

Granger Causality of Gaussian Signals from Noisy or Filtered Measurements ^{*}

Salman Ahmadi ^{*} Girish N. Nair ^{*} Erik Weyer ^{*}

^{*} *Department of Electrical and Electronic Engineering, University of
Melbourne, VIC 3010 Australia (e-mail:
ahmadis@student.unimelb.edu.au; {gnair, ewey}@unimelb.edu.au).*

Abstract: This paper investigates the assessment of Granger causality (GC) between jointly Gaussian signals based on noisy or filtered measurements. To do so, a recent rank condition for inferring GC between jointly Gaussian stochastic processes is exploited. Sufficient conditions are derived under which GC can be reliably inferred from the second order moments of the noisy or filtered measurements. This approach does not require a model of the underlying Gaussian system to be identified. The noise signals are not required to be Gaussian or independent, and the filters may be noncausal or nonminimum-phase, as long as they are stable.

Keywords: Granger causality, jointly Gaussian stochastic process, noise, filtering

1. INTRODUCTION

The analysis and control of large-scale dynamic systems become challenging as their internal and external interconnections grow in number. Lack of the knowledge about the correct qualitative relationships between subsystems can reduce the accuracy of identified models, and degrade the performance of controllers. Unraveling the relationships between time series is also of interest in other disciplines, such as econometrics and neuroscience. For instance, in econometrics, the relationship between domestic product and unemployment using their observed time series is investigated. In neuroscience, an important question is determining whether neural activity in one region of the brain affects, or is affected by, another region.

A powerful concept to address such problems is *Granger causality*. This notion was introduced by Clive Granger in economics in 1963 (Granger, 1963, 1980, 1988) inspired by Norbert Wiener's work on prediction in 1956 (Wiener, 1956). A signal x is said to cause another signal z if the optimal expected prediction error for a future value of z is reduced by knowing the past x and z , in comparison with if only the past values of z are known. Subsequently, Granger proposed a weaker definition, in terms of conditional probabilities in (Granger, 1980, 1988) as follows. Signal x is said to Granger cause z if the future of signal z and past x are conditionally dependent given past z at some time. Under a mean-square error prediction error, the first definition coincides with the second one for jointly Gaussian processes.

Granger causality is a statistical approach for inferring qualitative causal relationships from "in vivo", non-invasive observations and is useful when direct experimentation on a system is too expensive or risky. Contrary to the field of system identification, where it is known *a priori* that certain signals are inputs and others outputs,

^{*} This work was partially supported by the Australian Research Council via Future Fellowship grant FT140100527.

in causal inference the determination of such an input-output relationship is one of the main questions. Furthermore, feedback can be present among all signals, so that a categorisation into input and output may not be possible.

In (Caines and Chan, 1975), the idea of *feedback-freeness* for wide-sense stationary stochastic processes is introduced, which is shown to be equivalent to Granger non-causality under linear minimum mean-square error prediction. The relation between Granger non-causality and a linear time-invariant state space representation with a star graph structure as the topology of the network is considered in (Józsa et al., 2019), where it is shown that Granger non-causality is equivalent to the existence of such a representation.

The effect of additive noise on Granger causality has been discussed in (Solo, 2007; Nalatore et al., 2007, 2014; Smirnov, 2013; Newbold, 1978; Anderson et al., 2019). In (Nalatore et al., 2007, 2014), it is shown that adding uncorrelated white noise to the original signals can cause spurious causality between the noisy signals. Moreover, true causality between the original signals can be hidden due to noise. A denoising method via Kalman filter and expectation-maximization algorithm has been also proposed to mitigate the effects of noise. In (Anderson et al., 2019), the effect of additive noise on causality is investigated, assuming that the signals and the additive colored noises are mutually independent. The information-theoretic tool of transfer entropy is used in (Smirnov, 2013) to investigate Granger causality in the lack of accurate and exhaustive information such as latent variables and measurement noise.

The temporal order of signals and their time-reversed counterparts are considered in (Winkler et al., 2016) to infer causality based on Geweke's log likelihood ratio of variances of residuals (Geweke, 1982), and it is shown that time reversal testing can be robust to noise measurement. A nonlinear Cadzow method is proposed in (Rangarajan

and Rao, 2019) to estimate the parameters of the linear process in the presence of noise and then Geweke's measure of causality is exploited to infer the causality of processes corrupted by additive noise.

The impact of filtering on Granger causality has been investigated in (Sargent, 1987; Florin et al., 2010; Barnett and Seth, 2011; Seth et al., 2013; Solo, 2016; Anderson et al., 2019). In (Anderson et al., 2019), Granger causality of stochastic processes filtered by causal, minimum-phase linear transfer functions is investigated, and conditions introduced so that spurious causality is not inferred. In (Seth et al., 2013; Solo, 2016), it has been shown that when signals are filtered separately by minimum-phase causal linear filters, Granger causality is not affected. Furthermore, in (Solo, 2016), it has been shown that if the signals are separately filtered by causal nonminimum phase filters which can be represented as multiplication of an all-pass filter and a stable, minimum phase transfer function, then Granger causality is not affected as well. Otherwise, filtering can change the Granger causality.

