

EXACT SIMULATION OF NONCIRCULAR OR IMPROPER COMPLEX-VALUED STATIONARY GAUSSIAN PROCESSES USING CIRCULANT EMBEDDING

Adam M. Sykulski

Department of Statistical Science,
University College London, UK

Donald B. Percival

Applied Physics Laboratory,
University of Washington, Seattle, USA

ABSTRACT

This paper provides an algorithm for simulating improper (or noncircular) complex-valued stationary Gaussian processes. The technique utilizes recently developed methods for multivariate Gaussian processes from the circulant embedding literature. The method can be performed in $\mathcal{O}(n \log_2 n)$ operations, where n is the length of the desired sequence. The method is exact, except when eigenvalues of prescribed circulant matrices are negative. We evaluate the performance of the algorithm empirically, and provide a practical example where the method is guaranteed to be exact for all n , with an improper fractional Gaussian noise process.

Index Terms— Circulant embedding, improper, noncircular, complex-valued, fractional Gaussian noise

1. INTRODUCTION

Bivariate time series are commonly represented as complex-valued signals in numerous applications including wireless communication [1], meteorology and oceanography [2], and functional Magnetic Resonance Imaging (fMRI) [3]. There have been important methodological developments in machine learning and signal processing for complex-valued signals, see e.g. the books by [4, 5]. There is a growing interest in analyzing complex-valued processes that are said to be *improper* (see e.g. [6, 7, 8]), as opposed to *proper*, concepts which we define formally in this paper. Improper processes are sometimes referred to as noncircular processes, and proper processes as circular processes [4, 9]. A proper process can be interpreted as a second order process whose statistics are rotationally invariant when viewed in the complex plane, whereas improper processes are not, and the latter are seen to be more realistic models for applications [5, Section 1.9]. Numerous improper complex-valued stochastic models have been proposed, e.g. [6, 10, 11, 12, 13], for applications including climate and seismology, and in this paper we provide an algorithm for their efficient simulation.

For efficiently simulating Gaussian processes, the technique of circulant embedding [14, 15] is receiving increasing attention. The underlying concept is that the covariance matrix of any regularly sampled stationary process is Toeplitz, and hence can be embedded into a larger matrix that is *circulant*. A circulant matrix has the useful property that the matrix can be diagonalized via Fourier Transforms, such that Gaussian sequences can be simulated in $\mathcal{O}(n \log n)$ operations. This is faster than the alternative method of Cholesky factorization which requires $\mathcal{O}(n^2)$ operations, where n is the length of the desired sequence. The circulant embedding technique is exact, in that the simulated sequence has statistical properties that exactly match the model, there are however instances where exactness is not guaranteed, which has been the subject of further investigation, see e.g. [16, 17].

The method of circulant embedding was extended to simulating complex-valued processes in [18]. This procedure is only applicable to proper processes however. In this paper we extend the method to improper processes by incorporating techniques developed in [19] for multivariate processes. This connection was also recently mentioned in [20], although this paper is primarily focused on revisiting the problem of simulating proper processes. The contribution of this paper is to provide full details of the $\mathcal{O}(n \log_2 n)$ generating algorithm for improper processes, together with a practical example, and an empirical evaluation of its performance.

For improper processes modeled in the *frequency* domain, an exact simulation algorithm is provided in [21]. Here in this paper we detail an exact simulation algorithm for improper processes specified in the *time* domain by their covariance structure. The distinction is important, because processes specified analytically in one domain cannot in general be expressed analytically in the other. Therefore converting from one domain to the other often requires an approximation, thus losing the exactness of the procedure (see also [19, Section 3.4]). We also note approximate techniques for simulating improper processes in [22] using widely-linear filters [23]. Circulant embedding, on the other hand, can be used to exactly generate finite sequences from infinite-order non-Markovian Gaussian processes, as we shall demonstrate in this paper.

The work of A. M. Sykulski was supported by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme.

2. PRELIMINARIES

Consider a complex-valued stationary Gaussian process defined by $z_t = x_t + iy_t$, where $t \in \mathbb{Z}$. We assume without loss of generality, that z_t is zero mean, as a non-zero mean can be added to the simulated sequence *a posteriori*.

