Synthesis of Multivariate Stationary Series with Prescribed Marginal Distributions and Covariance using Circulant Matrix Embedding

Hannes Helgason[a], Vladas Pipiras[b,*], Patrice Abry[c,**]

[a]*School of Electrical Engineering, KTH, Stockholm (Sweden)*
[b]*Department of Statistics and Operations Research, University of North Carolina, Chapel Hill (USA)*
[c]*Laboratoire de Physique (CNRS UMR 5672), École Normale Supérieure de Lyon (France)*

Abstract

The problem of synthesizing multivariate stationary series $Y[n] = (Y_1[n], \ldots, Y_P[n])^T$, $n \in \mathbb{Z}$, with prescribed non-Gaussian marginal distributions, and a targeted covariance structure, is addressed. The focus is on constructions based on a memoryless transformation $Y_p[n] = f_p(X_p[n])$ of a multivariate stationary Gaussian series $X[n] = (X_1[n], \ldots, X_P[n])^T$. The mapping between the targeted covariance and that of the Gaussian series is expressed via Hermite expansions. The various choices of the transforms f_p for a prescribed marginal distribution are discussed in a comprehensive manner. The interplay between the targeted marginal distributions, the choice of the transforms f_p, and on the resulting reachability of the targeted covariance, is discussed theoretically and illustrated on examples. Also, an original practical procedure warranting positive definiteness for the transformed covariance at the price of approximating the targeted covariance is proposed, based on a simple and natural modification of the popular circulant matrix embedding technique. The applications of the proposed methodology are also discussed in the context of network traffic modeling. MATLAB codes implementing the proposed synthesis procedure are publicly available

*Work supported in part by the NSF grants DMS-0505628 and DMS-060866.
**Work supported by the *Young Research Team* award granted by Foundation del Duca, Académie des Sciences, Institut de France, 2007
During the course of this work, Hannes Helgason was employed at Laboratoire de Physique (CNRS UMR 5672), École Normale Supérieure de Lyon (France).

at http://www.hermir.org.

Keywords: Numerical Synthesis, Multivariate Gaussian Series, Circulant Matrix Embedding, Stationarity.

1. Introduction

1.1. Motivation

While studies in the early days of data analysis were often limited by a high cost of collecting data, the common problem in the modern era is that there is a plethora of data. Indeed, the amount of digital data being acquired and stored has increased dramatically due to the recent technological achievements in, for example, the development and manufacturing of computer hardware, sensors, wireless and digital networks, along with the popularity of social networking and online shopping over the Internet (e.g., Facebook, Twitter, Amazon). Examples are numerous: sensor network monitoring is nowadays increasingly used, e.g., in environmental or health surveillance [1, 2]; neural activity studied via functional magnetic resonance imaging, and hence the joint analysis of the time course of multiple neuron responses [3, 4]; human body monitoring in medicine [5]; Internet, where traffic might be collected jointly at numerous points of presence [6, 7, 8]; econometrics [9, 10]; etc. In many cases, these large data sets take the form of (not necessarily Gaussian) multivariate time series.

This induces major challenges to scientists and practitioners working in the field of data processing and statistics. Not only is there a need of fast algorithms for processing this ceaselessly rising pile of data, but there is also a necessity for new statistical methods for exploration, in particular, for cases where one has multiple measurements and different views of the same object or related objects of interest. Therefore, before heading out to the real world with new statistical tools and techniques, it is preferable to investigate their strengths and limitations under well-controlled and simplified situations by using academic mathematical models for the objects being studied and measured. However, sophisticated methods are usually analytically intractable, or prohibitively hard to study, even for the simplest models. In such cases, researchers might need to resort to Monte-Carlo simulations to analyze the methods and statistical problems through numerical experiments. Modern data and system analysis hence induces a crucial need for fast, accurate, and memory efficient numerical procedures enabling the synthesis of multivariate time series according to an a priori prescribed mathematical model.

1.2. Goals

In this context, the present contribution aims at studying the problem of, and at proposing a fast and efficient theoretical and practical procedure for synthesizing P-variate stationary series

$$Y[n] = (Y_1[n], \ldots, Y_P[n])^T, \quad n \in \mathbb{Z}, \tag{1}$$

whose marginal distributions, F_{Y_p}, and covariance structure, $R_Y[n] = EY[0]Y[n]^T - EY[0]EY[n]^T$, are targeted jointly and given a priori. Because this could be envisaged in different ways, the focus is here on constructions based on non-linear memoryless (or point) transforms $Y_p[n] = f_p(X_p[n])$, where $X[n] = (X_1[n], \ldots, X_P[n])^T$ is a multivariate stationary Gaussian series.

Note that not every joint selection of prescribed marginals and covariance structure is possible (see for example [11, 12], for the univariate case). This issue will play an important role throughout this article.

1.3. State-of-the-art

1.3.1. Univariate non-Gaussian synthesis

Most of the studies on synthesis of univariate non-Gaussian series $Y[n]$ are based on non-linear point transformations $Y[n] = f(X[n])$, with $X[n]$ a zero-mean unit variance stationary Gaussian series. One alternative approach is based on superposition of suitable Ornstein-Uhlenbeck processes and is particularly popular in applications to finance [13, 14].

The above constructions $Y[n] = f(X[n])$ are studied from several directions. One crucial issue is to be able to relate the covariance function r^Y of Y to r^X

that of X, as $r^Y[n] = \widetilde{g}(r^X[n])$. The pioneering work is that of Price [15], dating back to 1958, whose formulation in terms of differential equation however limits its practical use to a few possible choices for f (see also [16]). A number of more recent contributions [17, 18, 19, 11] establish this relationship between r^Y and r^X for general f by using Hermite polynomial expansions. Other approaches can be found in [20, 21] where the decomposition $X[n+k] = r_X(k)X[n] + Z[n,k]$, with a zero mean Gaussian series Z independent of X, is used, and in [22], studies specific analytically tractable cases.