To the best of our knowledge, in most of the literature on Granger causality with noisy signals, the noise terms are mutually independent from each other and from the underlying signals of interest. Furthermore, most analyses on Granger causality after filtering assume causal and minimum phase filters.

Here we investigate the effects of noise on Granger causality for the case where the additive noises can be dependent on each other and on the signals between which we wish to investigate causality (Theorem 4 and 5). Moreover, we propose sufficient conditions under which Granger causality can be inferred between signals filtered by non-causal and nonminimum phase filters (Theorem 6). The approach can be extended to nonlinear filters as well. Unlike previous works, we are not interested in investigating Granger causality or noncausality between the noisy/filtered processes. We wish to infer Granger causality between the original processes using noisy/filtered measurements. In order to derive such sufficient conditions, a rank-based method to infer Granger causality between jointly Gaussian stochastic processes introduced recently by the authors in (Ahmadi et al., 2019) is used. In fact, the sufficient conditions introduced in this paper guarantee that ranks of matrices constructed through noisy/filtered data are equal to the rank of a matrix representing Granger causality of underlying jointly Gaussian stochastic processes before noise corruption and/or filtering. The sufficient conditions involve bounds on the perturbations introduced by the impacts of noise/filters and the second-order statistics of noisy/filtered data.

The rest of the paper is organized as follows: in section 2, Granger causality between jointly Gaussian signals is formulated and recent results on assessing causality in terms of the rank of a matrix of covariances are mentioned. In section 3, the effect of noise is investigated and sufficient conditions are presented under which the Granger causality can be inferred correctly from the statistics of the noisy signals. Section 4 introduces sufficient conditions for causality investigation using the filtered signals. Section 5 concludes the paper.

Notation: Throughout this paper, the stochastic process segments $(x_k)_{k=\ell}^n$ and $(x_k)_{k=1}^n$ are denoted by x_ℓ^n and x^n , respectively. For $\ell \leq 1$, x_ℓ^n is written as $(x_k)_{k=1}^n$ and x_ℓ^n equals empty sequence when $\ell > n$ or $n < 1$. When clear from context, the full sequence $(x_k)_{k=1}^\infty$ is written as x . $\Gamma_{a,b}$ denotes the covariance between vectors a and b , the components of the covariance matrix Γ is denoted by γ and γ_{ab}^M indicates the maximum of the absolute value of the components of covariance matrix $\Gamma_{a,b}$.

2. PROBLEM FORMULATION OF GRANGER CAUSALITY FOR JOINTLY GAUSSIAN SIGNALS

In this section, the definition of Granger causality based on conditional probabilities is considered. Assuming the joint process is *partially finite-order Markov*, the definition reduces to the comparison of two conditional distributions as shown in the following. Then a rank-based necessary and sufficient condition is exploited to infer Granger causality between jointly Gaussian stochastic processes.

Definition 1. (Granger Causality (Granger, 1980)). Let $x_k, z_k, k = 1, 2, \dots$ be discrete-time stochastic processes. The stochastic process x is said to not *Granger cause (GC) z* if for all $k = 1, 2, \dots$

$$P(z_{k+1}|x^k, z^k) = P(z_{k+1}|z^k), \quad \text{with probability (w.p.) } 1, \quad (1)$$

where $P(\cdot|\cdot)$ denotes conditional probability measure. Otherwise, if there is a nonzero probability that

$$P(z_{k+1}|x^k, z^k) \neq P(z_{k+1}|z^k), \quad \text{for some } k \geq 1, \quad (2)$$

then x is said to Granger cause (GC) z .

If x does not GC z , it means that given the past and the present of z , the future value of z is always conditionally independent of the past and the present of x . On the other hand, x GC z means that given the past and the present of z at some time k , the future value of z at time $k + 1$ is influenced by the past and the present of x with a nonzero probability.

We make the following assumption:

Assumption 1. (Partial Markov- m). The stochastic process z is said to be partially Markov of order $m \geq 1$ in x and z if $P(z_{k+1}|x^k, z^k) = P(z_{k+1}|x_{k-m+1}^k, z_{k-m+1}^k)$, w.p.1.

Note that partial Markovianity is weaker than joint Markovianity, which is usually assumed in the literature e.g. (Geweke, 1982; Quinn et al., 2011, 2015; Kontoyiannis and Skoularidou, 2016).

Under the Assumption 1, (1) can be written as follows:

$$P(z_{k+1}|x_{k-m+1}^k, z_{k-m+1}^k) = P(z_{k+1}|z^k), \quad \text{w.p. } 1. \quad (3)$$

And hence, (3) can be used to infer Granger noncausality. Note that (3) is not a conditional independence due to the nonnestedness of the condition part of RHS in the condition part of LHS. For simplicity, we assume that both of the stochastic processes x and z are scalar-valued.