The process z_t is fully specified by the autocovariances and cross covariances of x_t and y_t respectively, given by

$$\begin{aligned} s_{xx}(\tau) &= \mathbb{E}\{x_{t+\tau}x_t\} \\ s_{yy}(\tau) &= \mathbb{E}\{y_{t+\tau}y_t\} \\ s_{xy}(\tau) &= \mathbb{E}\{x_{t+\tau}y_t\} \end{aligned} \quad (1)$$

for all $\tau \in \mathbb{Z}$, where $\mathbb{E}\{\cdot\}$ is the expectation operator. We note that $s_{xx}(\tau) = s_{xx}(-\tau)$ and $s_{yy}(\tau) = s_{yy}(-\tau)$, but in general $s_{xy}(\tau) \neq s_{xy}(-\tau)$. Equivalently, the process z_t can be fully specified by its own autocovariance and *complementary* covariance sequence, specified respectively by

$$\begin{aligned} s_{zz}(\tau) &= \mathbb{E}\{z_{t+\tau}z_t^*\} \\ r_{zz}(\tau) &= \mathbb{E}\{z_{t+\tau}z_t\} \end{aligned} \quad (2)$$

where z_t^* denotes the complex conjugate of z_t . The sequences $s_{zz}(\tau)$ and $r_{zz}(\tau)$ are in general complex-valued and have the properties that $s_{zz}(\tau) = s_{zz}^*(-\tau)$ (Hermitian symmetry) and $r_{zz}(\tau) = r_{zz}(-\tau)$ (symmetry). The sequence $r_{zz}(\tau)$ is sometimes also referred to as the relation sequence [8], or the pseudo-covariance [20]. Complex-valued processes that have the property that $r_{zz}(\tau) = 0$ for all τ are said to be *proper*, and are otherwise said to be *improper* [5]. In this paper we focus on the improper case where $r_{zz}(\tau) \neq 0$.

The easiest way to relate the sequences in (1) and (2) is by starting in the complex representation and then expanding, such that

$$\begin{aligned} s_{zz}(\tau) &= \mathbb{E}\{(x_{t+\tau} + iy_{t+\tau})(x_t - iy_t)\} \\ &= \mathbb{E}\{x_{t+\tau}x_t\} + \mathbb{E}\{y_{t+\tau}y_t\} \\ &\quad + i(\mathbb{E}\{x_t y_{t+\tau}\} - \mathbb{E}\{x_{t+\tau}y_t\}) \\ &= s_{xx}(\tau) + s_{yy}(\tau) + i[s_{xy}(-\tau) - s_{xy}(\tau)] \end{aligned} \quad (3)$$

$$\begin{aligned} r_{zz}(\tau) &= \mathbb{E}\{(x_{t+\tau} + iy_{t+\tau})(x_t + iy_t)\} \\ &= \mathbb{E}\{x_{t+\tau}x_t\} - \mathbb{E}\{y_{t+\tau}y_t\} \\ &\quad + i(\mathbb{E}\{x_t y_{t+\tau}\} + \mathbb{E}\{x_{t+\tau}y_t\}) \\ &= s_{xx}(\tau) - s_{yy}(\tau) + i[s_{xy}(-\tau) + s_{xy}(\tau)] \end{aligned} \quad (4)$$

where for the complex-valued process to be proper we therefore require $s_{xx}(\tau) = s_{yy}(\tau)$ and $s_{xy}(\tau) = -s_{xy}(-\tau)$.

3. THE METHOD

We provide the pseudo-code for the simulation algorithm of complex-valued improper processes in Algorithm 1. Then in Sections 3.1–3.3 we detail each line of the algorithm, followed by discussions in Section 3.4.

3.1. Converting from complex to bivariate

To generate a complex-valued sequence of length n of a zero-mean Gaussian process we first specify $s_{zz}(\tau)$ and $r_{zz}(\tau)$ for $\tau = 0, \dots, n$ (in Algorithm 1, line 1). This fully specifies the properties of the generated sequence as negative lags up to $-n$ are obtained through the properties $s_{zz}(\tau) = s_{zz}^*(-\tau)$ and $r_{zz}(\tau) = r_{zz}(-\tau)$, and lags larger than $\pm n$ are not required. The next step of our approach (in Algorithm 1, lines 2–5) is to specify the complex-valued process z_t in terms of the autocovariances and cross-covariances of x_t and y_t , which can be performed by inverting the relationships in (3) and (4) yielding

$$s_{xx}(\tau) = \Re\left\{\frac{1}{2}[s_{zz}(\tau) + r_{zz}(\tau)]\right\} \quad (5)$$

$$s_{yy}(\tau) = \Re\left\{\frac{1}{2}[s_{zz}(\tau) - r_{zz}(\tau)]\right\} \quad (6)$$