Another related direction concerns synthesizing series $Y[n]$ with prescribed marginal distribution F_Y and covariance function r^Y. To match the marginal distribution exactly, a common choice is to take the transformation $f(x) = F_Y^{-1}(\Phi(x))$, where $\Phi(x)$ denotes the distribution function of a standard Gaussian variable and F_Y^{-1} the generalized inverse of F_Y. For a targeted covariance r^Y, it is natural to write r^X as $r^X[n] = \widetilde{g}^{-1}(r^Y[n])$. It is known that $r^X[n]$ obtained as above does not necessarily define a valid covariance structure [17, 23]. Several constructions of valid covariance structures r^X which only approximate the targeted r^Y through $\widetilde{g}(r^X[n])$ in some optimal sense, have been proposed [17, 18, 11, 24]. Another synthesis method, based on the use of surrogates and an iterative algorithm, yields an approximation of the targeted covariance [25]. Methods based on “shuffling” can be found in [26, 27].

Real-world applications of the methods discussed above include, to name but a few, modeling of wind pressure on a building [28], sea wave height [27], Internet traffic [22, 8], neurotransmission signals [20], material properties [29], and sea clutter [30]. See also Section 6 below.

1.3.2. Multivariate non-Gaussian synthesis

The multivariate case seems to have received far less effort, even in the Gaussian case where there are basically two popular approaches: in [31] a fast but approximate method is provided which starts from a given spectral density matrix, and more recent developments [32, 33] provide fast and exact methods starting from the covariance. For non-Gaussian series, the univariate spectral representation method has been extended to the multivariate case in [29]. The method is iterative and starts from a given spectral density matrix of $Y[n]$ and is only valid for constructions of the type $Y_p[n] = F_{Y_p}^{-1}(\Phi(X_p[n]))$. A simple illustration of this method with a comparison to the method proposed here is given in Section 4.2.5.

A bivariate construction, extending the approach in [20, 21], has been proposed in [8] to target marginals with Erlang distributions. Its usefulness for other distributions however remains restricted to analytical tractability.

1.4. Contributions

In view of the goals listed in Section 1.2, the present work contributes to the existing literature described above in the following main ways.

First, we focus on synthesis of P-variate stationary non-Gaussian series. Most of the available literature concerns the univariate case. The single multivariate paper [29] focuses more broadly on multivariate random fields, and often lacks specifics of the P-variate case alone. Moreover, the approach presented here will be very different from that in [29]. For example, whereas a spectral density matrix is targeted in [29], we focus on covariances alone.

Second, we study other natural ways to match marginals than just the transformation $f(x) = F_Y^{-1}(\Phi(x))$. The choice of transformation affects which covariance structures can be achieved and how easily one can practically reach them. But, first and foremost, varying the transformation offers practitioners with a flexibility of synthesizing a variety of different series (having the same marginals and same covariances).

Third, in contrast to earlier work, we pay more attention to numerical issues. In a number of our examples, we work only with numerical values of Hermite coefficients of the transformation f_p of interest. In the past, examples were typically based on analytic forms of transformations and explicit forms of coefficients.

Fourth and related to the above, we show how the so-called series reversion can be used to obtain r^X from r^X, when only numerical values of Hermite coefficients are available. To the best of our knowledge, reversion has not been used yet in this area.

Fifth, for the actual synthesis, we work with the so-called circulant matrix embedding method. This is a fast, exact, and most popular method to generate stationary Gaussian series given its covariance structure. When inversion of targeted covariance structure does not lead to a valid covariance structure, we propose an original and natural modification to the circulant matrix embedding procedure, which provides an approximation to the targeted covariance. The procedure is akin to spectral truncation used in [17], and is argued to be optimal in a suitable sense. We should also emphasize that this seems to be the first work in the area to use the circulant matrix embedding.

We also note that the last four points are new to the area even for the univariate case. For this reason and for the sake of clarity, we shall illustrate some of our ideas with examples which are more univariate than multivariate in spirit.

Matlab routines implementing the synthesis procedure are publicly available at `http://www.hermir.org`.

1.5. Organization

The structure of this article is as follows. Section 2 describes the main framework, based on memoryless transformations $Y_p = f_p(X_p)$, and is split into several sections: (i) description of the relationship between the covariances of the targeted series Y_p and of the underlying Gaussian series X_p (Section 2.2), (ii) comparison of various possible choices for transforms f_p in matching prescribed marginals (Section 2.3), (iii) discussion of issues related to restrictions one is faced with when targeting jointly marginals and covariance (Section 2.4), and (iv) illustrations on concrete examples (Section 2.5 and Appendix B). Inversion of the relationship between covariances is discussed in Section 3. Synthesis procedures based on circulant matrix embedding can be found in Section 4; a simple illustration and a comparison to an existing method are given in Sections 4.2.4 and 4.2.5. In Section 6 we discuss the application of the proposed methodology in network traffic modeling. Some additional details about multivariate Gaussian synthesis are in Appendix A.

In Section 5, the analysis conducted in the earlier sections is extended to the case where the components of the targeted series Y are obtained not only from a single but from a collection of independent copies of auxiliary Gaussian series. Technical details are in Appendix C and Appendix D.

For practical convenience, we have provided a collection of popular marginals and transformations where the covariance relationship is given by a simple analytic formula which can easily be inverted (Appendix E). Finally, Appendix F gives a step-by-step summary gathering the main results of the paper and showing how one can apply the proposed methodology in practice.

2. Component-Wise Transformations of Gaussian Multivariate Series

2.1. Framework

We focus here on non-Gaussian stationary series $Y[n] = (Y_1[n], \ldots, Y_P[n])^T$ obtained as

$$Y_p[n] = f_p(X_p[n]), \quad n \in \mathbb{Z},\ 1 \le p \le P, \tag{2}$$

where $f_p : \mathbb{R} \mapsto \mathbb{R}$ are real functions, and $X[n] = (X_1[n], \ldots, X_P[n])^T$ is a Gaussian stationary multivariate series. Set also $f(x) = (f_1(x_1), \ldots, f_P(x_P))^T$ for $x = (x_1, \ldots, x_P)^T \in \mathbb{R}^P$.

