Granger causality has long been a prominent method for inferring causal interactions between stochastic variables for a broad range of complex physical systems. However, it has been recognized that a moving average (MA) component in the data presents a serious confound to Granger causal analysis, as routinely performed via autoregressive (AR) modeling. We solve this problem by demonstrating that Granger causality may be calculated simply and efficiently from the parameters of a state-space (SS) model. Since SS models are equivalent to autoregressive moving average models, Granger causality estimated in this fashion is not degraded by the presence of a MA component. This is of particular significance when the data has been filtered, downsampled, observed with noise, or is a subprocess of a higher dimensional process, since all of these operations—commonplace in application domains as diverse as climate science, econometrics, and the neurosciences—induce a MA component. We show how Granger causality, conditional and unconditional, in both time and frequency domains, may be calculated directly from SS model parameters via solution of a discrete algebraic Riccati equation. Numerical simulations demonstrate that Granger causality estimators thus derived have greater statistical power and smaller bias than AR estimators. We also discuss how the SS approach facilitates relaxation of the assumptions of linearity, stationarity, and homoscedasticity underlying current AR methods, thus opening up potentially significant new areas of research in Granger causal analysis.

DOI: 10.1103/PhysRevE.91.040101 PACS number(s): 02.50.Sk, 02.50.Ti, 05.45.Tp, 87.10.Mn

Wiener-Granger causality (GC), a powerful method for determining information transfer or directed functional connectivity between stochastic variables, is based on the premise that cause (a) precedes effect, and (b) contains unique information about effect [1–3]. Since its inception, it has steadily gained popularity in a broadening range of fields (see, e.g., [4]), due to its data-driven nature (few structural assumptions need be made about the data generation process), conceptual simplicity, spectral decomposition property, and ease of implementation. It is most commonly operationalized in a linear autoregressive (AR) modeling context, to the extent that it is frequently viewed as an essentially autoregressive method. This is unfortunate, first because it obscures the nonparametric essence of the original idea (indeed, GC may be formalized in purely information-theoretic terms [5–7]); more importantly, data from major application domains are frequently unsuited to AR modeling.

Time series data from diverse application domains often contain a strong moving average (MA) component, which may not be represented parsimoniously by a finite order AR model. Apart from any MA component intrinsic to the underlying signal, an MA component may be induced in an observed process by common data acquisition, sampling and preprocessing procedures: if an AR process is subsampled, aggregated, or observed with additive measurement noise, the resultant process will be an autoregressive moving average (ARMA) process [8]. A subprocess of an AR process will also generally be ARMA [9], as will a filtered AR process [10]. In general an ARMA process will have infinite AR model order—but of course in finite sample a finite, possibly large, model order must be selected, which will be reflected in GC statistics of reduced statistical power and increased bias.

In contrast to linear AR models, the class of multivariate ARMA models—or, equivalently [11], finite order linear state-space (SS) models—is closed under all of the above-mentioned operations. While the potential of SS modeling for GC inference has been remarked on [8,12–14], a rigorous derivation and demonstration of this potential has not yet been provided. Here we show how GC, conditional and unconditional, in both time and frequency domains, may be easily derived from SS parameters via solution of a discrete algebraic Riccati equation (DARE). We verify in simulation, for a bivariate ARMA process, the increase in statistical power and reduction in bias of GC estimated via SS, as compared to AR, modeling. We also discuss potential extensions of the SS approach beyond the usual stationary linear scenario, to encompass cointegrated processes, models with time-varying parameters, and nonlinear models, thus opening up new avenues of research and application. State-space Granger causal inference stands to significantly enhance our ability to identify and understand causal interactions and information flow in a wide range of complex dynamical systems.

I. STATE-SPACE MODELS

Our starting point [15] is a discrete-time, real-valued vector stochastic process \( y_t = [y_{1t} y_{2t} \cdots y_{mt}]^T \), \(-\infty < t < \infty\), of observations. The general time-invariant linear SS model without input for the observation process \( y_t \) is

\[
\begin{align*}
x_{t+1} &= A x_t + u_t & \text{state transition equation,} \\
y_t &= C x_t + v_t & \text{observation equation,}
\end{align*}
\]

where \( x_t \) is an (unobserved) \( m \)-dimensional state variable, \( u_t, v_t \) are zero-mean white noise processes, \( C \) is the observation matrix, and \( A \) is the state transition matrix. The parameters of