$$s_{xy}(\tau) = \Im\left\{\frac{1}{2}[r_{zz}(\tau) - s_{zz}(\tau)]\right\} \quad (7)$$

$$s_{xy}(-\tau) = \Im\left\{\frac{1}{2}[s_{zz}(\tau) + r_{zz}(\tau)]\right\}, \quad (8)$$

for $\tau = 0, \dots, n$, where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary part respectively. Note that $s_{xx}(\tau)$ and $s_{yy}(\tau)$ are defined for $\tau = -n, \dots, -1$ by symmetry, i.e. $s_{xx}(\tau) = s_{xx}(-\tau)$ and $s_{yy}(\tau) = s_{yy}(-\tau)$.

3.2. Generating the bivariate series

Having specified $s_{xx}(\tau)$, $s_{yy}(\tau)$, and $s_{xy}(\tau)$, for $\tau = -n, \dots, 0, \dots, n$, we proceed to use the method of [19] for multivariate Gaussian time series, setting the dimension parameter to $p = 2$ for bivariate series. We detail this within Algorithm 1, where we have simplified the algorithm provided in [19] for multivariate processes to bivariate processes.

In Algorithm 1, lines 6–8, the length- $2n$ vectors \mathbf{c}_{xx} , \mathbf{c}_{xy} , and \mathbf{c}_{yy} , are the top rows of the embedded circulant matrices, where the full matrices do not need to be formulated (thus reducing storage requirements from $4n^2$ to $2n$). The Fast Fourier Transform (FFT) of these vectors diagonalizes the circulant matrices, yielding the respective sequences of eigenvalues λ_{xx} , λ_{yy} , and λ_{xy} , which are computed in line 9 of Algorithm 1 via the operation

$$\lambda_{xx}(k) = \sum_{j=0}^{2n-1} c_{xx}(k) \exp\{-i2\pi j(k-1)/2n\}, \quad (9)$$

for $k = 1, \dots, 2n$, and similarly for $\lambda_{yy}(k)$ and $\lambda_{xy}(k)$. Note that $\lambda_{xx}(k)$ and $\lambda_{yy}(k)$ are real-valued, whereas $\lambda_{xy}(k)$ is in general complex-valued. As each FFT is of length $2n$, the ‘‘powers of two’’ FFT algorithm can be used when n is set to be a power of two.

Algorithm 1 Pseudo-code for the simulation algorithm for improper processes

```

1: INPUTS:  $s_{zz}(\tau), r_{zz}(\tau)$ , for  $\tau = 0, \dots, n$ 
2: for  $\tau = 0$  to  $n$  do
3:    $s_{xx}(\tau) \leftarrow \Re \left\{ \frac{1}{2} [s_{zz}(\tau) + r_{zz}(\tau)] \right\}$ ;    $s_{yy}(\tau) \leftarrow \Re \left\{ \frac{1}{2} [s_{zz}(\tau) - r_{zz}(\tau)] \right\}$ 
4:    $s_{xy}(\tau) \leftarrow \Im \left\{ \frac{1}{2} [r_{zz}(\tau) - s_{zz}(\tau)] \right\}$ ;    $s_{xy}(-\tau) \leftarrow \Im \left\{ \frac{1}{2} [s_{zz}(\tau) + r_{zz}(\tau)] \right\}$ 
5: end for
6:  $\mathbf{c}_{xx} \leftarrow [s_{xx}(0), s_{xx}(1), \dots, s_{xx}(n-1), s_{xx}(n), s_{xx}(n-1), \dots, s_{xx}(1)]$ 
7:  $\mathbf{c}_{yy} \leftarrow [s_{yy}(0), s_{yy}(1), \dots, s_{yy}(n-1), s_{yy}(n), s_{yy}(n-1), \dots, s_{yy}(1)]$ 
8:  $\mathbf{c}_{xy} \leftarrow [s_{xy}(0), s_{xy}(-1), \dots, s_{xy}(-(n-1)), s_{xy}(-n), s_{xy}(n-1), \dots, s_{xy}(1)]$ 
9:  $\lambda_{xx} \leftarrow \text{FFT}(\mathbf{c}_{xx})$ ;    $\lambda_{yy} \leftarrow \text{FFT}(\mathbf{c}_{yy})$ ;    $\lambda_{xy} \leftarrow \text{FFT}(\mathbf{c}_{xy})$ 
10: Generate 4 length- $2n$  i.i.d. Gaussian sequences with mean 0 and variance 1/2, denote as:  $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ 
11: for  $k = 1$  to  $2n$  do
12:    $\Sigma_k \leftarrow 2 \begin{pmatrix} 1 & \lambda_{xy}(k) / \sqrt{\lambda_{xx}(k)\lambda_{yy}(k)} \\ \lambda_{xy}^*(k) / \sqrt{\lambda_{xx}(k)\lambda_{yy}(k)} & 1 \end{pmatrix}$ ;    $A_k \leftarrow \text{chol}(\Sigma_k)$ 
13:    $\begin{bmatrix} g_x(k) \\ g_y(k) \end{bmatrix} \leftarrow A_k \begin{bmatrix} w_1(k) + iw_2(k) \\ w_3(k) + iw_4(k) \end{bmatrix}$ ;    $\begin{bmatrix} h_x(k) \\ h_y(k) \end{bmatrix} \leftarrow \frac{1}{\sqrt{2n}} \begin{bmatrix} \lambda_{xx}(k)g_x(k) \\ \lambda_{yy}(k)g_y(k) \end{bmatrix}$ 
14: end for
15:  $\mathbf{q}_x \leftarrow \text{FFT}(\mathbf{h}_x)$ ;    $\mathbf{q}_y \leftarrow \text{FFT}(\mathbf{h}_y)$ 
16:  $\mathbf{z}_1 \leftarrow [\Re\{q_x(1)\}, \dots, \Re\{q_x(n)\}] + i[\Re\{q_y(1)\}, \dots, \Re\{q_y(n)\}]$ 
17:  $\mathbf{z}_2 \leftarrow [\Im\{q_x(1)\}, \dots, \Im\{q_x(n)\}] + i[\Im\{q_y(1)\}, \dots, \Im\{q_y(n)\}]$ 
18: RETURN:  $\mathbf{z}_1, \mathbf{z}_2$ 