Theorem 1. (Ahmadi et al., 2019) Let x, z be jointly stationary Gaussian signals satisfying Assumption 1. Further assume that there is no deterministic relationship between x_{k-m+1}^k and z^k .

- x does not Granger cause z if and only if

$$\text{rank}(C_G^{x \rightarrow z}(m, k)) = \min\{m, k\}, \quad \forall k \geq 1, \quad (4)$$

where the *causality matrix* $C_G^{x \rightarrow z}(m, \ell), \ell \geq 1$ is defined as:

$$C_G^{x \rightarrow z}(m, \ell) := [\Gamma_{z^*, \tilde{x}} \quad \Gamma_{z^*, \tilde{z}} \quad \Gamma_{z^*, z^o}], \quad (5)$$

$z^* := z_{k-m+1}^{k+1}$, $\tilde{z} := z_{k-m+1}^k$, $z^o := z_{k-\ell+1}^{k-m}$, and $\tilde{x} := x_{k-m+1}^k$.

- if (4) holds, then z is marginally Markov- m , i.e.

$$P(z_{k+1}|z^k) = P(z_{k+1}|z_{k-m+1}^k). \quad (6)$$

■

The causality matrix in (5) depends on the cross-covariances between x and z and the autocovariances of z , but not on the autocovariances of x . For $k \geq \ell - 1 \geq m$, the matrices $\Gamma_{z^*, \tilde{x}}$, $\Gamma_{z^*, \tilde{z}}$, and Γ_{z^*, z^o} have size $(m+1) \times m$, $(m+1) \times m$, and $(m+1) \times (\ell - m)$, respectively.

The causality matrix $C_G^{x \rightarrow z}(m, k)$ has a fixed number of rows $(m+1)$ with a growing number $k+m$ of columns as $k > m$. Hence, Granger causality inference over a growing horizon $k+m$ is not a simple task, as expected. On the other hand, a sufficient condition to verify that x Granger causes z can be derived as follows:

Lemma 2. (Ahmadi et al., 2019) Let x, z be jointly Gaussian, stationary, scalar stochastic processes and z be partial Markov- m . Assume that $[x_{k-m+1}^k z^k]$ has positive definite covariance matrix at given time $k > m$, i.e. no deterministic relationship exists between x_{k-m+1}^k and z^k .

If there exists some $q \in (m, k]$ such that the matrix $C_G^{x \rightarrow z}(m, q)$ is full rank, then x Granger causes z . □

In the following sections, we use the rank-based approach to infer Granger causality between signals in the presence of noise and filtering.

3. GRANGER CAUSALITY INFERENCE USING DATA CORRUPTED BY ADDITIVE NOISE

In this section, the effect of noise on inferring Granger causality (GC) is investigated.

Consider two jointly Gaussian stationary stochastic processes x and z observed in noise:

$$x'_k := x_k + e_k^x, \quad (7)$$

$$z'_k := z_k + e_k^z. \quad (8)$$

The noise terms e_k^x and e_k^z are stationary, but can be correlated to each other and the signals x_k and z_k , and are not necessarily Gaussian. The signals x_k, z_k and the noise terms are jointly stationary.

From (5) it is obvious that if only the signal x is corrupted by noise independent of z , then the noise does not change the rank of the causality matrix $C_G^{x \rightarrow z}(m, k)$ because this matrix does not depend on the auto-covariances of x .

In the following, we introduce conditions under which GC between the original signals is not lost due to additive noise.

The covariances of the observed signals (x'_k and z'_k) are given by:

$$\gamma_{z'x'}(\kappa) = \gamma_{zx}(\kappa) + \gamma_{ze^x}(\kappa) + \gamma_{e^z x}(\kappa) + \gamma_{e^z e^x}(\kappa), \quad (9)$$

$$\gamma_{z'z'}(\kappa) = \gamma_{zz}(\kappa) + \gamma_{ze^z}(\kappa) + \gamma_{e^z z}(\kappa) + \gamma_{e^z e^z}(\kappa). \quad (10)$$

In order to infer GC between x and z , we need to obtain the covariances included in causality matrix $C_G^{x \rightarrow z}(m, q)$. The relation between causality matrix $C_G^{x \rightarrow z}(m, q)$ and the matrix created by the covariances of observed signals ($C^{x' \rightarrow z'}(m, q)$) is as follows:

$$C^{x' \rightarrow z'}(m, q) := C_G^{x \rightarrow z}(m, q) + C'(m, q), \quad (11)$$

where $C'(m, q)$ is obtained using (9) and (10).