For a series $Z[n]$, standing for either $X[n]$ or $Y[n]$, denote its covariance structure by

$$r^Z_{p,p'}[n] = EZ_p[0]Z_{p'}[n] - EZ_p[0]EZ_{p'}[n], \quad n \in \mathbb{Z},\ 1 \le p, p' \le P, \tag{3}$$

and let also $R_Z[n] = EZ[0]Z[n]^T - EZ[0]EZ[n]^T = (r^Z_{p,p'}[n])_{1 \le p,p' \le P}$ be the $P \times P$ auto-covariance matrices. Denote the correlation sequences by

$$\rho^Z_{p,p'}[n] = \frac{r^Z_{p,p'}[n]}{\sqrt{r^Z_{p,p}[0] r^Z_{p',p'}[0]}}. \tag{4}$$

We suppose that the series $X[n]$ is such that $EX_p[n] = 0$, $EX_p[n]^2 = r^X_{p,p}[0] = 1$, and therefore, $r^X_{p,p'}[n] = \rho^X_{p,p'}[n]$.

2.2. Relationships between covariances

We discuss here several relationships between the covariances of the series $X[n]$ and $Y[n]$ related through (2). These relations are useful in at least two ways: first, they characterize the covariance structure of the series $Y[n]$ in terms of that of $X[n]$ and, second, they will be used in inversion where the goal is to match a targeted covariance structure of $Y[n]$ by choosing a suitable one for $X[n]$.

2.2.1. Hermite polynomial expansions

Let $H_m(x)$, $m = 0, 1, \dots$, be univariate Hermite polynomials, that is, $H_0(x) = 1$, $H_m(x) = (-1)^m e^{x^2/2} d^m e^{-x^2/2}/dx^m$, $m \geq 1$. Recall that Hermite polynomials form an orthogonal basis of the Hilbert space $L^2(\mathbb{R}, e^{-x^2/2}dx)$. Suppose that the functions $f_p \in L^2(\mathbb{R}, e^{-x^2/2}dx)$ (or equivalently $Y_p[n] \in L^2(\Omega)$) so that they can be expanded in Hermite polynomials as

$$f_p(x) = \sum_{m=0}^{\infty} c_m^{(p)} H_m(x), \tag{5}$$

where the convergence of the series is in the norm induced by the inner-product $\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2/2}dx$ on the space $L^2(\mathbb{R}, e^{-x^2/2}dx)$. Using the relations $EH_m(Z) = 0$, $m \geq 1$, and

$$EH_m(Z)H_n(W) = \begin{cases} m!(EZW)^m, & \text{if } m = n, \\ 0, & \text{if } m \neq n, \end{cases}$$

for a Gaussian vector (Z, W) with standard Gaussian marginals and $m, n \geq 1$ (see Lemma 1.1. in [34]), we obtain that

$$r^Y_{p,p'}[n] = \sum_{m=1}^{\infty} c_m^{(p)} c_m^{(p')} m! (r^X_{p,p'}[n])^m. \tag{6}$$

This can also be written as

$$R_Y[n] = \sum_{m=1}^{\infty} C^{(m)} R_X^{(m)}[n] C^{(m)}, \tag{7}$$

where $C^{(m)} = diag(c_m^{(1)}, \dots, c_m^{(P)})$ and

$$R_X^{(m)}[n] = ((r^X_{p,p'}[n])^m)_{1 \leq p,p' \leq P} = E\tilde{H}_m(X[0])\tilde{H}_m(X[n])^T$$

is the auto-covariance matrix of the series $\tilde{H}_m(X[n]) = (H_m(X_1[n]), \ldots, H_m(X_P[n]))^T$.

2.2.2. Price theorem

Another useful relationship follows from Price theorem [15]. More precisely, under further suitable conditions on the functions $f_p(x)$ and $f_{p'}(x)$, and for $k \geq 1$, one has

$$\begin{aligned}\frac{\partial^k r^Y_{p,p'}[n]}{\partial(r^X_{p,p'}[n])^k} &= E\frac{\partial^k f_p(X_p[0])}{\partial x^k}\frac{\partial^k f_{p'}(X_{p'}[n])}{\partial x^k} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\partial^k f_p}{\partial x^k}(x_1)\frac{\partial^k f_{p'}}{\partial x^k}(x_2)\frac{1}{2\pi\sqrt{1-r^X_{p,p'}[n]^2}} \\ &\qquad \exp\left\{-\frac{x_1^2+x_2^2-2r^X_{p,p'}[n]x_1x_2}{2(1-r^X_{p,p'}[n]^2)}\right\}dx_1dx_2. \qquad (8)\end{aligned}$$

When $p = p'$, one set of sufficient conditions on function f_p for (8) to hold can be found in [23], Proposition 2 (though stated without a proof). Obtaining such conditions for (8) to hold, however, is not our focus here.

2.3. Matching marginals: Various possibilities

Apart from covariances, we are interested first of all in matching the prescribed marginals of the component series $Y_p[n]$. The bijective transformation

$$f(x) = F_Y^{-1}(\Phi(x)) \qquad (9)$$

of a standard Gaussian variable leads to a desired marginal distribution F_Y, where $F_Y^{-1}(u) = \inf\{x : F_Y(x) = u, 0 < u < 1\}$ is the generalized inverse; when the cumulative distribution function is strictly monotone, F_Y^{-1} is simply the inverse of F_Y. This transformation is the one commonly used in previous general constructions. We shall later refer to it as a *standard* transformation.

Other choices than (9) are possible to match the marginal, and will be given below. Unlike the standard transformation, for example, these other transformations may have explicit forms and hence might be easier to implement in practice (Section 2.5.1 below). Another important point to keep in mind is that marginal distribution and covariance structure do not characterize a non-Gaussian series. While there is only one Gaussian series with a given a priori prescribed covariance, there are infinitely many different non-Gaussian series that have the same covariance functions and the same marginal distributions. It is of crucial importance to have at disposal a procedure allowing for many different non-Gaussian series, even with identical marginals and covariances. For further discussion on this in the context of network traffic modeling, see Section 6. See also [28] where higher order moments are targeted as well.