```

The next step (Algorithm 1, line 10) is to generate four length- $2n$ i.i.d Gaussian sequences with mean 0 and variance 0.5, which we denote as $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$, and \mathbf{w}_4 . Then a length- $2n$ for loop is performed (Algorithm 1, line 11–14) to correlate the random variables correctly. First, in line 12, the 2×2 matrix Σ_k is computed using the eigenvalue sequences. A remark made in [19] is that if $\lambda_{xx}(k) = 0$ or $\lambda_{yy}(k) = 0$, then the off-diagonals in Σ_k should be set to 0, and this does not lose the exactness of the procedure. The next step is to compute the Cholesky decomposition, denoted as $A_k = \text{chol}(\Sigma_k)$. Then in Line 13, the matrix A_k is used to transform the random Gaussian sequences, $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$, and \mathbf{w}_4 , into two correlated complex-valued length- $2n$ vectors \mathbf{g}_x and \mathbf{g}_y . Finally, these are normalized by the eigenvalue sequences, λ_{xx} and λ_{yy} , yielding the two vectors \mathbf{h}_x and \mathbf{h}_y .

The final step (Algorithm 1, line 15) is to transform back from the frequency domain to the time domain, by performing the FFT (as in (9)) on \mathbf{h}_x and \mathbf{h}_y . This yields two length- $2n$ vectors \mathbf{q}_x and \mathbf{q}_y .

3.3. Generating the complex-valued series

To generate a complex-valued sequence of length n , with prescribed autocovariance $s_{zz}(\tau)$ and complementary covariance $r_{zz}(\tau)$, we take the real parts of the first n values of \mathbf{q}_x

and \mathbf{q}_y , to form a complex-valued sequence \mathbf{z}_1 , as given in Algorithm 1, line 16. In addition, we can recover a second sequence \mathbf{z}_2 by taking the first n values of the *imaginary* parts of \mathbf{q}_x and \mathbf{q}_y (Algorithm 1, line 17). Therefore to generate M length n sequences with the same $s_{zz}(\tau)$ and $r_{zz}(\tau)$, we only need to run the algorithm $\lceil M/2 \rceil$ times.

3.4. Discussion

The exactness of the approach of Algorithm 1 follows directly from the exactness of the method proposed in [19], together with the derivation of (3) and (4). Algorithm 1 is executable in $\mathcal{O}(n \log_2 n)$ operations, as all operations can be computed in $\mathcal{O}(n)$, except for the 5 Fourier transforms in Lines 9 and 15, which can be each computed in $\mathcal{O}(n \log_2 n)$ using the standard “powers of two” FFT algorithm.

