Telling cause from effect in deterministic linear dynamical systems
Supplementary Material

We have prepared an appendix to address the proofs for Proposition 3, Theorems 1 and 2 which we provide in the following sections. For this purpose we will use a few extra notations which we define here. We will use $\tau_N(A)$ for the normalised trace of a square matrix A of order N and we drop N when there is no confusion. In addition, we derive analytic expressions of the VAR model in the last section.

1. Proof of Proposition 3

Lemma 3. *For $f \in L^2(\mathcal{I})$ non-constant, such that $1/f \in L^2(\mathcal{I})$, we have*

$$\int_{\mathcal{I}} f(x)^2 dx \cdot \int_{\mathcal{I}} \frac{1}{f(x)^2} dx > 1$$

Proof. Using Cauchy-Schwartz inequality for the scalar product

$$\left\langle f(x)\,,\, \frac{1}{f(x)} \right\rangle = \int_{\mathcal{I}} f(x) \cdot \frac{1}{f(x)} dx = 1\,.$$

Inequality is strict since f and $1/f$ are not collinear (otherwise f would be constant). □

Lemma 4. *Let $f \in L^1(\mathcal{I})$ be positive, non-constant, such that $1/f \in L^1(\mathcal{I})$ and $\int_{\mathcal{I}} f(x)dx = 1$.*

Assume $\exists \alpha > 0, \forall x \in \mathcal{I}, f(x) \leq 2 - \alpha$,

then

$$\int_{\mathcal{I}} f(x)dx \cdot \int_{\mathcal{I}} \frac{1}{f(x)} dx \geq 1 + \alpha \int_{\mathcal{I}} (f(x) - 1)^2 dx$$

Proof. We denote $s(x) = f(x) - 1$. Then $\int_{\mathcal{I}} s(x)dx = 0$ and

$$\int_{\mathcal{I}} f(x)dx. \int_{\mathcal{I}} \frac{1}{f(x)} dx - 1 = \int_{\mathcal{I}} \frac{-s(x)}{1 + s(x)} dx$$

For $x > -1$, we have

$$\frac{-x}{1+x} \geq x^2 - x^3 - x. \tag{19}$$

Replacing $s(x)$ with x in (19) we get:

$$\int_{\mathcal{I}} f(x)dx \cdot \int_{\mathcal{I}} \frac{1}{f(x)} dx - 1 \geq \int_{\mathcal{I}} s(x)^2 (1 - s(x))dx.$$

Since $1 - s(x) = 2 - f(x) \geq \alpha > 0$,

$$\int_{\mathcal{I}} f(x)dx \cdot \int_{\mathcal{I}} \frac{1}{f(x)} dx - 1 \geq \alpha \int_{\mathcal{I}} s(x)^2 dx$$

□

Proof of Proposition 3. Using the definition of Spectral Dependency Ratios and Lemma 3 we get

$$\rho_{\mathbf{X}\to\mathbf{Y}} \rho_{\mathbf{Y}\to\mathbf{X}} = \frac{1}{\langle|\widehat{\mathbf{h}}_{\mathbf{X}\to\mathbf{Y}}|^2\rangle \langle 1/|\widehat{\mathbf{h}}_{\mathbf{X}\to\mathbf{Y}}|^2\rangle} < 1$$

Moreover, applying Lemma 4 to $f = |\widehat{\mathbf{h}}_{\mathbf{X}\to\mathbf{Y}}|^2 / \int_{\mathcal{I}} |\widehat{\mathbf{h}}_{\mathbf{X}\to\mathbf{Y}}|^2 = |\widehat{\mathbf{h}}_{\mathbf{X}\to\mathbf{Y}}|^2 / \|\mathbf{h}_{\mathbf{X}\to\mathbf{Y}}\|_2^2$ we get inequality (9). □

2. Proof of Theorem 1

To prove this theorem we rely on a theorem from (Janzing et al., 2010) and a corollary that we derive from it.

Theorem 3 (concentration of measure for finite dimensional linear relationships)**.** *(Janzing et al., 2010) Suppose Σ is a given covariance matrix and suppose $A \in M_{m\times n}(\mathbb{R})$ is also a given matrix. Then if one generates $\Sigma_X = U\Sigma U^\top$ by uniformly choosing an orthogonal matrix U from $O(n)$ then Σ_X together with A, satisfies trace condition in probability when n tends to infinity. More precisely for a given ε there exist $\delta := 1 - \exp(-\kappa(n-1)\varepsilon^2)$, κ being a constant where*

$$|\tau_m(A\Sigma_X A^\top) - \tau_n(\Sigma_X)\tau_m(AA^\top)| =$$
$$|\tau_m(AU\Sigma U^\top A^\top) - \tau_n(\Sigma)\tau_m(AA^\top)| \leq 2\varepsilon \|\Sigma\| \|AA^\top\|$$

holds with probability δ.

