

# Granger Causality of Gaussian Signals from Quantized Measurements

Salman Ahmadi, Girish N. Nair and Erik Weyer

**Abstract**—An approach is proposed for inferring Granger causality between jointly stationary, Gaussian signals from quantized data. First, a necessary and sufficient rank criterion for the equality of two conditional Gaussian distributions is proved. Assuming a partial finite-order Markov property, sufficient conditions are then derived under which Granger causality between them can be reliably inferred from the second order moments of the quantized processes. This approach does not require the statistics of the underlying Gaussian signals to be estimated, or a system model to be identified.

## I. INTRODUCTION

Causal inference is the determination of the qualitative cause-and-effect (or input versus output) relationships between two or more different signals over time. Sometimes these relationships are obvious beforehand, but in many critical applications they are not. For instance, in environmental monitoring the direction in which a pollutant spreads may be unknown to begin with, making it difficult to determine *a priori* which measurements are inputs and which are outputs. In large manufacturing plants, the root cause of alarm signals is commonly obscured by complex feedback loops. A misunderstanding of the correct causal relationships not only reduces the accuracy of the subsequently identified model, but could mislead decision-makers into poorly founded interventions.

In 1963, the econometrician C. Granger introduced a definition of causality in terms of statistical prediction [1], inspired by the work of N. Wiener [2]. A signal  $x$  is said to cause another signal  $z$  if at some time, the optimal expected prediction error for a future value of  $z$  is reduced by knowledge of past  $x$  and  $z$ , as compared to if the past values only of  $z$  are known. In subsequent work [3], [4], Granger proposed a looser definition in terms of conditional probabilities, whereby  $x$  is said to Granger cause (GC)  $z$  if, at some time, a future  $z$  and past  $x$  are conditionally dependent given past  $z$ ; i.e., given past  $z$ , past  $x$  can still influence the future of  $z$ . In the case of jointly Gaussian processes under a mean-square error prediction error, the first definition coincides with the second. These definitions allow causality from  $z$  to  $x$  as well, which would reflect mutual coupling between the two processes as they evolve over time.

As defined above, Granger causality is a ‘non-interventionist’ notion based on signals rather than systems. This suits applications where signals can only be measured,

and cannot easily be adjusted through experiments. In [5], a connection with linear systems theory is introduced through the idea of *feedback-freeness* for wide-sense stationary random processes, which is shown to be equivalent to Granger non-causality under linear minimum mean-square error prediction. A more restrictive version of feedback-free random processes having block-diagonal innovation covariance is introduced in [6] and its equivalence to strictly causal linear systems is discussed. The effect of sampling rate on Granger causality is investigated in [7]–[9] while the impact of filtering is discussed in [8], [9] and the effects of noise are addressed in [7], [9]–[11].

In this paper, we propose a method for inferring Granger causality between two jointly Gaussian, stationary signals using quantized measurements. Evidently, the nonlinearity introduced by the quantizers moves this problem beyond the linear systems realm of the literature above. The basis of our analysis is a necessary and sufficient rank condition for two conditional Gaussian distributions to be identical (Theorem 2.3). This extends a recent result of [12], [13] on Gaussian conditional independence and, to the best of our knowledge, has not appeared in the literature before. Assuming a *partial finite-order Markov* property, we then use a perturbation bound to derive sufficient conditions under which Granger causality between the unquantized signals can be inferred from the second order moments of their quantized versions (Theorem 3.3). In the case of high-resolution, uniform quantization, this condition reduces to an explicit inequality relating the quantizer resolutions to the statistics of the processes (Proposition 3.6 and Remark 3.7).

Unlike much of the literature on causal inference, e.g. [14]–[16], our approach does not require the statistics of the underlying Gaussian signals to be estimated, or a system model to be identified.

The rest of the paper is organized as follows: in section II, the problem of inferring Granger causality between a pair of jointly Gaussian signals is analyzed in terms of the rank of a matrix of covariances. In section III, the effects of quantization are investigated and conditions are presented under which the Granger causality can be inferred correctly from the statistics of the quantized data. Section IV concludes the paper.

**Notation:** Throughout this paper, we denote the random process segments  $(x_k)_{k=\ell}^n$  by  $x_\ell^n$ , and  $(x_k)_{k=1}^n$  by  $x^n$ . We use the conventions that  $x_\ell^n = (x_k)_{k=1}^n$  when  $\ell \leq 1$ , and equals the empty sequence when  $\ell > n$  or  $n < 1$ . Similarly,  $x^n$  is the empty sequence when  $n < 1$ . When clear from context, the full sequence  $(x_k)_{k=1}^\infty$  is written as  $x$ .

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

The authors are with the Department of Electrical and Electronic Engineering, University of Melbourne, VIC 3010, Australia. E-mails: ahmadis@student.unimelb.edu.au, {gnair, ewey}@unimelb.edu.au

## II. GRANGER CAUSALITY FOR JOINTLY GAUSSIAN, PARTIALLY MARKOV PROCESSES

In this section, we give the formal definition of Granger causality between discrete-time stochastic processes, in terms of conditional independence. Assuming that the process is *partially finite-order Markov*, this reduces to comparing two conditional distributions. Under the additional assumption of joint Gaussianity, we show that this comparison can be expressed in terms of the rank of a structured matrix of covariances.

Let  $x, z$  be discrete-time random processes on the time-axis  $k = 1, 2, \dots$ .