As discussed in [18, 19] and elsewhere, a potentially problematic aspect of implementing circulant embedding algorithms is that the eigenvalues λ_{xx} and λ_{yy} (in Algorithm 1, line 9) are not guaranteed to be nonnegative definite. Negative eigenvalues occur when the circulant matrices (in Algorithm 1, line 6–7) are not nonnegative definite, and are hence invalid covariance matrices. Two ways of dealing with negative eigenvalues are commonly suggested in the literature. The first is to over-sample and implement Algorithm 1 with $m > n$, where certain values of m may yield entirely

nonnegative eigenvalues. The second is to set negative eigenvalues to zero, which renders the methods inexact.

Note that λ_{xy} is complex-valued and nonnegative eigenvalues are not required here for the exactness of the procedure. We do however require that the matrix Σ_k is nonnegative definite for the Cholesky decomposition to be valid. Here, [19] suggest approximating Σ_k with a nonnegative definite matrix found by setting negative eigenvalues of Σ_k to zero.

In [19], the authors prove two scenarios in which the method is guaranteed to be exact, which we briefly discuss here. First we define the 2×2 covariance matrix

$$R_\tau = \begin{pmatrix} s_{xx}(\tau) & s_{xy}(\tau) \\ s_{xy}(-\tau) & s_{yy}(\tau) \end{pmatrix}$$

for all $\tau \in \mathbb{Z}$. The first scenario is when the matrices R_τ are absolutely summable, and the corresponding spectral density matrix is positive definite. Then the method is exact for large enough m . The second scenario is in time-reversible cases, which in the bivariate case simply means that $s_{xy}(\tau) = s_{xy}(-\tau)$, then the method is exact for any $m \geq n$ if the matrices R_τ , $\Delta R_\tau = R_\tau - R_{\tau+1}$, and $\Delta^2 R_\tau = R_\tau - 2R_{\tau+1} + R_{\tau+2}$, are nonnegative definite for $\tau \geq 0$.

At first glance these conditions appear hard to interpret, but intuition can be developed by looking in the complex domain. By examining (5)–(8) we can see that time-reversibility follows when $\Im\{s_{zz}(\tau)\} = 0$ for all τ . Then if we also have that $\Im\{r_{zz}(\tau)\} = 0$, it follows that R_τ is diagonal, with entries equal to $s_{zz}(\tau) + r_{zz}(\tau)$ and $s_{zz}(\tau) - r_{zz}(\tau)$. Thus if we meet the three conditions that

$$|r_{zz}(\tau)| \leq s_{zz}(\tau) \quad (10)$$

$$|r_{zz}(\tau) - r_{zz}(\tau + 1)| \leq s_{zz}(\tau) - s_{zz}(\tau + 1) \quad (11)$$

$$|r_{zz}(\tau) - 2r_{zz}(\tau + 1) + r_{zz}(\tau + 2)| \leq s_{zz}(\tau) - 2s_{zz}(\tau + 1) + s_{zz}(\tau + 2) \quad (12)$$

for $\tau \geq 0$, then the simulation method is exact. We will revisit these conditions in our applications section.

4. APPLICATION TO IMPROPER FRACTIONAL GAUSSIAN NOISE

In this section we test our algorithm on an improper fractional Gaussian noise (fGn) process. The complex-valued fGn process [24] is a stationary Gaussian process which, when regularly sampled at timesteps $t \in \mathbb{Z}$, has autocovariance

$$s_{zz}(\tau) = \frac{V_H}{2} A^2 (|\tau + 1|^{2H} + |\tau - 1|^{2H} - 2|\tau|^{2H}) \quad (13)$$

where $0 < H < 1$ is known as the Hurst parameter and controls the roughness or smoothness of the process, with a higher value leading to a smoother process. The normalizing parameter V_H is set as