The basic idea behind (9) is that $\Phi(X)$ is $U(0,1)$ (uniform on $(0,1)$) variable for a standard Gaussian variable X. Other transformations could be introduced by other constructions of $U(0,1)$ variables. Another way to get a $U(0,1)$ variable is to take $2(\Phi(|X|)-1/2)$ which leads to a transformation

$$f(x) = F_Y^{-1}\left(2(\Phi(|x|)-1/2)\right). \tag{10}$$

Note that this transformation is necessarily even, that is, $f(x) = f(-x)$, $x \in \mathbb{R}$. In the same spirit, another large class of transformations can be obtained as

$$f(x) = F_Y^{-1}\left(\xi\left(\Phi(x)\right)\right), \tag{11}$$

where $\xi : [0,1) \to [0,1)$ is a bijective mapping, continuously differentiable except perhaps on a finite number of points and with the derivative

$$\xi'(v) = \pm 1,$$

where it exists. For such mapping, there are finite number of points $0 = v_0 < v_1 < \cdots < v_L = 1$ such that ξ is bijective and continuously differentiable with derivative ξ' equal to $+1$ or -1 throughout each interval (v_{l-1}, v_l), $l = 1, \dots, L$. On each image $(\xi(v_{l-1}), \xi(v_l)) =: (u_{l-1}, u_l)$, with the definition $\xi(v_L) = \xi(1) := \lim_{v \to v_L^-} \xi(v)$, the inverse function ξ^{-1} is bijective and continuously differentiable with the derivative $(\xi^{-1})'$ equal to -1 or $+1$. Note now that ξ takes a $U(0,1)$ variable V to another $U(0,1)$ variable $\xi(V)$ since, for bounded functions h,

$$\begin{aligned}
Eh(\xi(V)) &= \int_0^1 h(\xi(v))dv \\
&= \sum_{l=1}^{L} \int_{(v_{l-1},v_l)} h(\xi(v))dv = \sum_{l=1}^{L} \int_{\xi^{-1}(u_{l-1},u_l)} h(\xi(v))dv \\
&= \sum_{l=1}^{L} \int_{(u_{l-1},u_l)} h(\xi(\xi^{-1}(u)))\ |(\xi^{-1})'(u)|du \\
&= \int_0^1 h(u)du,
\end{aligned}$$

by using the change of variables $v = \xi^{-1}(u)$ and the fact that $(\xi^{-1})'(u) = \pm 1$. This shows that the transformation (11) of a standard Gaussian variable indeed leads to a marginal distribution function F_Y. For a concrete example, one can take $\xi(v) = 1 - v$ or $\xi(v) = v - \frac{1}{2}\text{sign}\left(v - \frac{1}{2}\right)$, which interchanges the images of the function $\xi(v) = v$ on $[0, 1/2)$ and $[1/2, 1)$. More involved examples can be constructed by further swapping parts of the image of $\xi(v) = v$. For example, let $I_\ell = [a_\ell, a_{\ell+1})$, $\ell = 1, \dots, L$, with $a_\ell = (\ell - 1)/L$, be a partition of $[0, 1)$. Let also π be a permutation of $\{1, \dots, L\}$. Then, one can take ξ defined as $\xi(v) = v - a_\ell + a_{\pi(\ell)}$ for $v \in I_\ell$.

Note that different transformations (9), (10), and (11) are likely to have *different* Hermite expansions in general. Hence, the relationship (6) between covariances r^Y and r^X will be different, and will likely affect what covariance structure r^Y could be reached for various transformations. It remains an open, interesting, and, in our view, difficult question on whether and which particular transformation leads to the largest class of attainable covariances r^Y. In fact, we suspect that this might be the case with the standard transformation. But even if the standard transformation were to lead to the largest class of attainable covariances, other transformations should still be considered. As indicated above, other transformations may allow, for example, for easier implementation and having different time series with the same marginal and covariance structure. See also Section 2.5.2 below.

Section 2.4 below further expands this discussion by detailing the restrictions on the achievable covariance stemming from the goal of matching jointly prescribed marginals and covariances, while Section 2.5 illustrates, on concrete examples, the effects of varying f when targeting the same set of given marginal distributions.

2.4. Joint matching of marginals and covariance: Restrictions

According to Section 2.3 above, the prescribed marginals of component series $Y_p[n]$ can be matched exactly in a variety of ways. For any such fixed transformations and for a suitable choice of the series $X[n]$, however, not every targeted covariance structure of $Y[n]$ can be matched exactly. One reason for this is that the transformation (6) may already impose restrictions on the range of values that can be taken by the targeted covariances $r^Y_{p,p'}[n]$.

Before discussing this further, we rewrite relationship (6) in terms of correlation sequences

$$\rho^Y_{p,p'}[n] = \sum_{m=1}^{\infty} \frac{c_m^{(p)} c_m^{(p')} m!}{\sqrt{r^Y_{p,p}[0] r^Y_{p',p'}[0]}} (\rho^X_{p,p'}[n])^m. \tag{12}$$

Note that this relationship imposes $|\rho^Y_{p,p'}[n]| \leq |\rho^X_{p,p'}[n]|$. Indeed, this is so because

$$\begin{aligned}
|\rho^Y_{p,p'}[n]| &= \left| \sum_{m=1}^{\infty} \frac{c_m^{(p)} c_m^{(p')} m!}{\sqrt{r^Y_{p,p}[0] r^Y_{p',p'}[0]}} (\rho^X_{p,p'}[n])^m \right| \\
&\leq \sum_{m=1}^{\infty} \frac{|c_m^{(p)}|\sqrt{m!}}{\sqrt{r^Y_{p,p}[0]}} \cdot \frac{|c_m^{(p')}|\sqrt{m!}}{\sqrt{r^Y_{p',p'}[0]}} |\rho^X_{p,p'}[n]|^m \\
&\leq |\rho^X_{p,p'}[n]| \left(\sum_{m=1}^{\infty} \frac{(c_m^{(p)})^2 m!}{r^Y_{p,p}[0]} \right)^{1/2} \left(\sum_{m=1}^{\infty} \frac{(c_m^{(p')})^2 m!}{r^Y_{p',p'}[0]} \right)^{1/2} \\
&= |\rho^X_{p,p'}[n]| \sqrt{\rho^Y_{p,p}[0] \rho^Y_{p',p'}[0]} = |\rho^X_{p,p'}[n]|,
\end{aligned}$$

using the triangular inequality, the fact $|x|^m \leq |x|$, $m \geq 1$, for $|x| \leq 1$, and the Cauchy-Schwarz inequality (the inequality can also be found in [35]). Relationship (12) can be written as $\rho^Y_{p,p'}[n] = g_{p,p'}(\rho^X_{p,p'}[n])$, where the function $g_{p,p'}$ is defined as

$$g_{p,p'}(z) = \sum_{m=1}^{\infty} \frac{c_m^{(p)} c_m^{(p')} m!}{\sqrt{r_{p,p}^Y[0] r_{p',p'}^Y[0]}} z^m =: \sum_{m=1}^{\infty} b_m^{(p,p')} z^m. \tag{13}$$

It is clear that the range of $g_{p,p'}(z)$ for $z \in [-1,1]$ defines the range of reachable values for the correlation sequences $\rho_{p,p'}^Y[n]$. This range is in general not equal to $[-1,1]$. In practice, when given transforms f_p and targeted covariance R_Y, one should first check that the correlation sequences actually fall into this range. This can be done numerically, as discussed in Section 2.5 below. On the other, theoretical side, some results about this range are discussed in the remainder of this section.