*Definition 1 (Granger Causality [3]):* The random process  $x$  is said to not *Granger cause* (GC)  $z$  if

$$P(z_{k+1}|x^k, z^k) = P(z_{k+1}|z^k), \quad (1)$$

with probability (w.p.) 1,  $k = 1, 2, \dots$ ,

where  $P(\cdot|\cdot)$  denotes conditional probability measure. Otherwise, if for some  $k \geq 1$  there is a nonzero probability that  $P(z_{k+1}|x^k, z^k) \neq P(z_{k+1}|z^k)$ , then  $x$  is said to GC  $z$ .

In other words,  $x$  GC  $z$  means that there is a nonzero chance that at some time  $k$ , the next value of  $z$  could still be stochastically influenced by past and present  $x$ , even if all past and present values of  $z$  up to time  $k$  are known. If  $x$  does not GC  $z$ , the next value of  $z$  is always conditionally independent of past and present  $x$ , given past and present  $z$ .

Now let us assume the following:

*Assumption 2.1 (Partial Markov- $m$ ):* The random process  $z$  is said to be partially Markov of order  $m \geq 1$  in  $x$  and  $z$  if  $P(z_{k+1}|x^k, z^k) = P(z_{k+1}|x_{k-m+1}^k, z_{k-m+1}^k)$ .

*Remark 2.2:* This is weaker than being joint Markov- $m$  since the  $x$  process is not assumed to exhibit an analogous property.

With this restriction, we may rewrite the condition for  $x$  to not Granger cause  $z$  as

$$P(z_{k+1}|x_{k-m+1}^k, z_{k-m+1}^k) = P(z_{k+1}|z_{k-m+1}^k, z_{k-m}^k), \quad \text{w.p. } 1, k = 1, 2, \dots \quad (2)$$

As the conditioning term on the RHS is no longer nested (i.e. included) in that of the LHS, this is no longer a conditional independence relationship. In much of the literature e.g. [17]–[19], it is assumed that under non-causality the process  $z$  is also Markov- $m$ , so that  $z^{k-m}$  can be dropped from the RHS. However, finite-order partial Markovianity does not generally imply finite-order marginal Markovianity. In this paper, we do not make any *a priori* assumption on the Markovianity of  $z$ , but show that for jointly Gaussian signals, (2) and Assumption 2.1 imply that  $z$  is marginally Markov- $m$ , under a mild additional requirement.

For jointly Gaussian processes, it turns out that (2) can be expressed in terms of a rank condition. We first present this condition for general jointly Gaussian random vectors  $A, B, C, D$  with joint covariance matrix  $\Gamma$ . The covariance matrices between subsets of variables are denoted by matching subscripts, for instance  $\Gamma_{A,B} = \Gamma_{B,A}^T$  is the cross-

covariance matrix between  $A$  and  $B$ , while  $\Gamma_{[CD],[CD]}$  is the covariance matrix for the random vector  $[CD]$ .<sup>1</sup>

*Theorem 2.3:* Let  $A, B, C, D$  be jointly Gaussian random vectors with joint covariance matrix  $\Gamma$ , and suppose that the random vector  $[BCD]$  has positive definite covariance. Then

- 1) the conditional distributions  $P(A|C, B)$  and  $P(A|C, D)$  are identical if and only if

$$\text{rank}\left(\begin{bmatrix} \Gamma_{A,B} & \Gamma_{A,C} & \Gamma_{A,D} \\ \Gamma_{C,B} & \Gamma_{C,C} & \Gamma_{C,D} \end{bmatrix}\right) = \#C, \quad (3)$$

where  $\#C$  is the dimension of the random vector  $C$ , and

- 2) if (3) holds, then  $P(A|C, B) = P(A|C, D) = P(A|C)$ .

*Proof:* Refer to Appendix. ■

*Remark 2.4:* The first item in this result is a variation of a recent rank formula for Gaussian conditional independence, due to Sullivant [12], [13]. Under joint Gaussianity, the positive definiteness of  $\Gamma_{[BCD],[BCD]}$  excludes degenerate cases where deterministic (affine) relationships exist between  $B, C$  and  $D$ . With additional analysis, it can be shown that if  $D$  is constant (requiring this positive definiteness to be relaxed), then the formula of [12], [13] for conditional independence can be recovered as a special case. However, this extension is not needed for our present purposes. The second item in this result shows that if  $P(A|C, B) = P(A|C, D)$ , then conditional independence between  $A, B$  and between  $A, D$  immediately follow, given  $C$ .

Now suppose that  $x, z$  are jointly stationary Gaussian signals satisfying Assumption 2.1. Further assume that they are each scalar-valued, for simplicity, and that there is never a deterministic relationship between  $x_{k-m+1}^k$  and  $z^k$ . We may then apply Theorem 2.3 to (2) with  $A = z_{k+1}$ ,  $B = x_{k-m+1}^k$ ,  $C = z_{k-m+1}^k$  and  $D = z_{k-m}^k$ , to conclude that  $x$  does not GC  $z$  if and only if

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

where for fixed  $k$  and any  $\ell \geq 1$  we define a matrix

$$C_G^{x \rightarrow z}(m, \ell) := \begin{bmatrix} \Gamma_{z^*, \tilde{x}} & \Gamma_{z^*, \tilde{z}} & \Gamma_{z^*, z^o} \end{bmatrix} \quad (5)$$

of covariances, where

$$z^* := z_{k-m+1}^{k+1}, \quad \tilde{z} := z_{k-m+1}^k, \quad z^o := z_{k-\ell+1}^{k-m}, \quad \tilde{x} := x_{k-m+1}^k. \quad (6)$$

When  $k \geq \ell - 1 \geq m$ , the covariances  $\Gamma_{z^*, \tilde{x}}$  and  $\Gamma_{z^*, \tilde{z}}$  are  $(m+1) \times m$  matrices, and  $\Gamma_{z^*, z^o}$  is an  $(m+1) \times (\ell - m)$  matrix.

