) \leq \frac{\sqrt{\gamma_0\gamma_1} + 1}{\sqrt{(\gamma_0 - 1)(\gamma_1 - 1)}}. \tag{59}$$

These bounds (58) and (59) mean that the range of feasible GCs (θ_0, θ_1) and the range of feasible correlation $\cos\hat{\eta}$ are in a trade-off relationship in general.

Lemma 7. *The VAR with the correlation $|\cos(\hat{\eta})| < 1$ and Granger causality θ_0, θ_1 in the angular form is stable if and only if*

$$-\sin\hat{\eta} + |\hat{\gamma}_0\sin(\hat{\eta} - \theta_0) + \hat{\gamma}_1\sin(\hat{\eta} - \theta_1)| < \hat{\gamma}_0\hat{\gamma}_1\sin(\hat{\eta} - \theta_0 - \theta_1) < \sin\hat{\eta}. \tag{60}$$

Proof. Using the angular notation of the solution A with θ_0, θ_1, we have

$$\det(A) = \hat{\gamma}_0\hat{\gamma}_1\frac{\sin(\hat{\eta} - \theta_0 - \theta_1)}{\sin\hat{\eta}} \text{ and } \text{tr}(A) = \frac{\hat{\gamma}_0\sin(\hat{\eta} - \theta_0) + \hat{\gamma}_1\sin(\hat{\eta} - \theta_1)}{\sin\hat{\eta}}.$$

Inserting these into stability condition (23), we have the inequality (60). □

As well as Lemma 6, Lemma 7 leads the trade-off relationship between $\hat{\Sigma}$ in the angle from $\hat{\eta}$ and A in the angle from θ_0, θ_1. In general, this bound further narrows the upper and lower bounds given by Lemma 3. In general, correlation is limited to close to zero if one wishes for higher Granger causality. On the other hand, the two types of Granger causality are limited to close to zero if one wishes for a higher correlation in the absolute value.

6. Concluding Remarks

6.1. Summary and Potential Usage of the Algorithm

In this paper, we explored the relationship between the VAR parameters and timeseries statistics (Figure 1), and identified the trade-off limitation between the stationary covariance $\hat{\Sigma}$ and Granger causality G_0, G_1 (Lemma 6). This suggests that the following Algorithm 1 will generate a timeseries with desired statistics.

Algorithm 1: Compute a VAR parameter set for the desired statistics

---
Data: Desired timeseries statistics $(G_0, G_1, \hat{\Sigma}, \sigma_{0,0}, \sigma_{1,1})$ in the feasible range satisfy both inequalities (56) in Lemma 6 and (60) in Lemma 7.

1 Derive four sets of VAR parameters (A, Σ) for the given timeseries statistics $(G_0, G_1, \hat{\Sigma}, \sigma_{0,0}, \sigma_{1,1})$ by Lemma 5

2 Choose one of the four sets of VAR parameters (A, Σ).

Result: The VAR parameters (A, Σ).

This data-generation algorithm can be used to generate surrogate data [19], which can be used to test whether an empirical timeseries is a sample from a VAR with a given correlation and Granger causality. This algorithm is also useful in analyzing to what extent a class of VAR timeseries varies under the same statistics.

6.2. Validity of Granger Causality Estimated on Empirical Timeseries

Our analysis also warns that not all Granger causality (or transfer entropy) is "valid", in the sense that its underlying VAR model is not stable and thus the Granger causality is undefined in theory. In theory, we can identify some value of Granger causality for a finite empirical time series, which is generated by an underlying unstable VAR model without any stationary statistics. Such timeseries statistics will diverge in the long run, but it may be difficult to identify this with a finite empirical timeseries. This asymmetry— namely, that the Granger causality can be calculated numerically but does not guarantee the stability of the underlying VAR—is explicitly demonstrated by Theorem 1. For a given empirical timeseries $v = (v_0, v_1, \dots, v_T) \in \mathbb{R}^{2 \times (T+1)}$, one should calculate not just the Granger causality but also its validity by checking (1) $A \in \mathbb{R}_*^{2 \times 2}$, (2) $\Sigma \in \mathbb{R}_+^{2 \times 2}$, and (3) $\Sigma_{0,1} \le \hat{\Sigma}_{0,1}$, by calculating the maximum likelihood estimator of $(A(v), \Sigma(v))$, such as

$$A(v) := V_{1,0} V_{0,0}^{-1} \text{ and } \Sigma(v) = V_{1,1} - V_{1,0} V_{0,0}^{-1} V_{0,1},$$

where

$$V_{i,j} := T^{-1} \sum_{t=1}^{T} v_{t-1+i} v_{t-1+j}^{\top}.$$

In fact, $\Sigma(v)$ is always (semi-)positive definite, as it takes the form of the Schur complement of the (semi-)positive-definite matrix $\begin{pmatrix} V_{0,0} & V_{0,1} \\ V_{1,0} & V_{1,1} \end{pmatrix}$. Thus, this maximum likelihood estimator readily satisfies condition (2).