---

1. l.c.barnett@sussex.ac.uk
2. a.k.seth@sussex.ac.uk
the model (1) are \((A, C, Q, R, S)\), where
\[
\begin{bmatrix}
Q & S \\
S^T & R
\end{bmatrix} = \mathbb{E}
\begin{bmatrix}
u_i u_i' \\
u_i v_i'
\end{bmatrix}
\] (2)
is the noise covariance matrix. We assume that \(x_t, y_t\) are weakly stationary, which requires that the transition equation (1a) be stable, i.e., \(\lambda_{\text{max}}(A) < 1\), where \(\lambda_{\text{max}}(A)\) denotes the maximum of the absolute values of the eigenvalues of \(A\). We also assume that \(R\) is positive-definite. A process \(y_t\) satisfying a stable SS model (1) also satisfies a stable ARMA model; conversely, any stable ARMA process may be shown to satisfy a stable SS model of the form (1) [11].

Given an SS model (1), we define \(z_t \equiv \mathbb{E}[x_t | y_{t-1}]\), the projection of the state variable \(x_t\) on the space spanned by the infinite past \(y_{t-1} = [y_{t-1}', y_{t-2}' \ldots]'\) of the observation variable [11]. It is then not difficult to show [16] that the innovations \(\varepsilon_t \equiv y_t - \mathbb{E}[y_t | y_{t-1}]\) constitute a white noise process with positive-definite covariance matrix \(\Sigma \equiv \mathbb{E}[\varepsilon_t \varepsilon_t']\), and that in terms of the new state variable \(z_t\) we have an SS model
\[
z_{t+1} = Az_t + Ke_t, \quad \text{state equation}, \tag{3a}
y_t = Cz_t + \varepsilon_t, \quad \text{observation equation}, \tag{3b}
\]
for \(y_t\), where \(K\) is the Kalman gain matrix. The SS model (3) is said to be in innovations form, with parameters \((A, C, K, \Sigma)\). Note that innovations form constitutes a special case of (1), with \(u_t = Ke_t, v_t = \varepsilon_t\), so that \(Q = K\Sigma K^T, S = K\Sigma, R = \Sigma\).

We can write the state equation (3a) as \(z_t = (I - A\zeta)^{-1}Kz_t\cdot e_t\), where here \(z\) represents the back-shift operator [17], which yields the MA representation for the observation process
\[
y_t = H(z) \cdot e_t, \quad H(z) \equiv I + C(I - A\zeta)^{-1}Kz, \tag{4}
\]
with transfer function \(H(z)\). From (3) we have \(z_{t+1} = Bz_{t} + K y_{t}\), with \(B = A - K C\), from which we may derive the AR representation
\[
B(z) \cdot y_t = \varepsilon_t, \quad B(z) \equiv I - C(I - B\zeta)^{-1}Kz, \tag{5}
\]
where \(B(z) = H(z)^{-1}\). The model is minimum phase if \(\lambda_{\text{max}}(B) < 1\) (a stable inverse ARMA model for the process \(y_t\), then exists). We assume minimum phase from now on. From (4), the cross-power spectral density (CPSD) of \(y_t\) has the factorization [18]
\[
S(z) = H(z)\Sigma H^*(z) \tag{6}
\]
on \(|z| = 1\) in the complex plane.

Assuming stability, minimum phase, and positive-definite observation noise covariance [19], the DARE [20]
\[
P - APA^T = Q - (APC^T + S)(CPC^T + R)^{-1} \times (CPC^T + S^T) \tag{7}
\]
has a unique stabilizing solution for \(P\), and we have [21]
\[
\Sigma = CPC^T + R \quad \text{and} \quad K = (APC^T + S)\Sigma^{-1}. \tag{8a}
\]
\[
K = (APC^T + S)\Sigma^{-1}. \tag{8b}
\]
Given general SS parameters \((A, C, Q, R, S)\) then, corresponding innovations form parameters \(K, \Sigma\) may be obtained through (8) via solution of (7).