For jointly stationary Gaussian signals  $x$  and  $z$  satisfying Assumption 2.1, we call  $C_G^{x \rightarrow z}(m, \ell)$  the *causality matrix* from  $x$  to  $z$  over the interval  $[k - \ell + 1, \dots, k]$ . This matrix depends on the cross-covariances between  $x$  and  $z$  and the autocovariances of  $z$ , but not on the autocovariances of  $x$ . Notice that as it has a fixed number  $m+1$  of rows, but a growing number  $k+m$  of columns as  $k$  increases above

<sup>1</sup>For convenience we adopt this mild abuse of notation, rather than  $[C^T D^T]^T$ .

$m$ . Thus, confirming that  $x$  does not GC  $z$  over a growing horizon is not a simple task, as expected. On the other hand, a sufficient condition to verify that  $x$  GC  $z$  can be derived as follows:

**Lemma 2.5:** Let  $x, z$  be jointly Gaussian, stationary, scalar random processes and  $z$  be partial Markov- $m$ . Assume that  $[x_{k-m+1}^k z^k]$  has positive definite covariance 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  $z$ .

*Proof:* Suppose by hypothesis that the matrix  $C_G^{x \rightarrow z}(m, q)$  has full rank  $(m+1)$  for  $q \in (m, k]$ . Increasing  $q$  appends columns to  $C_G^{x \rightarrow z}(m, q)$  (to the partitioned matrix of  $[\Gamma_{A,D}^T \Gamma_{C,D}^T]^T$  in (3)), which can only keep the matrix full-rank. Hence  $C_G^{x \rightarrow z}(m, k)$  must also have rank  $m+1$ . This violates the noncausality condition (4), so  $x$  must GC  $z$ . ■

**Remark 2.6:** A distinguishing feature of this result is that it allows Granger causality to be inferred directly from the second-order statistics of the signals, without having to analyse or fit a linear dynamical model as is often done in the literature. For computational reasons, it is preferable to have a sufficient  $q$  in this Lemma that remains bounded as  $k$  increases, or at least grows slower than  $k$ . We conjecture that this can be achieved under suitable bounds on the auto- and cross-covariance terms at large time lags.

### III. INFERRING GRANGER CAUSALITY USING QUANTIZED DATA

In this section, we investigate the impact of quantized data on the inference of Granger causality between the jointly Gaussian, stationary processes  $x$  and  $z$  where  $z$  is partially Markov- $m$ . Using the second-order statistics of the quantized data, we construct a post-quantization matrix  $C^{x^Q \rightarrow z^Q}$  that mirrors the causality matrix  $C_G^{x \rightarrow z}$  (5) of the unquantized processes. We then show that if the difference between these two matrices is sufficiently small, then the full-rankness of  $C^{x^Q \rightarrow z^Q}$  implies that  $x$  Granger causes (GC)  $z$ . For high-resolution uniform quantizers, we derive this condition as an explicit inequality, and show that causality can be inferred if the resolution is sufficiently fine.

#### A. Relationship Between $C_G^{x \rightarrow z}$ and $C^{x^Q \rightarrow z^Q}$

We begin with the relationship between the scalar covariances  $\gamma_{z,x}$ ,  $\gamma_{z,z}$  of the unquantized signals, and  $\gamma_{z^Q,x^Q}$ ,  $\gamma_{z^Q,z^Q}$  of their quantized versions, where superscript  $Q$  denotes the quantized version. For convenience, in the following we suppress time lags and use  $w_1$  and  $w_2$  to denote  $x$  and/or  $z$ . We have

$$\gamma_{w_1^Q, w_2^Q} = \gamma_{w_1, w_2} + \gamma_{w_1, \epsilon_2} + \gamma_{\epsilon_1, w_2} + \gamma_{\epsilon_1, \epsilon_2}, \quad (7)$$

where  $\epsilon_1$  and  $\epsilon_2$  denote the quantization errors of  $w_1$  and  $w_2$  respectively. Now define the matrix

$$\begin{aligned} C^{x^Q \rightarrow z^Q}(m, \ell) &:= [\Gamma_{z^*Q\tilde{x}^Q} \quad \Gamma_{z^*Q\tilde{z}^Q} \quad \Gamma_{z^*Qz^Q}] \\ &\equiv C_G^{x \rightarrow z}(m, \ell) + \Gamma_\epsilon(m, \ell), \end{aligned} \quad (8)$$

where

$$\begin{aligned} \Gamma_\epsilon(m, \ell) &:= [\Gamma_{z^*, \epsilon_{\tilde{x}}} \quad \Gamma_{z^*, \epsilon_{\tilde{z}}} \quad \Gamma_{z^*, \epsilon_{z^Q}}] + \\ &[\Gamma_{\epsilon_{z^*}, \tilde{x}} \quad \Gamma_{\epsilon_{z^*}, \tilde{z}} \quad \Gamma_{\epsilon_{z^*}, z^Q}] + \\ &[\Gamma_{\epsilon_{z^*} \epsilon_{\tilde{x}}} \quad \Gamma_{\epsilon_{z^*} \epsilon_{\tilde{z}}} \quad \Gamma_{\epsilon_{z^*} \epsilon_{z^Q}}], \end{aligned} \quad (9)$$

which is the matrix version of the scalar relationship (7).