Note that we need to find the rank of the causality matrix $C_G^{x \rightarrow z}(m, q)$, but do not have access to x and z . As we can estimate the auto- and cross-covariances between observed signals x' and z' , the problem of interest is what we can infer about the rank of the causality matrix by using estimates of second-order statistics computed from noisy data. We first state Eckart-Young-Mirsky matrix approximation theorem from linear algebra:

Theorem 3. (Eckart and Young, 1936; Mirsky, 1960) Let the matrix $M \in \mathbb{R}^{l \times s}$ have rank r and singular value decomposition $M = \sum_{i=1}^r \sigma_i u_i v_i^T$, where $u_i, v_j, 1 \leq i, j \leq r$ are orthonormal vectors and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r (> 0)$ are the singular values.

If $p < r$, then

$$\min_{\text{rank}(X)=p} \|M - X\|_2 = \|M - M_p\|_2 = \sigma_{p+1}, \quad (12)$$

where $M_p = \sum_{i=1}^p \sigma_i u_i v_i^T$. ■

Using the Eckart-Young-Mirsky matrix approximation theorem, we know that if the matrix obtained using observed data $C^{x' \rightarrow z'}(m, q)$ is full rank and its smallest singular value satisfies

$$\sigma_{\min}(C^{x' \rightarrow z'}(m, q)) > \|C'(m, q)\|_2, \quad (13)$$

then $C_G^{x \rightarrow z}(m, q)$ remains full rank which means that x Granger causes z using Lemma 2.

Note that both sides of (13) depend on the noise processes. We first investigate whether (13) may ever be satisfied. We can show that if

$$\|C'(m, q)\|_2 < \frac{\sigma_{\min}(C_G^{x \rightarrow z}(m, q))}{2}, \quad (14)$$

then (13) holds. The RHS of (14) only depends on the underlying jointly Gaussian process. Hence, we have shown that there exist noise processes satisfying (13).

Upper bounds on $\|C'(m, q)\|_2$ can be obtained as follows:

$$|\gamma_{z'x'}(\kappa) - \gamma_{zx}(\kappa)| \leq \gamma_{ze^x}^M + \gamma_{e^z x}^M + \gamma_{e^z e^x}^M, \quad (15)$$

$$|\gamma_{z'z'}(\kappa) - \gamma_{zz}(\kappa)| \leq 2\gamma_{ze^z}^M + \gamma_{e^z e^z}^M, \quad (16)$$

where γ_{ab}^M denotes the maximum of absolute value of the covariances between signals a and b .

For $A = [a_{ij}] \in \mathbb{R}^{s \times t}$, we have $\|A\|_2 \leq \sqrt{s} \|A\|_\infty$ and $\|A\|_2 \leq \sqrt{t} \|A\|_1$ where $\|A\|_\infty := \max_{1 \leq i \leq s} \sum_{j=1}^t |a_{ij}|$, and $\|A\|_1 := \max_{1 \leq j \leq t} \sum_{i=1}^s |a_{ij}|$ (Golub and Van Loan, 1996). Therefore, we have:

$$\|C'(m, q)\|_\infty \leq m(\gamma_{ze^x}^M + \gamma_{e^z x}^M + \gamma_{e^z e^x}^M) + q(2\gamma_{ze^z}^M + \gamma_{e^z e^z}^M), \quad (17)$$

and moreover:

$$\|C'(m, q)\|_1 \leq (m+1) \max \left\{ \gamma_{ze^x}^M + \gamma_{e^z x}^M + \gamma_{e^z e^x}^M, 2\gamma_{ze^z}^M + \gamma_{e^z e^z}^M \right\}. \quad (18)$$

It follows that

Theorem 4. Let x, z be jointly Gaussian, stationary, scalar stochastic processes, and let z be partial Markov- m . Assume that $[x_{k-m+1}^k z^k]$ has positive definite covariance matrix at a given time $k > m$.

Suppose there exists some $q \in (m, k]$ such that the matrix $C^{x' \rightarrow z'}(m, q)$, involving covariances of the observed signals x'_k and z'_k , is full-rank. Then x Granger causes z , provided that one of the two following conditions is satisfied:

$$m(\gamma_{ze^x}^M + \gamma_{e^z x}^M + \gamma_{e^z e^x}^M) + q(2\gamma_{ze^z}^M + \gamma_{e^z e^z}^M) < \frac{\sigma'_{\min}}{\sqrt{m+1}}, \quad (19)$$

or

$$\max \{ \gamma_{ze^x}^M + \gamma_{e^z x}^M + \gamma_{e^z e^x}^M, 2\gamma_{ze^z}^M + \gamma_{e^z e^z}^M \} < \frac{\sigma'_{\min}}{(m+1)\sqrt{m+q}}, \quad (20)$$

where σ'_{\min} is the smallest singular value of $C^{x' \rightarrow z'}(m, q)$ and γ^M is the maximum of absolute value of the covariances between the corresponding signals. ■