2.4.1. Theoretical bounds for auto-correlation sequences

First consider (13) for $p = p'$, which corresponds to auto-correlation sequences. Note that $g_{p,p}(0) = 0$ and $g_{p,p}(1) = \sum_{m=1}^{\infty} (c_m^{(p)})^2 m!/r_{p,p}^Y[0] = 1$, where we used (6) and $r_{p,p}^X[0] = 1$. Since $b_m^{(p,p)} \geq 0$, this shows that the function $g_{p,p}(z)$ is monotonically increasing from $z = 0$, $g_{p,p}(0) = 0$ to $z = 1$, $g_{p,p}(1) = 1$. Thus, the range of function $g_{p,p}(z)$ for $z \in [0,1]$ is $[0,1]$, and any *positive* value of the targeted covariance can in principle be reached.

The situation turns out to be quite different for the range of the function $g_{p,p}(z)$ for $z \in [-1,0]$. It is generally not true that this range includes the whole interval $[-1,0]$. For example, with the even transformation (10), the Hermite coefficients $c_m^{(p)}$ are zero at odd $m = 1, 3, \ldots$. The function

$$g_{p,p}(z) = \sum_{k=1}^{\infty} \left(c_{2k}^{(p)} / \sqrt{r_p^Y[0]} \right)^2 (2k)! z^{2k} \tag{14}$$

is then even, and necessarily non-negative for all $z \in [-1,1]$. In particular, with this transformation, negative values of the targeted covariance can not be reached.

There are several attempts to quantify theoretically the range of $g_{p,p}(z)$ for $z \in [-1,0]$. In the case where f_p is monotone, Price theorem (Eq. (8)) with $k = 1$ shows, since $\frac{\partial f_p(u)}{\partial x} \frac{\partial f_p(v)}{\partial x} > 0$, $\forall u, v \in \mathbb{R}$, that $g_{p,p}(z)$ is monotonically increasing [12]. Then

$$g_{p,p}(z) \geq \xi_p^* \tag{15}$$

with

$$\xi_p^* = \frac{\mathrm{Cov}(f_p(X), f_p(-X))}{\mathrm{Var}(f_p(X))} = g_{p,p}(-1),$$

where X is a $\mathcal{N}(0,1)$ variable and the last equality follows by using the Hermite expansion (5) of $f_p(x)$. A sufficient condition for $\xi_p^* = -1$ to hold is that f_p is odd, $f_p(-x) = -f_p(x)$, $\forall x \in \mathbb{R}$ (see [12]). Otherwise, ξ_p^* is typically not equal to -1, as illustrated by the example in Figure 1 in Section 3.2.3 below.

The relation (15), however, is obviously not true for the even function (14), where $\xi_p^* = g_{p,p}(-1) = g_{p,p}(1) = 1$. Another inequality akin to (15) can be found in [17], at the end of Section III. It is unclear, however, whether that inequality actually provides a tight lower bound on targeted covariances.

2.4.2. Theoretical bounds for cross-correlation sequences

Less can be said about the range for functions $g_{p,p'}(z)$, $p \neq p'$. However, if f_p and $f_{p'}$ are both either monotonically increasing or decreasing, then $\frac{\partial f_p(u)}{\partial x}\frac{\partial f_{p'}(v)}{\partial x} > 0$, $\forall u, v \in \mathbb{R}$, and the same Price theorem shows that $g_{p,p'}(z)$ is monotonically increasing. In those cases, the interval $[g_{p,p'}(-1), g_{p,p'}(1)]$ is the reachable range for $\rho^Y_{p,p'}$. Let us further note that if f_p and $f_{p'}$ are equal, $g_{p,p'} \equiv g_{p,p}$, and the analysis in Section 2.4.1 will also hold for $g_{p,p'}$.

2.5. Examples

We give here several examples illustrating the ideas of the previous Sections 2.2, 2.3 and 2.4. We consider several prescribed marginal distributions, and discuss possible transformations and properties of resulting covariances of non-Gaussian series. The first two examples assume that $f_1 = \cdots = f_P$ in Eq. (2).

2.5.1. χ^2_1 marginals and negative correlations

The classical way a χ^2_1 marginal can be achieved is by taking $f_p(x) = x^2$. One can check that this corresponds to the construction (10), that is, $x^2 =$

$F^{-1}_{\chi^2_1}(2(\Phi(|x|) - 1/2))$. The relevant Hermite expansion here is

$$f_p(x) = x^2 = H_0(x) + H_2(x) \tag{16}$$

and hence the only nonzero coefficients in the expansion (5) are $c_0^{(p)} = c_2^{(p)} = 1$. Then (12) and $r_p^Y[0] = \mathrm{Var}(Y_p[n]) = 2$ give that

$$\rho^Y_{p,p'}[n] = (\rho^X_{p,p'}[n])^2 \tag{17}$$

and the corresponding function in (13) is $g_{p,p'}(z) = z^2 \equiv g(z)$. Eq. (17) shows, in particular, that with transformation (16), the series with χ^2_1 marginals can be only positively correlated. This is a special case of the general discussion around (14) in Section 2.4. Note that (17) can also be easily obtained using Price theorem (see Appendix B).

Another possibility to match the marginal is to take

$$f_p(x) = F^{-1}_{\chi^2_1}(\Phi(x)). \tag{18}$$

In fact, the function (18) significantly differs from (16). For example, observe that $f_p(x) \to 0$, as $x \to -\infty$, with (18), while $f_p(x) \to +\infty$, as $x \to -\infty$, with (16). The Hermite coefficients for the transformation (18) and hence the functions $g_{p,p'}(z) \equiv g(z)$ in (13) can only be computed numerically. The plot of the function $g(z)$ can be found in Section 3.2.3, Figure 1 (solid line). Note that, unlike the function $g(z) = z^2$ for (16), $g(z)$ for (18) now takes negative values. In particular, with the transformation (18) and in contrast to (16), when the series X_p has negative correlations, the series Y_p with χ^2_1 marginal will also have negative correlations.