#### B. Granger Causality Inference Under Quantization

Lemma 2.5 states that if for some  $q > m$  the causality matrix  $C_G^{x \rightarrow z}(m, q)$ , the entries of which correspond to scalar covariances of the unquantized Gaussian random processes, has full rank  $m+1$ , then  $x$  GC  $z$ . The question here is whether we can infer this causal relationship from the covariances of the quantized data. To answer this question, we need a classical result in linear algebra:

**Theorem 3.1:** (Eckart-Young-Mirsky Matrix Approximation [20], [21]) 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

$$\min_{\text{rank}(X)=p} \|M - X\|_2 = \|M - M_p\|_2 = \sigma_{p+1} \quad (10)$$

and

$$\min_{\text{rank}(X)=p} \|M - X\|_F = \|M - M_p\|_F = \sqrt{\sum_{i \geq p+1} \sigma_i^2}, \quad (11)$$

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

**Remark 3.2:** Unless otherwise stated,  $\|\cdot\|$  denotes either two-norm ( $\|\cdot\|_2$ ) or Frobenius norm ( $\|\cdot\|_F$ ) in the following.

**Theorem 3.3:** Let  $x, z$  be jointly Gaussian, stationary, scalar random processes and  $z$  be partial Markov- $m$ . Assume that  $[x_{k-m+1}^k z^k]$  has positive definite covariance at a given time  $k > m$ .

If there exists some  $q \in (m, k]$  such that the matrix  $C^{x^Q \rightarrow z^Q}(m, q)$  (8) involving the covariances of quantized data is full-rank with smallest singular value  $\sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q)) > \|\Gamma_\epsilon(m, q)\|$ , where  $\Gamma_\epsilon(m, q)$  is defined in (9), then the unquantized Gaussian signal  $x$  Granger causes the unquantized Gaussian signal  $z$ .

*Proof:* Denote the full-rank matrix  $C^{x^Q \rightarrow z^Q}(m, q)$  by  $M$ . If  $X$  is any other matrix of the same dimensions but lower rank, then Theorem 3.1 states that  $\|X - M\| \geq \sigma_{\min}(M)$ . Conversely, if  $\|X - M\| < \sigma_{\min}(M)$ , then  $X$  remains full-rank. Setting  $X = C^{x \rightarrow z}(m, q) \equiv C^{x^Q \rightarrow z^Q}(m, q) + \Gamma_\epsilon(m, q)$ , we see that if  $\sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q)) > \|\Gamma_\epsilon(m, q)\|$ , then  $C_G^{x \rightarrow z}(m, q)$  is guaranteed to be full-rank. Lemma 2.5 then implies that  $x$  GC  $z$ . ■

**Remark 3.4:** This result states that GC can be inferred from the statistics of the quantized data, provided that the quantization perturbation, as measured by  $\|\Gamma_\epsilon(m, q)\|$ , is smaller than  $\sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q))$ , which can be taken as a measure of how far  $C^{x^Q \rightarrow z^Q}(m, q)$  is from losing full rank. However, both sides of this inequality depend on

the quantization schemes. We can derive a condition that compares the size of the quantization perturbation to the statistics of the unquantized processes, as follows. We know that  $\sigma_{\min}(A) - \sigma_{\max}(B) \leq \sigma_{\min}(A + B)$ , ([22], Problem 7.3.P16). Therefore, we have:

$$\sigma_{\min}(C_G^{x \rightarrow z}(m, q)) - \sigma_{\max}(\Gamma_\epsilon(m, q)) \leq \sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q)), \quad (12)$$

Furthermore, we know that  $\|\Gamma_\epsilon(m, q)\|_2 = \sigma_{\max}(\Gamma_\epsilon(m, q))$ . So if

$$\|\Gamma_\epsilon(m, q)\|_2 < \sigma_{\min}(C_G^{x \rightarrow z}(m, q)) - \sigma_{\max}(\Gamma_\epsilon(m, q)), \quad (13)$$

then by (12) it is guaranteed to be smaller than  $\sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q))$  as required. Finally, note that since  $\|\Gamma_\epsilon(m, q)\|_2 = \sigma_{\max}(\Gamma_\epsilon(m, q))$ , the condition (13) may be equivalently written as

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

which is feasible with sufficiently fine quantization iff  $C_G^{x \rightarrow z}(m, q)$  is full-rank.

*Remark 3.5:* Note that if  $x$  GC  $z$ , then there exists some  $q$  such that  $C_G^{x \rightarrow z}(m, q)$  is full-rank. In this case, by swapping  $C_G^{x \rightarrow z}$  and  $C^{x^Q \rightarrow z^Q}$  in the proof of Theorem 3.3, it can be shown that  $C^{x^Q \rightarrow z^Q}(m, q)$  remains full-rank if  $\|\Gamma_\epsilon(m, q)\|_2 < \sigma_{\min}(C_G^{x \rightarrow z}(m, q))$ . Thus the full rank of  $C_G$  is preserved going from quantized to unquantized data as well as from unquantized to quantized, if the quantizers are designed to satisfy the stricter condition (14).

### C. Granger Causality Inference Under High-Resolution Quantization

In this section, Theorem 3.3 is explored in the regime of high-resolution quantization.