In the above theorem (and the rest of the document), $\|.\|$ applied to a matrix will refer to the operator norm. The following corollary is a direct consequence of the previous theorem:

Corollary 1. *Suppose Σ is a given covariance matrix and suppose $A \in M_{m\times n}(\mathbb{R})$ is also a given matrix. Then if one generates $A_U = AU$ by uniformly choosing an orthogonal matrix U from $O(n)$ then A_U together with Σ, satisfies trace condition in probability when n tends to infinity. More precisely for a given ε there exist $\delta :=$*

$1-\exp(-\kappa(n-1)\varepsilon^2)$, *$\kappa$ being a constant where*

$$|\tau_m(A_U \Sigma A_U^\top) - \tau_n(\Sigma_X)\tau_m(AA^\top)| =$$
$$|\tau_m(AU\Sigma U^\top A^\top) - \tau_n(\Sigma)\tau_m(AA^\top)| \le 2\varepsilon\|\Sigma\|\|AA^\top\|$$

holds with probability δ.

To prove the main theorem we will also need two lemmas that are stated below.

Lemma 5. *(Serre, 2010) For a given Hermitian matrix H and any principal submatrix of H, H', their spectral radius ρ_s satisfies*

$$\rho_s(H) \ge \rho_s(H').$$

Lemma 6. *(Gray, 2006) Let $f : [-\frac{1}{2}, \frac{1}{2}) \to \mathbb{R}$ $f \in L^1$ be a bounded function and suppose t_k is its Fourier series coefficients, i.e.*

$$t_k = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\nu) e^{i2\pi k\nu} d\nu, \quad t \in \mathbb{Z}.$$

Consider Toeplitz matrices T_n defined as

$$[T_n]_{ij} = t_{i-j} \quad i, j \in \{0, ..., n-1\}$$

with eigenvalues $\tau_{n,k} (0 \le k \le n-1)$. Then if t_i are absolutely summable we get:

$$\min_{x\in[-\frac{1}{2},\frac{1}{2})} f(x) \le \tau_{n,i} \le \max_{x\in[-\frac{1}{2},\frac{1}{2})} f(x)$$

Proof of Theorem 1. Without loss of generality and for the sake of simplicity we only consider the positive indices of the time series and we take the filter to be causal; other cases can be treated in a similar way. Then the following relation holds between input and output of the filter:

$$\forall i, \quad 0 \le i \le N-1 \quad Y_i = \sum_{j=0}^{m-1} b_j X_{i-j}$$

Formulated in terms of matrices the above relation can be represented as

$$\begin{bmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_{N-2} \\ Y_{N-1} \end{bmatrix} = B \begin{bmatrix} X_{-m+1} \\ X_{-m+2} \\ \vdots \\ X_{N-2} \\ X_{N-1} \end{bmatrix},$$

where B is a $N \times (N+m-1)$ matrix as follows:

$$\begin{bmatrix} b_{m-1} & b_{m-2} & \cdots & b_0 & 0 & \cdots & 0 & 0 \\ 0 & b_{m-1} & \cdots & b_1 & b_0 & \cdots & 0 & 0 \\ & & \ddots & & & & & \\ 0 & 0 & \cdots & b_{m-1} & \cdots & b_1 & b_0 & 0 \\ 0 & 0 & \cdots & 0 & b_{m-1} & \cdots & b_1 & b_0 \end{bmatrix}$$

We define $\Sigma_X^i \in M_{m\times m}(\mathbb{R})$ to be the covariance matrices as follows:

$$\forall i \quad 0 \le i \le N-1 \quad 0 \le j, k \le m-1$$
$$[\Sigma_X^i]_{jk} = \mathbb{C}\mathrm{ov}(X_{i+j}, X_{i+k})$$