Remark 1. Theorem 4 states that Granger causality of the stochastic processes can be inferred through the second-order statistics of noisy data provided that the perturbation due to noisy data measured by $\|C'(m, q)\|_2$ is smaller than σ'_{\min} . Under such circumstances, the full rank of $C^{x' \rightarrow z'}(m, q)$ implies full rank of $C_G^{x \rightarrow z}(m, q)$. Thus x Granger causes z by Lemma 2. Note that we cannot interpret the rank of $C^{x' \rightarrow z'}(m, q)$ as an indicator of Granger causality between noisy signals since the assumptions in Theorem 1 and Lemma 2 are not necessarily satisfied anymore. However, we use this matrix to investigate the rank of the causality matrix $C_G^{x \rightarrow z}(m, q)$, which is a criterion to infer whether x Granger causes z as stated in Theorem 1 and Lemma 2.

Remark 2. In Theorem 4, the noise is not necessarily Gaussian. They can be dependent on each other and/or on the stochastic processes x and z . However, we need *a priori* knowledge of the maximum covariances between such noises and between the noises and the processes x and z to infer Granger causality.

Note that both sides of (19) and (20) depend on the second-order statistics of the noise. In the following, we derive a condition in terms of the underlying Gaussian stochastic process. Using (18) and (14), we can find a sufficient condition on the second-order noise statistics as follows, such that we can infer the causality using second-order statistics of corrupted signals through Theorem 4.

$$\max \{ \gamma_{ze^x}^M + \gamma_{e^z x}^M + \gamma_{e^z e^x}^M, 2\gamma_{ze^z}^M + \gamma_{e^z e^z}^M \} < \frac{\sigma_{\min}(C_G^{x \rightarrow z}(m, q))}{2(m+1)\sqrt{m+q}}. \quad (21)$$

Note that the RHS of (21) depends just on the underlying jointly Gaussian stochastic process and its LHS depends on the second-order statistics of noises.

Now let us consider the case where the sum of covariances between the additive noise and the underlying processes decay exponentially in κ , i.e.

$$|\gamma_{ze^x}(\kappa) + \gamma_{e^z x}(\kappa) + \gamma_{e^z e^x}(\kappa)| \leq a_{zx} e^{-b_{zx}|\kappa|}, \quad (22)$$

and

$$|\gamma_{ze^z}(\kappa) + \gamma_{e^z z}(\kappa) + \gamma_{e^z e^z}(\kappa)| \leq a_{zz} e^{-b_{zz}|\kappa|}. \quad (23)$$

where $a_{zx}, a_{zz}, b_{zx}, b_{zz} > 0$. Note that such bounds can be obtained by imposing similar bounds on $\gamma_{ze^x}(\kappa), \gamma_{e^z x}(\kappa), \gamma_{e^z e^x}(\kappa), \gamma_{ze^z}(\kappa), \gamma_{e^z z}(\kappa), \gamma_{e^z e^z}(\kappa)$ individually.

For the stochastic processes with exponentially decaying covariances, tighter upper bounds on the norm of $C'(m, q)$ can be obtained. We have the following result:

Theorem 5. Let x, z be jointly Gaussian, stationary, scalar stochastic processes, and let z be partial Markov- m . Assume that $[x_{k-m+1}^k z^k]$ has positive definite covariance matrix at a given time $k > m$ and that (22)-(23) hold. Suppose there exists some $q \in (m, k]$ such that the matrix $C^{x' \rightarrow z'}(m, q)$, involving covariances of the observed signals x'_k and z'_k , is full-rank. Then x Granger causes z , provided that one of the two following conditions is satisfied:

$$\max \{ f(a_{zx}, b_{zx}, m, i) + f(a_{zz}, b_{zz}, q, i) \} < \frac{\sigma'_{\min}}{\sqrt{m+1}}, \quad (24)$$

or

$$\max \{ g(a_{zx}, b_{zx}), g(a_{zz}, b_{zz}) \} < \frac{\sigma'_{\min}}{\sqrt{m+q}}, \quad (25)$$

where the maximum taken over $i = 0, \dots, m - \lceil \frac{m+1}{2} \rceil$ in (24), $\lceil \cdot \rceil$ is the ceiling function, σ'_{\min} is the smallest singular value of $C^{x' \rightarrow z'}(m, q)$,

$$f(a, b, c, d) := a \frac{e^{b(\lceil \frac{m+1}{2} \rceil + d)}(e^{-b} + e^{-bc}) - e^b - 1}{1 - e^b}, \quad (26)$$

$$g(a, b) := a \frac{e^{-b(m - \lceil \frac{m-1}{2} \rceil)} - e^b - 1 + e^{-b\lceil \frac{m-1}{2} \rceil}}{1 - e^b}, \quad (27)$$

and $a_{zx}, a_{zz}, b_{zx}, b_{zz} > 0$ are defined in (22)-(23). ■

Remark 3. Theorem 5 is a refined version of Theorem 4 which introduces tighter upper bounds on $\|C'(m, q)\|_2$ in (13) while excluding periodic noise processes. Granger causality of such periodic processes can be inferred through Theorem 4.