Suppose first that zero-mean, jointly Gaussian scalar signals  $w_1$  and  $w_2$  are passed through mid-tread uniform quantizers with quantization intervals of length  $\Delta_{w_1}$  and  $\Delta_{w_2}$ , respectively. For simplicity, we assume an infinite number of quantization intervals on each axis. The auto- and cross-covariances of uniformly quantized bivariate Gaussian signals are derived in [23], in the form of infinite sums. By analysing the dominant terms in these sums for sufficiently small  $k_i := \frac{\Delta_{w_i}}{\gamma_{w_i}}$ ,  $i = 1, 2$  where  $\gamma_{w_i}$  (with a single subscript) denotes the standard deviation of  $w_i$ , we can show that the last three terms on the RHS of (7) can be expressed as follows:

$$\gamma_{w_1 \epsilon_2} = O(e^{\frac{-2\pi^2}{k_2^2}}), \quad (15)$$

$$\gamma_{\epsilon_1 w_2} = O(e^{\frac{-2\pi^2}{k_1^2}}), \quad (16)$$

$$\gamma_{\epsilon_1 \epsilon_2} = O(k_1 k_2 e^{-4\pi^2 (\frac{1}{2}(\frac{m_1^f}{k_1})^2 + \frac{1}{2}(\frac{m_2^f}{k_2})^2 - \rho \frac{m_1^f m_2^f}{k_1 k_2})}), \quad (17)$$

where  $(m_1^f, m_2^f)$  is the minimum of  $f(m_1, m_2) := (\frac{m_1}{k_1})^2 + (\frac{m_2}{k_2})^2 - 2\rho \frac{m_1 m_2}{k_1 k_2}$  over positive integer two-dimensional lattice  $(m_1, m_2) \in \mathbb{N}^2$  and  $\rho$  is the correlation coefficient related to the corresponding Gaussian signals. And  $O(\cdot)$  is Bachmann-Landau  $O$ -notation.

Note that (7) for the case where  $w_1$  and  $w_2$  are the same (variance instead of covariance) can be represented as:

$$\gamma_{w_1^Q}^2 = \gamma_{w_1}^2 + \frac{k_1^2 \gamma_{w_1}^2 (= \Delta_{w_1}^2)}{12} + O(k_1^2 \gamma_{w_1}^2 e^{\frac{-2\pi^2}{k_1^2}}). \quad (18)$$

In order to have an explicit relation between the auto and cross-covariances above and  $\sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q))$  mentioned in Theorem 3.3, we know that:

$$\|\Gamma_\epsilon(m, q)\|_2 \leq \sqrt{m+1} \|\Gamma_\epsilon(m, q)\|_\infty, \quad (19)$$

and if we restrict  $\sqrt{m+1} \|\Gamma_\epsilon(m, q)\|_\infty < \sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q))$ , the condition  $\sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q)) > \|\Gamma_\epsilon(m, q)\|_2$  holds. Hence, in the following we try to find an upper bound on  $\sqrt{m+1} \|\Gamma_\epsilon(m, q)\|_\infty$  and then such an upper bound is considered as a lower bound on  $\sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q))$ , which can be calculated by quantized signals.

$$\begin{aligned} \|\Gamma_\epsilon(m, q)\|_\infty &:= \max_{1 \leq i \leq m+1} \sum_{j=1}^{m+q} |\gamma_{\epsilon_{ij}}| \\ &= \max_{2 \leq i \leq m+1} \sum_{j=1}^{m+q} |\gamma_{\epsilon_{ij}}| \\ &\leq m \sup_{\kappa} |\gamma_{x^Q z^Q}(\kappa) - \gamma_{xz}(\kappa)| + (q-1) \times \\ &\quad \sup_{\kappa} |\gamma_{z^Q z^Q}(\kappa) - \gamma_{zz}(\kappa)| + |\gamma_{z^Q}^2 - \gamma_z^2|. \end{aligned} \quad (20)$$

In high resolution quantization regime,  $\|\Gamma_\epsilon(m, q)\|_\infty$  can be written as follows using (15)-(18):

$$\|\Gamma_\epsilon(m, q)\|_\infty < \frac{\Delta_z^2}{12}. \quad (21)$$

We have:

$$\|\Gamma_\epsilon(m, q)\|_2 < \frac{\sqrt{m+1} \Delta_z^2}{12} < \sigma_{\min}(C^{x^Q \rightarrow z^Q}(m, q)). \quad (22)$$

Hence, the following can be stated:

*Proposition 3.6:* Let  $x, z$  be jointly Gaussian, stationary, scalar random processes and  $z$  be partial Markov- $m$ . Assume that  $[x_{k-m+1}^k z^k]$  has positive definite covariance at a given time  $k > m$ .

Suppose there exists some  $q \in (m, k]$  such that the matrix  $C^{x^Q \rightarrow z^Q}(m, q)$  (8), involving covariances of the high-resolution, uniformly quantized data, is full-rank. Then  $x$  causes  $z$ , provided that the quantizer interval length  $\Delta_z$  for signal  $z$  satisfies

$$\Delta_z < \left( \frac{12\sigma_{\min}^Q}{\sqrt{m+1}} \right)^{\frac{1}{2}}, \quad (23)$$

where  $\sigma_{\min}^Q$  is the smallest singular value of  $C^{x^Q \rightarrow z^Q}(m, q)$ .

*Remark 3.7:* This result gives an explicit formula in the high-resolution regime for deciding how finely quantized  $x$  and  $z$  should be in order to infer causality from  $C^{x^Q \rightarrow z^Q}(m, q)$ , which can be constructed through quantized signals available. As in the quantization perturbation bound

in Theorem 3.3, the RHS depends on the quantization scheme. Using (14), (19) and (21), we can establish a quantizer-independent RHS as follows:

$$\Delta_z < \left( \frac{6\sigma_{\min}}{\sqrt{m+1}} \right)^{\frac{1}{2}}, \quad (24)$$

where  $\sigma_{\min} := \sigma_{\min}(C_G^{x \rightarrow z}(m, q))$ .

Note the LHS depends on  $\Delta_z$ , suggesting that  $\Delta_z$  plays a more critical role than  $\Delta_x$ . Even if  $\Delta_x$  is coarse, the inequality can be satisfied by choosing a sufficiently fine  $\Delta_z$ . Conversely, if  $\Delta_x$  is small,  $\Delta_z$  would still have to be small to satisfy (24). This is related to the fact that in the causality matrix  $C_G^{x \rightarrow z}(m, q)$ ,  $x$  appears only in cross-covariances with  $z$ , whereas  $z$  also appears in auto-covariances with lag zero.