Since the time series under consideration are weakly stationary it is obvious that Σ_X^i is independent of i and we can replace any appearance of it with Σ. If we take $\Sigma_{X_{0:N-1}}, \Sigma_{Y_{0:N-1}} \in M_{N\times N}(\mathbb{R})$ to be the covariance matrices for $X_{0:N-1}$ and $Y_{0:N-1}$ respectively, then we have

$$\Sigma_{Y_{0:N-1}} = B\Sigma_{X_{-m+1:N-1}} B^\top$$

Also define $\Sigma_{Y_{0:N-1}}^U$ to be the covariance matrix of the output for FIR $\mathcal{S}'$ with $\mathbf{b}' = U^\top \mathbf{b}$. Furthermore assume the spectrum of the output for this filter is S_{yy}^U. One can write the diagonal elements of $\Sigma_{Y_{0:N-1}}^U$ based on the above equation as follows:

$$[\Sigma_{Y_{0:N-1}}^U]_{ii} = \mathbf{b}^\top U\Sigma U^\top \mathbf{b}$$

which is therefore equal to the normalized trace of $\Sigma_{Y_{0:N-1}}^U$. Taking $A = \mathbf{b}^\top$ in corollary 1 for a randomly selected U we get

$$|\mathbf{b}^\top U\Sigma U^\top \mathbf{b} - \frac{1}{N}\tau_m(\Sigma)\langle \mathbf{b}, \mathbf{b}\rangle| \le 2\varepsilon\|\Sigma\|\sqrt{\langle \mathbf{b}, \mathbf{b}\rangle}$$

and hence

$$|\tau(\Sigma_{Y_{0:N-1}}^U) - \frac{1}{N}\tau_m(\Sigma)\|\mathbf{b}\|_2^2| \le 2\varepsilon\|\Sigma\|\|\mathbf{b}\|_2^2 \qquad (20)$$

with probability δ. On the other hand the elements of diagonals of Σ are $C_X(0)$. Therefore:

$$\frac{1}{N}\tau_m(\Sigma) = \frac{mNC_X(0)}{mN} = P(\mathbf{X})$$

Since Σ is a principal submatrix of $\Sigma_{X_{0:N-1}}$ therefore by corollary 5

$$\|\Sigma\| = \rho(\Sigma) \le \rho(\Sigma_{X_{0:N-1}}).$$

Because $C_X(\tau)$'s are absolutely summable, based on lemma 6 we get

$$\rho(\Sigma_{X_{0:N-1}}) \le \max_\nu S_{xx}(\nu),$$

and then inequality (20) can be rewritten as

$$|\frac{\tau(\Sigma_{Y_{0:N-1}}^U)}{P(\mathbf{X})\|\mathbf{b}\|_2^2} - 1| \le 2\frac{\varepsilon}{P(X)}\|\Sigma\|$$

which completes the proof. □

3. Proof of Theorem 2

In this section we give a proof that the TDR (see (12)) asymptotically approaches the SDR (see (6)). We first state and prove two lemmas that are used to derive this result. As before suppose $\{X_t\}$ and $\{Y_t\}$ are given input and output of an LTI filter that are related through the impulse response function $\{h_t\}$. According to the definition of the truncated linear systems (see Definition 2) of order N for the linear system above we get the following matrix relationship:

$$\begin{bmatrix} Y'_{-N} \\ Y'_{-N+1} \\ \vdots \\ Y'_{N-2} \\ Y'_{N-1} \end{bmatrix} = \tag{21}$$

$$\begin{bmatrix} h_0 & h_{-1} & \cdots & h_{-2N+1} \\ h_1 & h_0 & \cdots & h_{-2N+2} \\ & \vdots & & \\ h_{2N-2} & h_{2N-3} & \cdots & h_{-1} \\ h_{2N-1} & h_{2N-2} & \cdots & h_0 \end{bmatrix} \begin{bmatrix} X_{-N} \\ X_{-N+1} \\ \vdots \\ X_{N-2} \\ X_{N-1} \end{bmatrix}.$$

According to definition 2 we can write the following TDR for the linear equation (21):

$$r_{\mathbf{X}_N \to \mathbf{Y}_N} = \frac{\tau(\Sigma_{\mathbf{Y}_N})}{\tau(\Sigma_{\mathbf{X}_N})\tau(H^N {H^N}^T)} \tag{22}$$

Define

$$T_N := \tau(H^N {H^N}^\top). \tag{23}$$

Now we show that T_N converges to $\|\mathbf{h}\|_2^2$ the energy of the impulse response.