II. GRANGER CAUSALITY

Granger causality is commonly expressed in terms of prediction error. Here we follow the standard formulation of Geweke [22,23]. Suppose that an observable process \(y_t\) is partitioned into subprocesses: \(y_t = [y_{t1}', y_{t2}', y_{t3}']',\) within the “observable universe of information” represented by the process \(y_t, \text{GC from } y_{t2} \text{ to } y_{t1} (\text{conditional on } y_{t3})\) quantifies the extent to which the past of \(y_{t2}\) improves prediction of the future of \(y_{t1}\) over and above the extent to which \(y_{t1}\) (along with \(y_{t3}\)) already predicts its own future. Now the best prediction (in the least-squares sense) of \(y_{t1}\), given the entire universe of past information \(y_{t-1}\), is the projection \(\mathbb{E}[y_{t1} | y_{t-1}]\). This prediction may be contrasted with the prediction \(\mathbb{E}[y_{t1} | y_{t2}']\) of \(y_{t1}\) based on the reduced universe of past information \(y_{t-1}'\), where \(y_{t2}' = [y_{t1}', y_{t2}']',\) omits \(y_{t2}\) from the observable information set. In Geweke’s formulation, predictive power is quantified by the generalized variances [24,25] of the associated full and reduced residual errors (innovations), \(\varepsilon_t \equiv y_t - \mathbb{E}[y_t | y_{t-1}]\) and \(\varepsilon_t^0 \equiv y_t - \mathbb{E}[y_t | y_{t-1}']\), respectively. Specifically, Geweke [23] defines the time-domain GC from \(y_2\) to \(y_1\) conditional on \(y_3\) (for unconditional GC we may take \(y_3\) to be empty) as
\[
F_{y_2 \rightarrow y_1 | y_3} = \ln \frac{|\Sigma_{11}^n|}{|\Sigma_{11}|}, \tag{9}
\]
where \(\Sigma = \mathbb{E}[\varepsilon_t \varepsilon_t']\) and \(\Sigma_{ij} = \mathbb{E}[\varepsilon_t^i \varepsilon_t^j]'.\) In a maximum likelihood (ML) framework, the corresponding statistic is just the log-likelihood ratio [26,27] for the nested AR models associated with the projections \(\mathbb{E}[y_t | y_{t-1}]\) and \(\mathbb{E}[y_t | y_{t-1}']\).

In the unconditional case (i.e., \(y_3\) is empty) \(f_{y_2 \rightarrow y_1}\) is given by [22]
\[
f_{y_2 \rightarrow y_1} = \ln \frac{|\Sigma_{11}|}{|\Sigma_{11}^n|} = \ln \left| \frac{\Sigma_{11}(\zeta)}{H(\zeta)} \right|^2 = \ln \frac{H(\zeta)}{H^*(\zeta)}, \tag{10}
\]
where \(\Sigma_{ij} = \Sigma_{ii} - \Sigma_{ik} \Sigma_{kk}^{-1} \Sigma_{kl}\) denotes a partial covariance matrix. For the conditional case, we define the process \(\tilde{y}_t \equiv [y_{t1}', y_{t2}', y_{t3}']',\) the frequency-domain causality from \(y_2\) to \(y_1\) conditional on \(y_3\) is then defined [23] as the unconditional causality
\[
f_{y_2 \rightarrow y_1 | y_3} = f_{y_{t2}' \rightarrow y_{t1} | y_{t3}}. \tag{11}
\]
Let \(\tilde{S}(z), \tilde{H}(z), \) and \(\tilde{\Sigma}\) be, respectively, the CPSD, transfer function, and innovations covariance matrix of the process \(\tilde{y}_t\). We note first that, since the innovations process \(\varepsilon_t^0\) is white, \(\tilde{S}_{11}(z)\) is just the flat spectrum \(\Sigma_{11}\). The AR representation for the reduced process \(y_{t2}'\) is \(B(y_t) \cdot y_{t2}' = \varepsilon_t^0\) where \(B(\zeta) = H(\zeta)^{-1}\) is the inverse transfer function of the reduced model.
so we have \( \tilde{y}_t = \tilde{B}(z) \cdot y_t \), where
\[
\tilde{B}(z) = \begin{bmatrix}
B_{11}(z) & 0 & 0 \\
0 & 1 & 0 \\
B_{31}(z) & 0 & B_{33}(z)
\end{bmatrix}.
\]