#### IV. CONCLUSIONS

Granger causality inference for jointly stationary Gaussian, partially Markov- $m$  signals using quantized data has been investigated in this paper. A rank criterion to assess the Granger causality between Gaussian signals was proposed. Conditions under which causality can be inferred by using just the statistical properties of quantized signals instead of estimating the statistics or model of the underlying Gaussian signals have been introduced.

Future work will focus on determining lower bounds and tighter upper bounds on the coarsest quantizer resolution that permits Granger causality to be inferred from quantized data, and also on relaxing the Gaussian assumption. In this case, it is likely that information-theoretic approaches based on *directed information* (see e.g. [24]) will prove useful.

#### V. APPENDIX

*Proof of Part 1 of Theorem 2.3:* The conditional Gaussian random vectors  $A|C, B$  and  $A|C, D$  have distributions as follows:

$$\begin{aligned} A|C, B &\sim \mathcal{N}(\mu_{\text{cond}}^{A|C, B}, \Gamma_{\text{cond}}^{A|C, B}), \\ A|C, D &\sim \mathcal{N}(\mu_{\text{cond}}^{A|C, D}, \Gamma_{\text{cond}}^{A|C, D}), \end{aligned}$$

where

$$\begin{aligned} \mu_{\text{cond}}^{A|C, B} &:= \mu_A + \Gamma_{A, [CB]} \Gamma_{[CB], [CB]}^{-1} \begin{bmatrix} C - \mu_C \\ B - \mu_B \end{bmatrix}, \\ \Gamma_{\text{cond}}^{A|C, B} &:= \Gamma_{A, A} - \Gamma_{A, [CB]} \Gamma_{[CB], [CB]}^{-1} \Gamma_{[CB], A}, \\ \mu_{\text{cond}}^{A|C, D} &:= \mu_A + \Gamma_{A, [CD]} \Gamma_{[CD], [CD]}^{-1} \begin{bmatrix} C - \mu_C \\ D - \mu_D \end{bmatrix}, \\ \Gamma_{\text{cond}}^{A|C, D} &:= \Gamma_{A, A} - \Gamma_{A, [CD]} \Gamma_{[CD], [CD]}^{-1} \Gamma_{[CD], A}, \end{aligned}$$

and  $\Gamma_{A, [CD]} := [\Gamma_{A, C} \quad \Gamma_{A, D}]$  and  $\Gamma_{[CD], A} := [\Gamma_{C, A}^T \quad \Gamma_{D, A}^T]^T$ .

In order for two Gaussian distributions to be identical, it is necessary and sufficient that they should have the same mean vector and covariance matrix. Let us begin with the equality of mean vectors ( $\mu_{\text{cond}}^{A|C, B} = \mu_{\text{cond}}^{A|C, D}$ ):

$$\begin{bmatrix} \Gamma_{A, C} & \Gamma_{A, B} \end{bmatrix} \begin{bmatrix} \Gamma_{C, C} & \Gamma_{C, B} \\ \Gamma_{B, C} & \Gamma_{B, B} \end{bmatrix}^{-1} \begin{bmatrix} C - \mu_C \\ B - \mu_B \end{bmatrix} =$$

$$\begin{bmatrix} \Gamma_{A, C} & \Gamma_{A, D} \end{bmatrix} \begin{bmatrix} \Gamma_{C, C} & \Gamma_{C, D} \\ \Gamma_{D, C} & \Gamma_{D, D} \end{bmatrix}^{-1} \begin{bmatrix} C - \mu_C \\ D - \mu_D \end{bmatrix}$$

As  $\Gamma_{[BCD], [BCD]}$  is positive definite, this linear relationship between  $B - \mu_B$ ,  $C - \mu_C$  and  $D - \mu_D$  implies that their three coefficients must be zero. Noting that the inverses above and of  $\Gamma_{B, B}$  and  $\Gamma_{D, D}$  also must exist, Problem 1.6.7 of [25] implies that  $\Gamma_{C, C} - \Gamma_{C, B} \Gamma_{B, B}^{-1} \Gamma_{B, C}$  and  $\Gamma_{C, C} - \Gamma_{C, D} \Gamma_{D, D}^{-1} \Gamma_{D, C}$  are also invertible. Thus the three zero-coefficient equations can be written in the following form:

$$\begin{aligned} \Gamma_{A, C}(K_2 - K_1) - \Gamma_{A, D} \Gamma_{D, D}^{-1} \Gamma_{D, C} K_2 \\ + \Gamma_{A, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} K_1 = \mathbf{0}, \end{aligned} \quad (25)$$

$$-\Gamma_{A, C} K_1 \Gamma_{C, B} + \Gamma_{A, B} + \Gamma_{A, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} K_1 \Gamma_{C, B} = \mathbf{0}, \quad (26)$$

and

$$-\Gamma_{A, C} K_2 \Gamma_{C, D} + \Gamma_{A, D} + \Gamma_{A, D} \Gamma_{D, D}^{-1} \Gamma_{D, C} K_2 \Gamma_{C, D} = \mathbf{0}, \quad (27)$$

where

$$\begin{aligned} K_1 &:= (\Gamma_{C, C} - \Gamma_{C, B} \Gamma_{B, B}^{-1} \Gamma_{B, C})^{-1}, \\ K_2 &:= (\Gamma_{C, C} - \Gamma_{C, D} \Gamma_{D, D}^{-1} \Gamma_{D, C})^{-1}. \end{aligned} \quad (28)$$

It can be shown that if (25), (26) and (27) are satisfied, the equality of the conditional covariance matrices  $\Gamma_{\text{cond}}^{A|C, B}$  and  $\Gamma_{\text{cond}}^{A|C, D}$  also holds. In other words, the conditional Gaussian probability density functions are equal ( $P(A|C, B) = P(A|C, D)$ ) if and only if (25), (26), and (27) are satisfied.