$$V_H = \frac{\Gamma(H)\Gamma(1-H)}{\pi\Gamma(2H+1)} \quad (14)$$

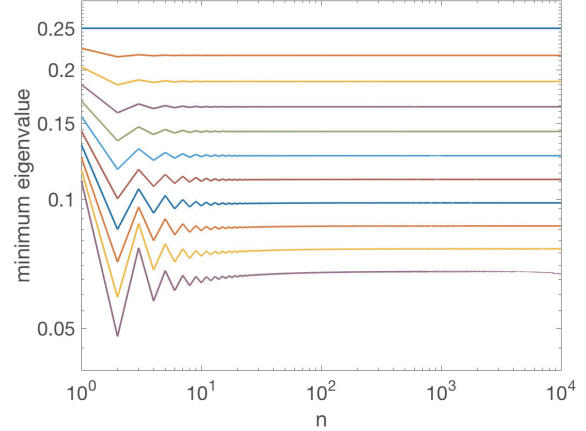


Fig. 1. The minimum eigenvalue for $\{\lambda_{xx}, \lambda_{yy}\}$ obtained over different values of n in line 9 of Algorithm 1 for the models of (13) and (15) where $A = 1$ and $B = 1/\sqrt{2}$. Each line represents a different value of the Hurst parameter H , and corresponding normalizing constant V_H in (14), where the top line is $H = 0.5$ and in each subsequent line H increases in increments of 0.05. The bottom line is $H = 0.9999$, as the process is not defined for $H = 1$, only in the limit as $H \rightarrow 1$.

which is required to regulate the behavior of fractional Brownian motion (fBm) [25], where fGn is defined as the increment process of a regularly sampled fBm.

Setting $H = 0.5$ recovers a regular Gaussian white noise process. When $H > 0.5$, then $s_{zz}(\tau) > 0$ and the process is said to be *persistent*, as it is positively correlated in time. Conversely, if $H < 0.5$ then $s_{zz}(\tau) < 0$ and the process is *anti-persistent* and negatively correlated in time. fGn with $H > 0.5$ is a commonly used long memory model in numerous applications, see e.g. [26].

For our improper process, we propose to model the complementary covariance to follow the same form of $s_{zz}(\tau)$ such that

$$r_{zz}(\tau) = \frac{V_H}{2} B^2 (|\tau + 1|^{2H} + |\tau - 1|^{2H} - 2|\tau|^{2H}) \quad (15)$$

where $B^2 < A^2$. It is easy to show that this simple model is a valid process by either transforming into the frequency domain and verifying that the spectral matrix is nonnegative definite [27], or by transforming to a bivariate representation using (5)–(8).

It also follows that the conditions in (10)–(12) are satisfied by (13) and (15) when $H \geq 0.5$. This can be verified by first observing that $s_{zz}(\tau)$ is positive, decreasing for $\tau \geq 0$, and convex, such that the right hand side of (10)–(12) is positive for all $\tau \geq 0$ in each of the three cases. Then as $r_{zz}(\tau)$ is a re-scaled version of $s_{zz}(\tau)$, the terms on the left hand side are always a factor of B^2/A^2 smaller than the right hand side, and the three inequalities are hence satisfied. Finally, as $\Im\{s_{zz}(\tau)\} = 0$ and $\Im\{r_{zz}(\tau)\} = 0$, it follows from the proof of [19] that the circulant embedding procedure of Algo-

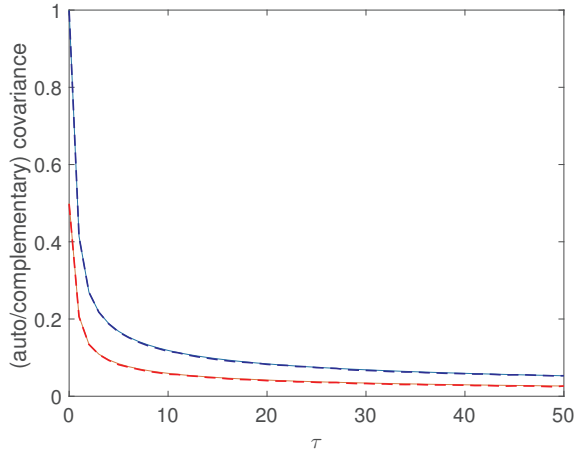


Fig. 2. The thin solid lines are the autocovariance and cross covariance models of (13) (in blue) and (15) (in red) respectively, for the first 50 lags. We have set the parameters to $H = 0.75$, $A = 1/\sqrt{V_H}$, and $B = A/\sqrt{2}$. Overlaid (in thick dashed lines) are the averaged unbiased auto and complementary covariance estimates for 1000 time series of length $n = 1000$, simulated using Algorithm 1.

rithm 1 is guaranteed to be exact for all n with this process. As an additional verification, in Fig. 1 we plot the minimum eigenvalue for $\{\lambda_{xx}, \lambda_{yy}\}$ obtained in Line 9 of Algorithm 1 over different values of $H \geq 0.5$ and n (with $A = 1$ and $B = 1/\sqrt{2}$), and they are seen to be positive everywhere.