But \( y_t = H(z) \cdot \epsilon_t \), so that \( \tilde{y}_t = B(z)H(z) \cdot \epsilon_t \) and it follows that \( \tilde{H}(z) = \hat{B}(z)H(z) \) and \( \Sigma = \Sigma \). From (11) and (12), after some matrix algebra, we may derive
\[
f_{y_1 \rightarrow y_1}(z) = \ln \left| \Sigma^{-1}_1 \right| = \ln \left| \Sigma^{-1}_1 - [\hat{H}_1(z) \hat{H}_3(z)] \right| = \ln \left| \Sigma^{-1}_1 - [\tilde{H}_1(z) \tilde{H}_3(z)] \right|.
\]

where
\[
[\hat{H}_1(z) \hat{H}_3(z)] = \begin{bmatrix}
B_{11}(z) & B_{13}(z) \\
B_{31}(z) & B_{33}(z)
\end{bmatrix} \begin{bmatrix}
H_{12}(z) & H_{13}(z) \\
H_{22}(z) & H_{33}(z)
\end{bmatrix}.
\]

Suppose now that we have an SS model in innovations form (3) for the observation process \( y_t \), satisfying the assumptions of stability, minimum phase, and positive-definite innovations covariance. The reduced process \( y^0_t \) satisfies the observation equation:
\[
y^0_t = C^0 z_t + v^0_t,
\]
where \( C^0 = [C_1^0 \ C_3^0]^T \) and \( v^0_t = [\epsilon_1^0 \ \epsilon_3^0]^T \). Together with the original state transition equation (3a), (16) constitutes an SS model for \( y^0_t \). Note that in general this “reduced SS model” will not be in innovations form (it is for this reason that we write \( v^0_t \) rather than \( \epsilon^0_t \)). It is, however, stable (since \( A \) is stable) and the noise covariance \( R^0 = E[v^0_t v^0_t]^T \) is positive-definite (since \( \Sigma \) is positive-definite). By the minimum phase assumption, \( y_t \) is invertible ARMA, so that the subprocess \( y^0_t \) is also invertible ARMA [9] and the reduced SS model is thus minimum phase. Now the reduced Kalman gain \( K^0 \) [which enters into the expression for \( B^0(z) \)], and the reduced innovations covariance \( \Sigma^0 \), may be obtained by solving the DARE (7) for the general form SS (3a) and (16). Thus, given innovations form SS parameters \( A,C,K,\Sigma \), GCs in both time and frequency domain, conditional and unconditional, are readily calculated.

\( F_{y_1 \rightarrow y_1} \) vanishes precisely when coefficients with block indices 12 vanish at all lags in the AR representation of the process \( y_t \). For an SS model, setting \( H_{12}(z) \equiv 0 \) in (5) yields the necessary and sufficient condition
\[
C_1^0 B^0 K^2 = 0, \quad k = 0,1,2,\ldots.
\]

By the Cayley–Hamilton theorem, (17) need be satisfied just for \( k = 0,1,\ldots,m - 1 \). In particular, \( (B,C) \) observable \( \Rightarrow K_2 = 0 \) while \( (B,K) \) controllable \( \Rightarrow C_1 = 0 \). In terms of the MA representation (4), performing block inversion of \( B(z) = H(z)^{-1} \), the noncausality condition is found to be \( H_{12}(z) \equiv 0 \) in the conditional case, or just \( H_{12}(z) \equiv 0 \) in the unconditional case [in which case (17) holds with \( B \) replaced by \( A \)].

AR model parameters are generally estimated from data by standard regression techniques such as ordinary least-squares (OLS) or so-called Levinson-Wiggins-Robinson (LWR) algorithms [30–33], which yield pseudo-ML estimates [34].

Though we are not primarily concerned with identification of SS models, for which a large literature exists [21,35], many identification procedures yield (asymptotically) pseudo-ML estimates for model parameters. State-space–subspace (SS-SS) algorithms, in particular, are noniterative and highly efficient [36]. See Supplemental Material [37] for further discussion on GC estimation and statistical inference for AR and SS models, establishing that SS-SS GC estimation is in general more computationally efficient and numerically stable than AR GC estimation.

III. SIMULATION EXPERIMENT

To examine the performance of state-space Granger causal inference, we generated and analyzed simulated time series data using a bivariate AR(1) process for which GC values may be computed analytically. The AR data were then filtered through an invertible weighted MA filter of order \( r \) which leaves causalities invariant [10]. By construction, the resultant ARMA(\( r,1 \)) processes \( y_t = [y_{1t}^0 y_{2t}^0]^T \) satisfy \( F_{y_1 \rightarrow y_2} = F \geq 0 \), while \( F_{y_2 \rightarrow y_1} = 0 \). We consider a causal model with (fixed) \( F > 0 \), and a null model with \( F = 0 \). MA order is allowed to vary from \( r = 0 \) (pure AR) to \( r = 9 \) (strong MA component). See Supplemental Material [37] for details of the model, GC estimation, and statistical procedures.