2.5.2. Same marginal and covariance but different series

Further using the χ^2_1 marginal example, let us now choose an always positive targeted covariance, for example, an AR(1) correlation $\rho^Y[n] = \phi^{|n|}$ with a positive parameter $0 < \phi < 1$. It can be easily shown that using both transformations (16) and (18), time series with the desired χ^2_1 marginal and AR(1) covariance can be synthesized exactly. For (16), it is straightforward to check analytically, since $\rho^X[n] = (\sqrt{\phi})^{|n|}$ is a valid covariance structure, and for (18), the procedure proposed in the present contribution enables to check this numerically in a simple manner. This example can be readily extended to the multivariate case. Therefore, we have here easily synthesized time series that have identical marginal and covariance, but are yet different in the sense that their joint higher order statistics differ. This possibility is of major practical importance, enabling practitioners to test the impact of higher order statistics on the system they are studying, while keeping the first and second order moments fixed.

2.5.3. Log-normal (LN) marginals and negative correlation

Let us now consider another example of major practical importance: time series with log-normal marginals, chosen as the emblematic representative of heavy tailed distributions. A natural choice here is to take $f_p(x) = e^{\sigma x + \mu}$, $\sigma > 0$, $\mu \in \mathbb{R}$. One can check that this choice corresponds to the standard transformation (9), that is, $f_p(x) = e^{\sigma x+\mu} = F_{LN}^{-1}(\Phi(x))$. Consider the case $\mu = 0$. Using the well-known identity

$$e^{xt - t^2/2} = \sum_{m=0}^{\infty} \frac{t^m}{m!} H_m(x), \tag{19}$$

one gets

$$f_p(x) = e^{\sigma x} = \sum_{m=0}^{\infty} \frac{e^{\sigma^2/2}\sigma^m}{m!} H_m(x),$$

from which the Hermite coefficients are $c_m^{(p)} = e^{\sigma^2/2}\sigma^m/m!$. Eq. (12) and $r_p^Y[0] = \mathrm{Var}(Y_p[n]) = e^{2\sigma^2} - e^{\sigma^2} =: c_\sigma^2$ now give

$$\rho_{p,p'}^Y[n] = \sum_{m=1}^{\infty} \frac{e^{\sigma^2}}{c_\sigma^2 m!} (\sigma^2 \rho_{p,p'}^X[n])^m = \frac{e^{\sigma^2}}{c_\sigma^2} (e^{\sigma^2 \rho_{p,p'}^X[n]} - 1). \tag{20}$$

Changing to general μ only introduces the same multiplicative factor e^μ to each component Y_p and therefore the relationship between the correlation sequences stays the same. Relation (20) can also be obtained using Price theorem (see Appendix B).

Another possibility to get log-normal marginals is to consider transformation (10), that is,

$$f_p(x) = F_{LN}^{-1}\left(2(\Phi(|x|) - 1/2)\right). \tag{21}$$

This is a different transformation from the one considered above. As discussed around (14) in Section 2.4, with transformation (21) the series with log-normal marginals can only be positively correlated. Plots of the function $g_{p,p}(z) \equiv g(z)$ for transformation (21) can be found in Section 3.2.3, Figures 2 and 3.

2.5.4. One component series with Gaussian marginal

Consider $f_1(x) = x$ and suppose that $f_p(x)$ is nonlinear for $p \geq 2$. In other words, the first component series has a Gaussian marginal while other components have not. Since $f_1(x) = H_1(x)$, the only nonzero coefficient in its Hermite expansion is $c_1^{(1)} = 1$. From Eq. (6), we have

$$r_{1,p}^Y[n] = c_1^{(p)} r_{1,p}^X[n].$$

In particular, if $c_1^{(p)} = 0$, then necessarily $r_{1,p}^Y[n] = 0$. For example, this occurs with $f_p(x) = x^2$. Thus, a Gaussian series cannot be correlated with a series having χ_1^2 marginal obtained through transformation x^2. If the transformation $f_p(x) = F_{\chi_1^2}^{-1}(\Phi(x))$ is used instead for χ_1^2 marginal, we can check numerically that $c_1^{(p)} \neq 0$ and hence that the two series can be correlated. This again points to the fact that various possibilities discussed in Section 2.3 lead to series with very different properties.

3. Inversion

The synthesis of $Y[n]$, with a priori targeted covariance $r_{p,p'}^Y$, will rely on applying the transforms f_p to series $X[n]$, synthesized using the circulant embedding procedure which is based on $r_{p,p'}^X$ (Section 4 below). This implies the need to express $r_{p,p'}^X$ in terms of $r_{p,p'}^Y$, that is, to invert the relation (6). Issues related to this practical inversion are addressed in the present section. Inversion *per se* does not guarantee a valid covariance structure. The handling of this issue is postponed to Section 4.

3.1. Examples of inversion through explicit formula

As seen from the examples in Section 2.5, in some cases inversion can be carried out through an explicit formula.

(a) With χ_1^2 marginals and if $f_p(x) = x^2$, the relation (17) can be be inverted as $\rho^X_{p,p'}[n] = \pm\sqrt{\rho^Y_{p,p'}[n]}$. Using (4) with $r^Y_{p,p}[0] = 2$ and $r^X_{p,p}[0] = 1$, the formula for the covariances becomes

$$r^X_{p,p'}[n] = \pm\sqrt{\frac{r^Y_{p,p'}[n]}{2}}.$$

This inversion formula restricts input covariance structures $r^Y_{p,p'}$ to be nonnegative. Moreover, the output is not necessarily a valid covariance structure.