Lemma 7. *Assume $\|\mathbf{h}\|_2^2 < +\infty$, then*

$$\lim_{N\to+\infty} T_N = \|\mathbf{h}\|_2^2$$

Proof. To show

$$\lim_{N\to\infty} \sum_{k=-2N+1}^{2N-1} |h_k|^2 \frac{2N-1}{2N} = \sum_{k\in\mathbb{Z}} |h_k|^2, \tag{24}$$

define

$$a_{N,k} := \begin{cases} |h_k|^2 \frac{2N-1}{2N} & \text{for } |k| \le 2N-1 \\ 0 & \text{otherwise} \end{cases}$$

Then (24) is equivalent to

$$\lim_{N\to\infty} \sum_{k\in\mathbb{Z}} a_{N,k} = \sum_{k\in\mathbb{Z}} \lim_{N\to\infty} a_{N,k}\,,$$

which follows from the monotone convergence theorem (Yeh, 2006) since $(a_{N,k})_N$ is a monotonically increasing sequence for each k.

□

In order to get the main result, we also need to prove that Y_k''s in (21) are asymptotically converging to Y_k's in the following sense:

Lemma 8. *Suppose an LTI filter $\mathcal{S}$ with zero mean weakly stationary processes as input ($\{X_t\}$) and output ($\{Y_t\}$) and impulse response function $\{h_t\}$ has been given. Then for the truncated linear systems we have:*

$$\lim_{N\to\infty} |\tau(\Sigma_{\mathbf{Y}_N}) - \tau(\Sigma_{\mathbf{Y}'_N})| = 0,$$

Proof. Let us define a symmetric positive semi-definite bilinear form on the vector space of truncated stochastic processes $\mathbf{W}_{-N:N-1}$ by

$$\langle \mathbf{W}_N, \mathbf{Z}_N \rangle := \tau(\Sigma_{\mathbf{W}'_N \mathbf{Z}'_N})\,,$$

where Σ_{WZ}^N denotes the cross-covariance matrix between the two truncated processes (note that our notation Σ_W is simply an abbreviation for Σ_{WW}). The inner product induces a semi-norm by

$$\|W\| := \sqrt{\langle W, W\rangle}\,.$$

Using the triangle inequality one easily obtains

$$|\|W\| - \|Z\|| \le \|W - Z\|\,.$$

Hence,

$$|\sqrt{\tau(\Sigma_{YY})} - \sqrt{\tau(\Sigma_{Y'Y'})}| \le \sqrt{\tau(\Sigma_{Y-Y'})}. \tag{25}$$

To prove the lemma it is enough to show that the r.h.s. of (25) approaches zero when N goes to infinity. So we show that each element of diagonal of $\Sigma_{Y-Y'}$ tends to zero when N tends to infinity. With overload of notation, in this case define $\{h_t^{(j)}\}$ as follows

$$h_t^{(j)} = \begin{cases} 0 & \text{if } -N \le t+j \le N-1 \\ h_t & \text{otherwise.} \end{cases}$$

Then for the j-th element of diagonal of $\Sigma_{Y-Y'}$ we have

$$[\Sigma_{Y-Y'}]_{jj} = \mathbb{E}\Big[(Y_j - Y_j')^2\Big] = \tag{26}$$

$$\mathbb{E}\Big[(\sum_{l=-\infty}^{\infty} X_{j-l} h_l^{(j)})^2\Big] = \mathbb{E}\Big[(\sum_{\substack{l\ge N-j \\ l<-N-j}} X_{j-l} h_l)^2\Big] \tag{27}$$

Since autocorrelation function attains its maximum at $t=0$ and

$$\forall i,j \in \mathbb{Z}, \quad \mathbb{E}(X_i X_j) \le \sqrt{\mathbb{E}(X_i^2)\mathbb{E}(X_j^2)}$$

we get:

$$\forall i,j \in \mathbb{Z}, \quad \mathbb{E}(X_i X_j) \le \mathbb{E}(X_0^2).$$

As a result we have:

$$[\Sigma_{Y-Y'}]_{jj} = \mathbb{E}\Big[(\sum_{\substack{l\geq N-j\\ l<-N-j}} X_{j-l}h_l)^2\Big] \leq$$
$$\sum_{\substack{l,l'\geq N-j\\ l,l'<-N-j}} \mathbb{E}(X_0^2)h_l h_{l'} = \mathbb{E}(X_0^2) \sum_{\substack{l,l'\geq N-j\\ l,l'<-N-j}} h_l h_{l'} \leq$$
$$\mathbb{E}(X_0^2)(\sum_{\substack{l\geq N-j\\ l<-N-j}} h_l)^2 \leq \mathbb{E}(X_0^2)(\sum_{\substack{l\geq N-j\\ l<-N-j}} |h_l|)^2.$$

Since $\{h_t\}$ is absolutely convergent, it follows that $[\Sigma_{Y-Y'}]_{jj}$ can be arbitrarily reduced by increasing N. Then it follows that $\tau(\Sigma_{Y-Y'})$ approaches to zero when N tends to infinity. This concludes the proof. □