We now simulate regularly sampled discrete sequences from the model of (13) and (15). As the process is not Markovian to any finite order, and hence cannot be represented as an autoregressive process of finite order, then we cannot recursively simulate the sequences in $\mathcal{O}(n)$ operations, and as such we turn to our circulant embedding algorithm. We first simulate 1000 time series each of length $n = 1000$. We fix the parameters in (13) and (15) to $H = 0.75$, $A = 1/\sqrt{V_H}$, and $B = A/\sqrt{2}$, such that the variance is normalized to 1. We display plots of $s_{zz}(\tau)$ and $r_{zz}(\tau)$ in Fig. 2 along with the average unbiased estimators of the autocovariance and complementary covariance for the 1000 simulated time series. The exactness of the method appears to hold, however to test this more robustly we repeat the experiment over values of n ranging from 10 to 1000 (in increments of 10). For each value of n , we simulated 1000 time series and report the root mean square difference between $s_{zz}(\tau)$ and its averaged unbiased estimate $\hat{s}_{zz}(\tau; n)$ (as performed in [18]) given by

$$RMS(s_{zz}(\tau); n) = \sqrt{\frac{1}{n} \sum_{\tau=0}^{n-1} |\hat{s}_{zz}(\tau; n) - s_{zz}(\tau)|^2}. \quad (16)$$

We similarly compute $RMS(r_{zz}(\tau); n)$ with each value of n . These metrics are displayed in Fig. 3, and are seen to decay as n increases. Even for small n however, the root mean square distance is less than 0.02, supporting the statement that for this process the method is an exact procedure for all n .

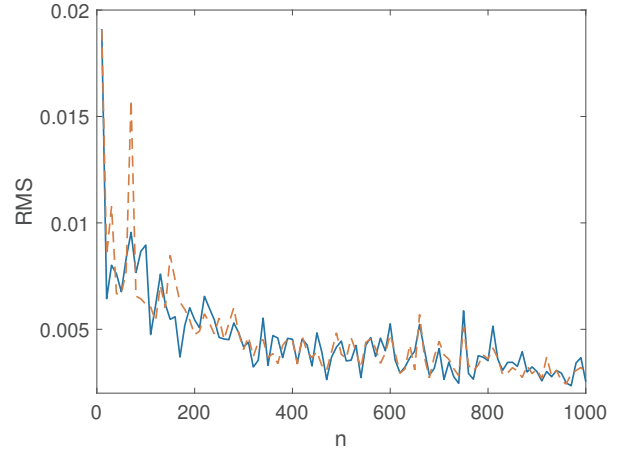


Fig. 3. The blue-solid line is the root mean square difference of the theoretical and observed autocovariance, calculated using (16), averaging over 1000 time series simulated using Algorithm 1, with the same parameter values as in Fig. 2. The analysis is repeated over values of n ranging from 10 to 1000 in increments of 10. The red-dashed line is the same analysis, this time reporting the root mean square difference of the theoretical and observed complementary covariance.

5. CONCLUSIONS

Improper or noncircular Gaussian processes have been receiving increasing attention in signal processing [28]. In this paper, an algorithm for generating improper complex-valued sequences has been proposed for stationary Gaussian processes specified in the time domain. This builds on previous methods for proper processes [18, 20], and with improper processes modeled in the frequency domain [21]. Our method is a reformulation of [19], where we have simplified the approach to bivariate processes, and then converted into the complex domain. We discussed how theoretical conditions under which exactness is guaranteed extend to the complex representation. An example of such a model was given, by proposing a simple improper complex fractional Gaussian noise process, and simulation evidence supported the exactness of the algorithm.

Circulant embedding simulation methods are only exact when computed eigenvalues from circulant matrices are non-negative. As reported in [29] for spatial processes, the issue of negative eigenvalues appears to grow going from one to two dimensions (and indeed higher). As an example, we implemented the improper periodic covariance kernel for complex-valued time series used in [12] for climate modeling, and found negative eigenvalues were unavoidable with any n . Setting these to zero, the root mean square metric of (16) was found to be much higher, at around 0.05 for $n = 1000$. Therefore, pending more use in practice, the limitations of the simulation algorithm remain unclear. Important avenues of future investigation are to find more general conditions under which exactness does and does not hold, as well as to see if the error bounds of approximate methods can be controlled theoretically when exact $\mathcal{O}(n \log n)$ simulation is not achievable.