Substituting (25) in (27) can be used to express the three equations as the two following equations:

$$\Gamma_{A, B} - \Gamma_{A, C} K_1 \Gamma_{C, B} + \Gamma_{A, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} K_1 \Gamma_{C, B} = \mathbf{0}, \quad (29)$$

$$\Gamma_{A, D} - \Gamma_{A, C} K_1 \Gamma_{C, D} + \Gamma_{A, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} K_1 \Gamma_{C, D} = \mathbf{0}. \quad (30)$$

We know that a matrix  $A$  is zero if and only if  $\text{rank}(A) = 0$ . Using Guttman rank additivity formula [26] for (29) and (30), we have:

$$\begin{aligned} \text{rank}(\Phi) &= \text{rank}(\Gamma_{C, C} - \Gamma_{C, B} \Gamma_{B, B}^{-1} \Gamma_{B, C}) = \#C, \\ \text{rank}(\Upsilon) &= \text{rank}(\Gamma_{C, C} - \Gamma_{C, B} \Gamma_{B, B}^{-1} \Gamma_{B, C}) = \#C, \end{aligned}$$

where

$$\begin{aligned} \Phi &:= \begin{bmatrix} \Gamma_{C, C} - \Gamma_{C, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} & \Gamma_{C, B} \\ \Gamma_{A, C} - \Gamma_{A, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} & \Gamma_{A, B} \end{bmatrix}, \\ \Upsilon &:= \begin{bmatrix} \Gamma_{C, C} - \Gamma_{C, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} & \Gamma_{C, D} \\ \Gamma_{A, C} - \Gamma_{A, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} & \Gamma_{A, D} \end{bmatrix}. \end{aligned}$$

There are two partitioned matrices whose the first column is the same in the equations above. To deal with the ranks of these two matrices, we augment two matrices of our interest through Lemma 6 of [27] as follows:

$$\text{rank}(\Xi) \leq \text{rank} \left( \begin{bmatrix} \Gamma_{C, D} & \Gamma_{C, C} - \Gamma_{C, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} \\ \Gamma_{A, D} & \Gamma_{A, C} - \Gamma_{A, B} \Gamma_{B, B}^{-1} \Gamma_{B, C} \end{bmatrix} \right)$$

$$\begin{aligned}
& + \text{rank} \left( \begin{bmatrix} \Gamma_{C,C} - \Gamma_{C,B} \Gamma_{B,B}^{-1} \Gamma_{B,C} & \Gamma_{C,B} \\ \Gamma_{A,C} - \Gamma_{A,B} \Gamma_{B,B}^{-1} \Gamma_{B,C} & \Gamma_{A,B} \end{bmatrix} \right) \\
& - \text{rank} \left( \begin{bmatrix} \Gamma_{C,C} - \Gamma_{C,B} \Gamma_{B,B}^{-1} \Gamma_{B,C} \\ \Gamma_{A,C} - \Gamma_{A,B} \Gamma_{B,B}^{-1} \Gamma_{B,C} \end{bmatrix} \right) \\
& = \text{rank}(\Upsilon) + \text{rank}(\Phi) - \#C = \#C,
\end{aligned} \tag{31}$$

where

$$\Xi := \begin{bmatrix} \Gamma_{C,D} & \Gamma_{C,C} - \Gamma_{C,B} \Gamma_{B,B}^{-1} \Gamma_{B,C} & \Gamma_{C,B} \\ \Gamma_{A,D} & \Gamma_{A,C} - \Gamma_{A,B} \Gamma_{B,B}^{-1} \Gamma_{B,C} & \Gamma_{A,B} \end{bmatrix}. \tag{32}$$

We also have:

$$\begin{aligned}
\text{rank}(\Upsilon) & \geq \max \left\{ \text{rank} \left( \begin{bmatrix} \Gamma_{C,D} \\ \Gamma_{A,D} \end{bmatrix} \right), \right. \\
& \quad \left. \text{rank} \left( \begin{bmatrix} \Gamma_{C,C} - \Gamma_{C,B} \Gamma_{B,B}^{-1} \Gamma_{B,C} \\ \Gamma_{A,C} - \Gamma_{A,B} \Gamma_{B,B}^{-1} \Gamma_{B,C} \end{bmatrix} \right) \right\} \\
& \geq \text{rank} \left( \begin{bmatrix} \Gamma_{C,D} \\ \Gamma_{A,D} \end{bmatrix} \right)
\end{aligned} \tag{33}$$

where Theorem 2.9 of [28] is used to derive (33). Exploiting (33), we have:

$$\begin{aligned}
\text{rank}(\Xi) & \geq \max \left\{ \text{rank} \left( \begin{bmatrix} \Gamma_{C,D} \\ \Gamma_{A,D} \end{bmatrix} \right), \text{rank}(\Phi) \right\} \\
& = \max \left\{ \text{rank}(\Upsilon), \text{rank}(\Phi) \right\} \\
& = \#C.
\end{aligned} \tag{34}$$

Equations (31) and (34) yield:

$$\#C \leq \text{rank}(\Xi) \leq \#C. \tag{35}$$

Thus, we can establish:

$$\text{rank}(\Xi) = \#C. \tag{36}$$

We can conclude that if the conditional Gaussian probability density functions are equal ( $P(A|C, B) = P(A|C, D)$ ), (25), (26) and (27) are satisfied and in turn show that the rank criterion (36) holds. Therefore, one direction is proved.