More specifically, $\pm\sqrt{r[n]}$ may not be nonnegative definite for (nonnegative) autocovariance function $r[n]$. For example, the function $r[n]$ with $r[0] = 1$, $r[n] = 0$, $n \geq 2$, is nonnegative definite if and only if $|r[1]| \leq 1/2$ (Example 2.1.1 on p. 48 in [36]). For a finite length sequence $r[n]$, $n = 0, \dots, N-1$, with $r[0] = 1$, $r[n] = 0$, $2 \leq n \leq N-1$, the sequence is nonnegative definite if and only if $|r[1]| \leq -1/(2\cos(N\pi/(N+1)))$ (see [37]). The same condition has to hold for the sequence $\sqrt{r[n]}$ in order for it to be nonnegative definite. Therefore $r[1]$ would have to satisfy the stronger inequality $r[1] \leq 1/(2\cos(N\pi/(N+1)))^2$, where the right side tends to $1/4$ as $N \to \infty$.

(b) With log-normal marginals and if $f_p(x) = e^{\sigma x + \mu}$, the relation (20) can be inverted as

$$\rho^X_{p,p'}[n] = \frac{1}{\sigma^2}\log(c_\sigma^2 e^{-\sigma^2}\rho^Y_{p,p'}[n] + 1). \tag{22}$$

This relationship imposes $\rho^Y_{p,p'}[n] > -c_\sigma^{-2}e^{\sigma^2}$. Note that the parameter μ does not appear in this restriction. As in part (a) above, we also expect that the output in (22) is not necessarily a valid covariance structure.

3.2. Series reversion

In many cases, the relationship (6) between covariances is available only in the series form with numerically computed coefficients and an explicit formula for inversion cannot be obtained. This is the case, for example, in Section 2.5.1

when using transformation (18), and in Section 2.5.3 when using transformation (21). In these instances, one possible way to invert (6) is to use the so-called series reversion (see, for example, Section 1.7 in [38]). We next discuss series reversion in two practical situations of interest. We note again that the use of reversion in the context of non-Gaussian series simulation appears to be new.

3.2.1. Reversion when first Hermite coefficient is nonzero

Inversion of (6) is equivalent to inverting the function $g_{p,p'}(z)$ defined in (13). To simplify notation, we omit indices p, p' and write

$$g(z) = \sum_{m=1}^{\infty} b_m z^m.$$

Suppose that

$$b_1 \neq 0. \tag{23}$$

The radius of convergence of the series defining the function g (e.g. in the sense of [38]) is at least 1 and, in particular, is positive. Then, by Theorem 2.4c in [38], p. 97, the function $g(z)$ has an inverse g^{-1} in a small neighborhood of $z = 0$ (that is, $g(z)$ is one-to-one and $g^{-1}(g(z)) = z$ in this neighborhood), and the inverse g^{-1} can be expressed as

$$g^{-1}(w) = \sum_{m=1}^{\infty} d_m w^m$$

for small enough neighborhood of $w = 0$, where the sequence $\{d_m\}_{m\geq1}$ is defined by the so-called reversion of the sequence $\{c_m\}_{m\geq1}$ (Section 1.7 of [38]). Reversion is a straightforward operation where the coefficient d_m is expressed in terms of $b_1, \ldots, b_m$, for example, $d_1 = b_1^{-1}$, $d_2 = -b_1^{-3} b_2$, $d_3 = b_1^{-5}(2b_2^2 - b_1 b_3)$, etc. Formally, the coefficients d_m can be obtained by substituting the series $g^{-1}(w)$ and $g(z)$ into $g^{-1}(g(z)) = z$ and equating the coefficients d_m at the powers of z of both sides of the equation. Numerically, the coefficients are easy to compute using Theorem 1.7c in [38], p. 47. An implementation of sequence reversion is included in the publicly available software package mentioned in Section 8.

A theoretical and technical issue with the reversion above is that the radius of convergence of g^{-1} cannot be determined easily from the coefficients b_m of g. (The best one knows seems to be the lower bound found after the proof of Theorem 2.4b in [38], p. 99. Though even this is difficult to use just having the coefficients b_m.) A related difficulty is that it is also difficult to determine in theory where exactly g^{-1} defines the inverse of g.

From a practical perspective, the issues above could be sidestepped at the expense of the following more careful analysis. By plotting $(z, g(z))$ for $z \in [-1, 1]$, one can determine a range $R_{0,g}$ of values w where g^{-1} is expected. By plotting $(g^{-1}(w), w)$ for $w \in R_{0,g}$ (and where $g^{-1}(w)$ is obtained by reversion), one can determine the range $R_g \subset R_{0,g}$ where $g^{-1}(w)$ can be taken as the inverse of g. The inverse $g^{-1}(w)$ could then be used when the targeted values $\rho^Y_{p,p'}[n]$ fall in this range R_g. This is further discussed in Section 3.2.3 below.

3.2.2. Series reversion for even functions

The approach in Section 3.2.1 above assumes that the first Hermite coefficient $c_1^{(p)}$ of f_p is nonzero as in (23). This is not the case, for example, when $f_p(x)$ is even, for which the function $g_{p,p'}(z)$ has the form (14). Dropping indices p, p', we thus have

$$g(z) = \sum_{k=1}^{\infty} b_{2k} z^{2k} = \sum_{k=1}^{\infty} a_k y^k =: h(y),$$

where $a_k = c_{2k}$ and $y = z^2$. In this case and when

$$b_2 \neq 0, \tag{24}$$

the following natural approach can be followed.

Use the reversion of $\{a_k\}_{k\geq 1}$ ($a_1 \neq 0$ by (24)) to obtain

$$h^{-1}(w) = \sum_{k=1}^{\infty} d_k w^k.$$

Then $h^{-1}(g(z)) = h^{-1}(h(y)) = y = z^2$ so that

$$z = \pm \sqrt{\sum_{k=1}^{\infty} d_k g(z)^k}$$

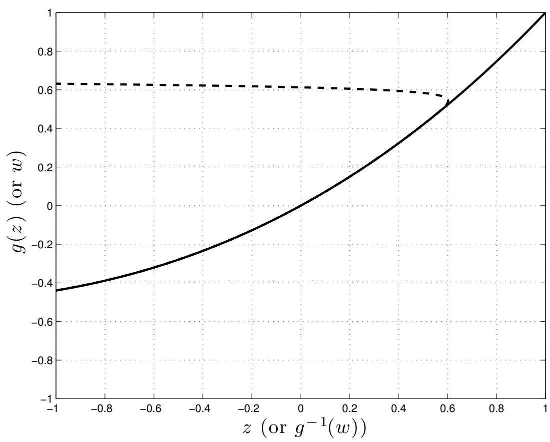

Figure 1: Plots of $g(z)$ and $g^{-1}(w)$ when $f_p(x)$ is given by (18). The curve $(g^{-1}(w), w)$ (broken line) overlaps with the curve $(z, g(z))$ until $w \approx 0.47$.

or $g^{-1}(w) = \pm\sqrt{h^{-1}(w)}$.