In the following section, the impact of filtering on inferring Granger causality is addressed.

4. GRANGER CAUSALITY INFERENCE UNDER FILTERING

Consider filtered jointly Gaussian stationary stochastic processes x and z :

$$x_k^f = h_{xx}(k) * x_k + h_{xz}(k) * z_k, \quad (28)$$

$$z_k^f = h_{zx}(k) * x_k + h_{zz}(k) * z_k, \quad (29)$$

where x_k^f and z_k^f denote the filtered signals, $h_{xx}(k), h_{xz}(k), h_{zx}(k)$ and $h_{zz}(k)$ are impulse responses of stable filters, and $*$ denotes convolution operator. The filters can be non-causal and non-minimum phase.

Let us consider the relation between the causality matrix $C_G^{x \rightarrow z}(m, q)$ of the unfiltered processes, and its counterpart $C^{x^f \rightarrow z^f}(m, q)$ constructed from the filtered signals. The difference between these two matrices is called $C^f(m, q) := C_G^{x \rightarrow z}(m, q) - C^{x^f \rightarrow z^f}(m, q)$. Theorem 3 implies that if the smallest singular value of $C^{x^f \rightarrow z^f}(m, q)$ is greater than $\|C^f(m, q)\|_2$, then the causality matrix $C_G^{x \rightarrow z}(m, q)$ constructed by the signals before filtering remains full

rank. By Lemma 2, this implies that x Granger causes (GC) z .

Next we find an upper bound on $\|C^f(m, q)\|_2$. To do so, we investigate the perturbations $\gamma_{z^f x^f}(\kappa) - \gamma_{zx}(\kappa)$ and $\gamma_{z^f z^f}(\kappa) - \gamma_{zz}(\kappa)$. Since the filters are stable and the jointly Gaussian stochastic processes are stationary, the covariances between filtered signals, which can be obtained by standard linear convolution in terms of the covariances of the unfiltered signals, in the form of infinite sums are convergent. In the following, upper bounds on the difference between the covariances of the filtered signals and the covariances of the original signals are derived.

An upper bound $\Delta_{zx}(\kappa)$ on $|\gamma_{z^f x^f}(\kappa) - \gamma_{zx}(\kappa)|$ can be obtained as follows:

$$\begin{aligned} \Delta_{zx}(\kappa) := & \sum_{i=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |h_{xx}(i)h_{zx}(n)| |\gamma_{xx}(n + \kappa - i)| \\ & + \sum_{i=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} |h_{xx}(i)h_{zz}(r)| |\gamma_{zx}(r + \kappa - i)| I_{r \neq i} \\ & + \sum_{j=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |h_{xz}(j)h_{zx}(n)| |\gamma_{zx}(j - n - \kappa)| I_{j \neq n + 2\kappa} \\ & + \sum_{j=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} |h_{xz}(j)h_{zz}(r)| |\gamma_{zz}(-\kappa + j - r)| \\ & + |\gamma_{zx}(\kappa)| \left| \sum_{j=-\infty}^{\infty} h_{xz}(j)h_{zx}(j - 2\kappa) + \right. \\ & \left. \sum_{i=-\infty}^{\infty} h_{xx}(i)h_{zz}(i) - 1 \right|, \end{aligned} \quad (30)$$

where I is the indicator function. And upper bound $\Delta_{zz}(\kappa)$ on $|\gamma_{z^f z^f}(\kappa) - \gamma_{zz}(\kappa)|$ can be derived as follows:

$$\begin{aligned} \Delta_{zz}(\kappa) := & \sum_{n=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} |h_{zx}(n)h_{zx}(i)| |\gamma_{xx}(\kappa - i + n)| + \\ & \sum_{n=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} |h_{zx}(n)h_{zz}(j)| |\gamma_{zx}(j - \kappa - n)| + \\ & \sum_{r=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} |h_{zz}(r)h_{zx}(i)| |\gamma_{zx}(\kappa - i + r)| + \\ & \sum_{r=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} |h_{zz}(r)h_{zz}(j)| |\gamma_{zz}(\kappa - j + r)| I_{j \neq r} + \\ & |\gamma_{zz}(\kappa)| \sum_{r=-\infty}^{\infty} |h_{zz}^2(r) - 1|. \end{aligned} \quad (31)$$