ACKNOWLEDGMENTS

The authors would like to thank Dr Swati Chandna for her insightful comments and expertise in this area.

6. REFERENCES

- [1] K. W. Martin, "Complex signal processing is not complex," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 9, pp. 1823–1836, 2004.
- [2] J. Gonella, "A rotary-component method for analysing meteorological and oceanographic vector time series," *Deep-Sea Res.*, vol. 19, no. 12, pp. 833–846, 1972.
- [3] D. B. Rowe, "Modeling both the magnitude and phase of complex-valued fMRI data," *NeuroImage*, vol. 25, no. 4, pp. 1310–1324, 2005.
- [4] D. P. Mandic and V. S. L. Goh, *Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models*, John Wiley & Sons, 2009.
- [5] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals*, Cambridge University Press, 2010.
- [6] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, 1997.
- [7] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, 2003.
- [8] P. Rubin-Delanchy and A. T. Walden, "Kinematics of complex-valued time series," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4189–4198, 2008.
- [9] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Process. Lett.*, vol. 15, pp. 841–844, 2008.
- [10] J. Navarro-Moreno, M. D. Estudillo-Martínez, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "Estimation of improper complex-valued random signals in colored noise by using the hilbert space theory," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2859–2867, 2009.
- [11] A. Mohammadi and K. N. Plataniotis, "Improper complex-valued multiple-model adaptive estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1528–1542, 2015.
- [12] F. Tobar and R. E. Turner, "Modelling of complex signals using Gaussian processes," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, pp. 2209–2213, 2015.
- [13] A. M. Sykulski, S. C. Olhede, and J. M. Lilly, "A widely linear improper complex autoregressive process of order one," *IEEE Trans. Signal Process.*, accepted, 2016.
- [14] R. B. Davies and D. S. Harte, "Tests for Hurst effect," *Biometrika*, vol. 74, no. 1, pp. 95–101, 1987.
- [15] A. T. A. Wood and G. Chan, "Simulation of stationary Gaussian processes in $[0, 1]^d$," *Journal of Comput. Graph. Stat.*, vol. 3, no. 4, pp. 409–432, 1994.
- [16] C. R. Dietrich and G. N. Newsam, "Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix," *SIAM J Sci. Comput.*, vol. 18, no. 4, pp. 1088–1107, 1997.
- [17] P. F. Craigmile, "Simulating a class of stationary Gaussian processes using the Davies-Harte algorithm, with application to long memory processes," *Journal of Time Ser. Anal.*, vol. 24, no. 5, pp. 505–511, 2003.
- [18] D. B. Percival, "Exact simulation of complex-valued Gaussian stationary processes via circulant embedding," *Signal Process.*, vol. 86, no. 7, pp. 1470–1476, 2006.
- [19] H. Helgason, V. Pipiras, and P. Abry, "Fast and exact synthesis of stationary multivariate Gaussian time series using circulant embedding," *Signal Process.*, vol. 91, no. 5, pp. 1123–1133, 2011.
- [20] J. F. Couerjolly and E. Porcu, "Fast and exact simulation of complex-valued stationary Gaussian processes through embedding circulant matrix," *arXiv preprint:1604.00362*, 2016.
- [21] S. Chandna and A. T. Walden, "Simulation methodology for inference on physical parameters of complex vector-valued signals," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5260–5269, 2013.
- [22] P. Rubin-Delanchy and A. T. Walden, "Simulation of improper complex-valued sequences," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5517–5521, 2007.
- [23] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, 1995.
- [24] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM rev.*, vol. 10, no. 4, pp. 422–437, 1968.
- [25] J. M. Lilly, A. M. Sykulski, J. J. Early, and S. C. Olhede, "Fractional Brownian motion, the Matérn process, and stochastic modeling of turbulent dispersion," *arXiv preprint:1605.01684*, 2016.
- [26] V. Paxson, "Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic," *ACM SIGCOMM Comp. Comm. R.*, vol. 27, no. 5, pp. 5–18, 1997.
- [27] A. M. Sykulski, S. C. Olhede, J. M. Lilly, and J. J. Early, "Stochastic modelling and estimation of stationary complex-valued signals," *arXiv preprint: 1306.5993*, 2016.
- [28] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101–5125, 2011.
- [29] T. Gneiting, H. Ševčíková, D. B. Percival, M. Schlather, and Y. Jiang, "Fast and exact simulation of large Gaussian lattice systems in \mathbb{R}^2 : Exploring the limits," *Journal Comput. Graph. Stat.*, 2012.