Based on 10,000 sample simulations of \( y_t \) for 1000 time steps, model orders were estimated and empirical distributions obtained for AR and SS estimators \( \hat{F}_{y_1 \rightarrow y_2} \), for both null and causal models. Firstly, Fig. 1 confirms that estimated AR model order increases more rapidly with increasing MA order than SS model order. Bias, and statistical power at \( \alpha = 0.05 \), were estimated from the empirical GC distributions. The key outcome (Fig. 2) is that both bias and power scale more slowly with increasing MA order for the SS than the AR causal estimators.

IV. DISCUSSION

We have described and experimentally validated an approach to Granger causal inference, based on state-space modeling. When the data manifest a moving average component—a ubiquitous feature of real-world data—this approach offers

![FIG. 1. (Color online) AR and SS mean model orders plotted against MA order, for the bivariate ARMA process \( y_t \).](image-url)
improved statistical power and reduced bias as compared to standard autoregressive methods.

Our primary theoretical contribution is to demonstrate that GC for a state-space process, conditional and unconditional, in time and frequency domains, may be simply expressed in terms of SS model parameters, the calculation involving solution of a single DARE. The resulting estimation of GC from empirical time series data is simpler and more computationally efficient and numerically stable than AR-derived causal estimation. Detailed simulation experiments verified the gain in statistical power and reduction in bias resulting from SS as compared to AR modeling of an ARMA process.

These results are important, because diverse application domains furnish time series data likely to feature a moving average component, either intrinsic to the data generating process, or arising from common data acquisition or manipulation operations. Indeed, the potential of SS modeling in this respect has, as noted previously, been recognized. Two recent publications, [38] and [39], deploy SS methods specifically to address the issue of measurement noise. The former utilizes a SS model to construct a denoised signal, from which Granger causalities are subsequently obtained. The latter reframes a putative AR model as a SS model which is, however, reconsidered as an AR model for Granger causality estimation. In both cases causalities could have been estimated directly from the SS model parameters via the method presented here. In general, SS-based Granger causal inference can be expected to outperform standard AR-based inference for ARMA data. Given the ready availability of effective and efficient state-space system identification procedures, we conclude that in the light of our results, state-space modeling should become the default approach to Granger causal analysis.

One important example, where the use of GC has remained controversial, is in functional magnetic resonance imaging (fMRI) of the brain, where the time series are highly downsampled and filtered (by hemodynamic processes) reflections of the neural dynamics of interest [4,12,40]. Other important scenarios where MA components are to be expected include climate science and economics.

The state-space approach also promises an exciting potential for extension beyond linear, stationary, and homoscedastic assumptions. For example, subspace algorithms may be adapted for cointegrated processes [41] and SS models have been developed for heteroscedastic noise variance [42]. A powerful feature of the Kalman filter underlying state prediction is that it applies to nonstationary systems, which raises the possibility of a systematic approach to GC analysis for systems where parameters vary over time (see, e.g., [43], where a SS model and Kalman filter is used to estimate coefficients for a time-varying AR model). Finally, nonlinear state-space systems already constitutes a mature field with a large literature. Bilinear SS models, in particular, form the basis for direct causal modeling (DCM) [44], an approach to causal analysis of coupled dynamical systems sometimes seen as a rival to GC analysis. The state-space setting provides a common framework in which to compare GC and DCM, helping to systematize statistical approaches to characterization of causal connectivity [45].

ACKNOWLEDGMENTS

The authors are grateful to the Dr. Mortimer and Theresa Sackler Foundation, which supports the Sackler Centre for Consciousness Science.


[15] Notational conventions: vector quantities are written in lowercase bold, matrices in uppercase. Superscript “T” denotes the transpose, “∗” conjugate transpose, and |·| the determinant of a (complex) matrix; superscript “R” refers to a “reduced” model.


[17] In the frequency domain $z = e^{-iω}$ where $-π < ω ≤ π$ is the phase angle (we note that the inverse $z^{-1}$ is sometimes used for the back-shift operator, particularly in the signal processing literature).


[19] In fact, these conditions may be relaxed somewhat (see, e.g., [46, 47]).