Let us now assume that maxima of the absolute values of auto- and cross-covariances of the original processes are known a priori (γ_{xx}^M , γ_{zx}^M and γ_{zz}^M). By replacing the auto- and cross-covariances in (30) and (31) with their maxima, we can find upper bounds on $\Delta_{zx}(\kappa)$ and $\Delta_{zz}(\kappa)$ which do not depend on κ , and are denoted by Δ_{zx}^M and Δ_{zz}^M , respectively. We can then obtain upper bounds on the norms of the perturbation between the causality matrix $C_G^{x \rightarrow z}(m, q)$ and $C^{x^f \rightarrow z^f}(m, q)$ using similar techniques as in Section 3 as follows:

$$\|C^f(m, q)\|_{\infty} \leq m\Delta_{zx}^M + q\Delta_{zz}^M, \quad (32)$$

and

$$\|C^f(m, q)\|_1 \leq (m + 1) \max \{ \Delta_{zx}^M, \Delta_{zz}^M \}. \quad (33)$$

Hence, we have:

Theorem 6. Let x, z be jointly Gaussian stationary, scalar stochastic processes, and let z be partial Markov- m . Assume that $[x_{k-m+1}^k z^k]$ has positive definite covariance matrix at a given time $k > m$.

Suppose there exists some $q \in (m, k]$ such that the matrix $C^{x^f \rightarrow z^f}(m, q)$, involving covariances of the filtered signals, is full-rank. Then x Granger causes z , provided that one of the two following conditions is satisfied:

$$m\Delta_{zx}^M + q\Delta_{zz}^M < \frac{\sigma_{\min}^f}{\sqrt{m+1}}, \quad (34)$$

or

$$\max \{ \Delta_{zx}^M, \Delta_{zz}^M \} < \frac{\sigma_{\min}^f}{(m+1)\sqrt{m+q}}, \quad (35)$$

where σ_{\min}^f is the smallest singular value of $C^{x^f \rightarrow z^f}(m, q)$ and Δ_{zx}^M and Δ_{zz}^M are the upper bounds on (30) and (31) derived using the maximum magnitudes of the covariances (γ_{xx}^M , γ_{zx}^M and γ_{zz}^M). ■

Remark 4. Note that conditions (34) and/or (35) guarantee that if the matrix $C^{x^f \rightarrow z^f}(m, q)$ is full rank, then the causality matrix $C_G^{x \rightarrow z}(m, q)$, which is constructed by the second-order statistics of the underlying Gaussian stochastic processes, is also full rank. Lemma 2 then implies that x Granger causes z . Note that the matrix $C^{x^f \rightarrow z^f}(m, q)$ cannot be used to infer Granger causality between filtered signals due to violations of assumptions of Theorem 1 and Lemma 2. We use this matrix as a mathematical object carrying information about the rank of the causality matrix $C_G^{x \rightarrow z}(m, q)$.

Depending on the underlying jointly Gaussian stochastic process, the filters can be chosen such that we are able to infer the Granger causality through Theorem 6. As mentioned in Section 3, we have the relation

$$\|C^f(m, q)\|_2 < \frac{\sigma_{\min}(C_G^{x \rightarrow z}(m, q))}{2} \quad (36)$$

guaranteeing that $\sigma_{\min}(C^f(m, q)) > \|C^f(m, q)\|_2$. Therefore, if the parameters of the filters satisfy

$$\max \{ \Delta_{zx}^M, \Delta_{zz}^M \} < \frac{\sigma_{\min}(C_G^{x \rightarrow z}(m, q))}{2(m+1)\sqrt{m+q}}, \quad (37)$$

then the sufficient condition (35) in Theorem 6 holds.

Remark 5. If the covariances decay to zero with lag κ , tighter conditions can be obtained. For instance, if they decay exponentially with κ , then the sufficient conditions in Theorem 6 become

$$\max_{i=0, \dots, m} \left\{ \sum_{\kappa=1-i}^{m-i} \Delta_{zx}^f(\kappa) + \sum_{\kappa=1-i}^{q-i} \Delta_{zz}^f(\kappa) \right\} < \frac{\sigma_{\min}^f}{\sqrt{m+1}}, \quad (38)$$

or

$$\max_{i=1, \dots, m} \left\{ \sum_{\kappa=i}^{i-m} \Delta_{zx}^f(\kappa), \sum_{\kappa=j}^{m-j} \Delta_{zz}^f(\kappa) \right\} < \frac{\sigma_{\min}^f}{\sqrt{m+q}}. \quad (39)$$

where $\Delta_{zx}^f(\kappa)$ and $\Delta_{zz}^f(\kappa)$ are bounds on the covariance magnitudes.

Remark 6. Filters in the approach introduced in this paper can be noncausal and/or nonminimum phase. It is

required to know bounds on the magnitudes of the correlation coefficients of the underlying jointly Gaussian stochastic processes a priori to infer the Granger causality.

Remark 7. Note that the approach can be used for non-linear filters as well e.g. Volterra filters.

5. CONCLUSION

This paper studies Granger causality between jointly Gaussian, partially Markov- m signals using the fact that Granger causality between such stochastic processes can be determined by a full-rank condition of a matrix constructed by covariances between the signals.