6.3. Future Work

In this paper, we limit the VAR model to be bivariate for simplicity of analysis. We expect it is possible to generalize the current result to any higher dimensional VAR model. In such a generalization, feasible boundaries for the stable VAR models may require further effort to understand.

Author Contributions: Conceptualization, S.H. and T.T.; Methodology, S.H. and T.T.; Software, S.H.; Validation, S.H. and T.T.; Formal Analysis, S.H.; Investigation, S.H.; Writing—Original Draft Preparation, S.H.; Writing—Review and Editing, S.H. amd T.T.; Visualization, S.H.; Supervision, S.H.; Project Administration, S.H.; Funding Acquisition, S.H. Both authors have read and agreed to the published version of the manuscript.

References

1. Heider, F.; Simmel, M. An experimental study of apparent behavior. *Am. J. Psychol.* **1944**, *57*, 243–259. [CrossRef]
2. Bassili, J.N. Temporal and spatial contingencies in the perception of social events. *J. Personal. Soc. Psychol.* **1976**, *33*, 680–685. [CrossRef]
3. Dittrich, W.H.; Lea, S.E. Visual perception of intentional motion. *Perception* **1994**, *23*, 253–268. [CrossRef] [PubMed]
4. Tremoulet, P.D.; Feldman, J. Perception of animacy from the motion of a single object. *Perception* **2000**, *29*, 943–951. [CrossRef] [PubMed]
5. Scholl, B.J.; Tremoulet, P.D. Perceptual causality and animacy. *Trends Cogn. Sci.* **2000**, *4*, 299–308. [CrossRef]
6. Takahashi, K.; Watanabe, K. Synchronous motion modulates animacy perception. *J. Vis.* **2015**, *15*, 1–17. [CrossRef] [PubMed]
7. Granger, C.W. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **1969**, *37*, 424–438. [CrossRef]
8. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **2000**, *85*, 461–464. [CrossRef] [PubMed]
9. Barnett, L.; Barrett, A.B.; Seth, A.K. Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys. Rev. Lett.* **2009**, *103*, 238701. [CrossRef] [PubMed]
10. Eichler, M. Granger causality and path diagrams for multivariate time series. *J. Econom.* **2007**, *137*, 334–353. [CrossRef]
11. Freeman, J.R. Granger causality and the times series analysis of political relationships. *Am. J. Political Sci.* **1983**, pp. 327–358. [CrossRef]
12. Joerding, W. Economic growth and defense spending: Granger causality. *J. Dev. Econ.* **1986**, *21*, 35–40. [CrossRef]
13. Seth, A.K. Measuring autonomy and emergence via Granger causality. *Artif. Life* **2010**, *16*, 179–196. [CrossRef] [PubMed]
14. Okazaki, S.; Hirotani, M.; Koike, T.; Bosch-Bayard, J.; Takahashi, H.K.; Hashiguchi, M.; Sadato, N. Unintentional interpersonal synchronization represented as a reciprocal visuo-postural feedback system: a multivariate autoregressive modeling approach. *PLoS ONE* **2015**, *10*, e0137126. [CrossRef] [PubMed]
15. Bressler, S.L.; Seth, A.K. Wiener–Granger causality: A well established methodology. *Neuroimage* **2011**, *58*, 323–329. [CrossRef] [PubMed]
16. Ding, M.; Chen, Y.; Bressler, S.L. 17 Granger causality: Basic theory and application to neuroscience. In *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*; Wiley Online Library: Hoboken, NJ, USA, 2006; Volume 437.
17. Porta, A.; Faes, L. Wiener–Granger causality in network physiology with applications to cardiovascular control and neuroscience. *Proc. IEEE* **2015**, *104*, 282–309. [CrossRef]
18. Seth, A.K.; Barrett, A.B.; Barnett, L. Granger causality analysis in neuroscience and neuroimaging. *J. Neurosci.* **2015**, *35*, 3293–3297. [CrossRef] [PubMed]
19. Prichard, D.; Theiler, J. Generating surrogate data for time series with several simultaneously measured variables. *Phys. Rev. Lett.* **1994**, *73*, 951. [CrossRef]