To derive the result regarding the asymptotic behaviour of the trace condition in the truncated linear systems and the equivalence of trace condition (see postulate 2) to SIC, we will also need one of the convergence theorems due to Szegö:

Theorem 4 (Szegö's convergence theorem)**.** *(Gray, 2006) Let $f : [-\frac{1}{2}, \frac{1}{2}) \to \mathbb{R}$ $f \in L^1$ be a bounded function and suppose t_k's are its Fourier series coefficients, i.e.*

$$t_k = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\nu)e^{i2\pi k\nu}d\nu, \quad t \in \mathbb{Z}.$$

Consider Toeplitz matrices T_n defined as

$$[T_n]_{ij} = t_{i-j} \quad i,j \in \{0,...,n-1\}$$

with eigenvalues $\tau_{n,k}(0 \leq k \leq n-1)$. Then if T_n's are Hermitian, i.e. $t_i = \bar{t}_i$ for any i, then for any continuous function F we have:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} F(\tau_{n,k}) = \int_{-\frac{1}{2}}^{\frac{1}{2}} F(f(\nu))d\nu$$

We are ready to state our convergence theorem:

Theorem 5. *For a given truncated linear time series, $r_{\mathbf{X}'_N\to\mathbf{Y}'_N}$ asymptotically approaches to the spectral values of time series on infinite domain. As a result the spectral density based estimator coincides with the trace based estimator in the limit, and more precisely*

$$\lim_{N\to\infty} \tau(\Sigma_{\mathbf{X}'_N}) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{xx}(\nu)d\nu,$$
$$\lim_{N\to\infty} \tau(\Sigma_{\mathbf{Y}'_N}) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{yy}(\nu)d\nu,$$
$$\text{and} \qquad \lim_{N\to\infty} T_N = \int_{-\frac{1}{2}}^{\frac{1}{2}} |\widehat{\mathbf{h}}(\nu)|^2 d\nu,$$

where T_N is defined as in (23). And eventually:

$$\lim_{n\to\infty} r_{\mathbf{X}'_N\to\mathbf{Y}'_N} = \rho_{\mathbf{X}\to\mathbf{Y}} \qquad \lim_{n\to\infty} r_{\mathbf{Y}'_N\to\mathbf{X}'_N} = \rho_{\mathbf{Y}\to\mathbf{X}}$$

Proof. Both $\Sigma_{\mathbf{X}'_N}$ and $\Sigma_{\mathbf{Y}'_N}$ are hermitian Toeplitz matrices and based on theorem 4 where F has been chosen as identity function and also applying lemma 8 we get:

$$\lim_{N\to\infty} \tau(\Sigma_{\mathbf{X}_N}) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{xx}(\nu)d\nu \tag{28}$$
$$\lim_{N\to\infty} \tau(\Sigma_{\mathbf{Y}_N}) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{yy}(\nu)d\nu \tag{29}$$

Moreover by Plancherel's theorem and lemma 7 it follows that:

$$\lim_{N\to\infty} T_N = \|\mathbf{h}\|_2^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} |\widehat{\mathbf{h}}(\nu)|^2 d\nu \tag{30}$$

□

This theorem therefore shows that the trace ratios calculated for windowed version of time series are nothing but estimates of the spectral ratios and therefore justifies that these two different methods for causal inference are indeed consistent with each other.

4. Analytic expressions of the VAR model

Using basic properties of the Z-transform, we can derive from (13) the following analytic expressions of the input PSD S_{xx}:

$$S_{xx}(\nu) = |\widehat{n}(\nu)|^2 = |\widetilde{n}(\exp(2\pi\mathbf{i}\nu))|^2,$$

with

$$\widetilde{n}(z) = \frac{1}{1-\sum_k a_k z^{-k}}.$$

Moreover, the transfer function corresponding to the mechanism in (15) is

$$\widetilde{m}(z) = \frac{\sum_k c_k z^{-k}}{1 - \sum_k b_k z^{-k}}.$$

As a consequence, testing SIC on the VAR model in the forward direction amounts (when neglecting the filtered noise ξ) to test independence between

$$|\widehat{h}(\nu)|^2 = |\widetilde{m}(\exp(2\pi \mathbf{i}\nu))|^2 \tag{31}$$

and

$$S_{xx}(\nu) = |\tilde{n}(\exp(2\pi \mathbf{i}\nu)|^2 \,, \tag{32}$$

which are parametrized by the coefficients $\{b_k, c_k\}$ and $\{a_k\}$ respectively.