Exploiting the properties of the rank-based condition for Granger causality, the Granger causality between signals corrupted by additive noise terms was investigated. Sufficient conditions involving the second-order statistics of the noisy signals were derived which guaranteed Granger causality between the noise free jointly Gaussian stochastic processes.

Furthermore, impacts of filtering of stochastic processes were investigated. Sufficient conditions were introduced under which Granger causality of the stochastic processes before filtering can be inferred using the second-order statistics of the filtered signals. The stable filters can be non-causal and nonminimum-phase. Note that the approach introduced in this paper can be exploited to investigate Granger causality of jointly Gaussian processes under the simultaneous impacts of both noise and filtering.

The approach does not require the statistics of the underlying Gaussian signals to be estimated, or a system model to be identified, unlike most literature addressing the inference of Granger causality.

REFERENCES

- Ahmadi, S., Nair, G.N., and Weyer, E. (2019). Granger causality of Gaussian signals from quantized measurements. In *58th IEEE Conference on Decision and Control*, 3587–3592. IEEE.
- Anderson, B.D.O., Deistler, M., and Dufour, J.M. (2019). On the sensitivity of Granger causality to errors-in-variables, linear transformations and subsampling. *Time Ser. Anal.*, 40, 102–123.
- Barnett, L. and Seth, A.K. (2011). Behaviour of Granger causality under filtering: theoretical invariance and practical application. *J. Neurosci. Methods*, 201, 404–419.
- Caines, P. and Chan, C. (1975). Feedback between stationary stochastic processes. *IEEE Trans. Autom. Contr.*, 20, 498–508.
- Eckart, C. and Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1, 211–218.
- Florin, E., Gross, J., Pfeifer, J., Fink, G.R., and Timmermann, L. (2010). The effect of filtering on Granger causality based multivariate causality measures. *Neuroimage*, 50, 577–588.
- Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *J. Am. Stat. Assoc.*, 77, 304–313.
- Golub, G. and Van Loan, C. (1996). *Matrix computations*. Johns Hopkins University Press, Baltimore.
- Granger, C.W.J. (1963). Economic process involving feedback. *Inf. Control*, 6, 28–48.
- Granger, C.W.J. (1980). Testing for causality: a personal viewpoint. *J. Econ. Dyn. Control*, 2, 329–352.
- Granger, C.W.J. (1988). Some recent development in a concept of causality. *J. Econom.*, 39, 199–211.
- Józsa, M., Petreczky, M., and Camlibel, M.K. (2019). Relationship between Granger noncausality and network graph of state-space representations. *IEEE Trans. Autom. Contr.*, 64, 912–927.
- Kontoyiannis, I. and Skoularidou, M. (2016). Estimating the directed information and testing for causality. *IEEE Trans. Inform. Theory*, 62, 6053–6067.
- Mirsky, L. (1960). Symmetric gauge functions and unitarily invariant norms. *Quart. J. Math. Oxford*, 11, 50–59.
- Nalatore, H., Ding, M., and Rangarajan, G. (2007). Mitigating the effects of measurement noise on Granger causality. *Phys. Rev. E*, 75, 031123.
- Nalatore, H., Sasikumar, N., and Rangarajan, G. (2014). Effect of measurement noise on Granger causality. *Phys. Rev. E*, 90, 062127.
- Newbold, P. (1978). Feedback induced by measurement errors. *Int. Econ. Rev.*, 19, 787–791.
- Quinn, C.J., Coleman, T.P., Kiyavash, N., and Hatsopoulos, N.G. (2011). Estimating the directed information to infer causal relationships in ensemble neural spike train recordings. *J. Comput. Neurosci.*, 30, 17–44.
- Quinn, C.J., Kiyavash, N., and Coleman, T.P. (2015). Directed information graphs. *IEEE Trans. Inform. Theory*, 61, 6887–6909.
- Rangarajan, P. and Rao, R.P.N. (2019). Estimation of vector autoregressive parameters and Granger causality from noisy multichannel data. *IEEE Trans. Biomed. Eng.*, 66, 2231–2240.
- Sargent, T.J. (1987). *Macroeconomic theory*. Academic Press, Boston.
- Seth, A.K., Chorley, P., and Barnett, L.C. (2013). Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. *Neuroimage*, 65, 540–555.
- Smirnov, D.A. (2013). Spurious causalities with transfer entropy. *Phys. Rev. E*, 87, 042917.
- Solo, V. (2007). On causality I: sampling and noise. In *46th IEEE Conference on Decision and Control*, 3634–3639. IEEE.
- Solo, V. (2016). State-space analysis of Granger-Geweke causality measures with application to fMRI. *Neural Comput.*, 26, 914–949.
- Wiener, N. (1956). *The theory of prediction*. McGraw-Hill, New York.
- Winkler, I., Panknin, D., Bartz, D., Müller, K.R., and Haufe, S. (2016). Validity of time reversal for testing Granger causality. *IEEE Trans. Signal Process.*, 64, 2746–2760.