Due to paper length restriction, proofs of the converse of part 1 and part 2 of Theorem 2.3 are not mentioned here. ■

## REFERENCES

- [1] C.W.J. Granger, Economic Process Involving Feedback, *Inf. Control*, vol. 6, 1963, pp 28-48.
- [2] N. Wiener, *The Theory of Prediction*, McGraw-Hill, NY; 1956.
- [3] C.W.J. Granger, Testing for Causality: a Personal Viewpoint, *J. Econ. Dyn. Control*, vol. 2, 1980, pp 329-352.
- [4] C.W.J. Granger, Some Recent Development in a Concept of Causality, *J. Econom.*, vol. 39, 1988, pp 199-211.
- [5] P. Caines, C. Chan, Feedback Between Stationary Stochastic Processes, *IEEE Trans. Autom. Contr.*, vol. 20, 1975, pp 498-508.
- [6] P. Caines, Weak and Strong Feedback Free Processes, *IEEE Trans. Autom. Contr.*, vol. 21, 1976, pp 737-739.
- [7] V. Solo, On causality I: Sampling and Noise, in *46th IEEE Conference on Decision and Control*, New Orleans, LA, 2007, pp 3634-3639.
- [8] V. Solo, State-Space Analysis of Granger-Geweke Causality Measures with Application to fMRI, *Neural Comput.*, vol. 26, 2016, pp 914-949.
- [9] B.D.O. Anderson, M. Deistler and J.-M. Dufour, On the Sensitivity of Granger Causality to Errors-In-Variables, Linear Transformations and Subsampling, *J. Time Ser. Anal.*, vol. 40, 2019, pp 102-123.
- [10] H. Nalatore, M. Ding, G. Rangarajan, Mitigating the Effects of Measurement Noise on Granger Causality, *Physical Review E*, vol. 75, 2007, pp 031123.
- [11] H. Nalatore, N. Sasikumar, G. Rangarajan, Effect of Measurement Noise on Granger Causality, *Physical Review E*, vol. 90, 2014, pp 062127.
- [12] S. Sullivant, Algebraic Geometry of Gaussian Bayesian Networks, *Adv. in Appl. Math.*, vol. 40, 2008, pp 482-513.
- [13] S. Sullivant, *Algebraic Statistics*, American Mathematical Society, Rhode Island, USA; 2018.
- [14] M. Laghate and D. Cabric, Learning Wireless Networks Topologies Using Asymmetric Granger Causality, *IEEE J. Sel. Topics Signal Process.*, vol. 12, 2018, pp 233 - 247.
- [15] C.A. Sims, Money, Income and Causality, *Am. Econ. Rev.*, vol. 62, 1972, pp 540-552.
- [16] J.F. Geweke, Measures of Conditional Linear Dependence and Feedback Between Time Series, *J. Am. Stat. Assoc.*, vol. 79, 1984, pp 907 - 915.
- [17] I. Kontoyiannis and M. Skoularidou, Estimating the Directed Information and Testing for Causality, *IEEE Trans. Inform. Theory*, vol. 62, 2016, pp 6053-6067.
- [18] C.J. Quinn, N. Kiyavash, and T.P. Coleman, Directed Information Graphs, *IEEE Trans. Inform. Theory*, vol. 61, 2015, pp 6887-6909.
- [19] C.J. Quinn, T.P. Coleman, N. Kiyavash, and N.G. Hatsopoulos, Estimating the Directed Information to Infer Causal Relationships in Ensemble Neural Spike Train Recordings, *J. Comput. Neurosci.*, vol. 30, 2011, pp 17-44.
- [20] C. Eckart and G. Young, The Approximation of One Matrix by Another of Lower Rank, *Psychometrika*, vol. 1, 1936, pp 211-218.
- [21] L. Mirsky, Symmetric Gauge Functions and Unitarily Invariant Norms, *Quart. J. Math. Oxford*, vol. 11, 1960, pp 50-59.
- [22] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK; 2013.
- [23] B. Widrow, I. Kollar, *Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*, Cambridge University Press, Cambridge, UK; 2008.
- [24] P.O. Amblard, O.J.J. Michel, The Relation Between Granger Causality and Directed Information Theory: A Review, *Entropy*, vol. 15, 2013, pp 113-143.
- [25] B. Noble, J.W. Daniel, *Applied Linear Algebra*, Prentice Hall, Englewood Cliffs, NJ; 1988.
- [26] L. Guttman, Enlargement Methods for Computing the Inverse Matrix, *Ann. Math. Statist.*, vol. 17, 1946, pp 336-343.
- [27] M. Lundquist, W. Barrett, Rank Inequalities for Positive Semidefinite Matrices, *Linear Algebra Appl.*, vol. 248, 1996, pp 91-100.
- [28] J.R. Schott, *Matrix Analysis for Statistics*, John Wiley & Sons, New York, NY; 2017.

Minerva Access is the Institutional Repository of The University of Melbourne

**Author/s:**

Ahmadi, S;Nair, GN;Weyer, E

**Title:**

Granger Causality of Gaussian Signals from Quantized Measurements

**Date:**

2020-03-12

**Citation:**

Ahmadi, S., Nair, G. N. & Weyer, E. (2020). Granger Causality of Gaussian Signals from Quantized Measurements. Proceedings of the 2019 IEEE 58th Conference on Decision and Control (CDC), 2019-December, pp.3587-3592. IEEE. <https://doi.org/10.1109/CDC40024.2019.9030255>.

**Persistent Link:**

<http://hdl.handle.net/11343/241850>