3.2.3. Numerical illustrations

We illustrate the ideas above through several examples.

(a) Consider the case of χ_1^2 marginals where the transformation $f_p(x)$ is given by (18). Recall that for this transformation, $g_{p,p'}(z) \equiv g(z)$ or its inverse do not have explicit forms, and only numerical values of the coefficients $c_m^{(p,p')} \equiv c_m$ are available. In Figure 1, we provide the plots of $g(z)$ and its inverse $g^{-1}(w)$ obtained by reversion. As discussed in Section 2.4, note that the range of $g(z)$, $z \in [-1, 1]$, is not the whole interval $[-1, 1]$ but rather $R_{0,g} = [g(-1), 1]$ with $g(-1)$ around -0.44. As indicated in Section 3.2.1 above, note also that $g^{-1}(w)$ defines the inverse of $g(z)$ not on the whole interval $R_{0,g}$ but rather on a smaller domain $R_g = [g(-1), \eta^*]$, where η^* is about 0.47.

(b) Consider the case of lognormal marginals where the transformation $f_p(x)$ is given by (21) and $\mu = 0$. This is another example where no explicit forms are available for $g_{p,p'}(z) \equiv g(z)$ or its inverse. The functions $f_p(x)$ and $g(z)$ are even and we use here reversion as in Section 3.2.2 above. In Figures 2 and 3, we similarly provide plots of $g(z)$ and its inverse $g^{-1}(w)$ obtained by reversion. We

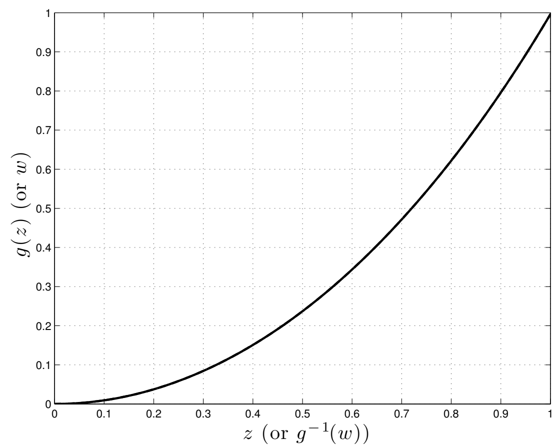

Figure 2: Plots of $g(z)$ and $g^{-1}(w)$ when $f_p(x)$ is given by (21) for $\sigma = 1$. The two curves overlap.

plot only $g(z)$ for $z \in [0, 1]$ and its inverse with positive values. We also consider two choices of $\sigma = 1$ and $\sigma = 1.5$ for illustration. For $\sigma = 1$, $g^{-1}(w)$ provides an inverse on the whole interval $R_{0,g} = [0, 1]$. But for $\sigma = 1.5$, as with χ_1^2 marginals above, $g^{-1}(w)$ provides an inverse on a smaller domain $R_g = [0, \eta^*]$, where η^* is about 0.36.

3.2.4. Reversion versus direct inversion

A number of questions arise naturally in view of numerical illustrations and general considerations above. Why should reversion be used in the first place? Why should it be used if it only leads to the inverse of $g(z)$ on a smaller domain R_g of the desired full domain $R_{0,g}$ for the inverse? What can be done if inversion is needed for $R_{0,g} \setminus R_g$? Which of the domains, R_g or $R_{0,g} \setminus R_g$ is more relevant?

Regarding these questions, reversion provides a convenient and numerically explicit way to define the inverse of the function $g(z)$. Another, simple-minded possibility would be to solve $g(z) = w$ numerically for a fixed w to obtain an inverse $z = g^{-1}(w)$. This is, however, numerically cumbersome if inversion is needed for many values of w. This is the case when long non-Gaussian series need to be synthesized and for which $z = r^X_{p,p'}[n] (= \rho^X_{p,p'}[n])$ need to be com-

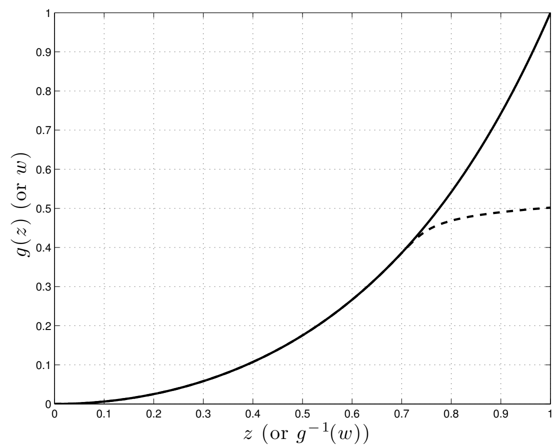

Figure 3: Plots of $g(z)$ and $g^{-1}(w)$ when $f_p(x)$ is given by (21) for $\sigma = 1.5$. The curve $(g^{-1}(w), w)$ (broken line) overlaps with the curve $(z, g(z))$ until $w \approx 0.36$.

puted through inversion from $w = \rho^Y_{p,p'}[n]$ for $n = 1, \ldots, N$ with large N. The reversion allows computing these $r^X_{p,p'}[n]$ in an efficient manner. Note also that most of $r^X_{p,p'}[n]$ are expected to be small (close to zero) and hence to fall in the domain R_g where $g^{-1}(w)$ obtained through reversion coincides with the inverse of g. In this sense, the domain R_g is much more relevant than $R_{0,g} \setminus R_g$. If inversion is needed on $R_{0,g} \setminus R_g$, it could be achieved by solving $g(z) = w$ numerically. Again, this is expected to be needed only for a small number of values of $w = \rho^Y_{p,p'}[n]$.

4. Approximations of Targeted Covariances (Through Circulant Embedding)

As mentioned in Section 2.4, starting from a valid positive definite covariance $r^Y_{p,p'}$ and a set of transforms f_p, there is no way to guarantee that $r^X_{p,p'}$ obtained through inversion constitutes a valid positive definite covariance structure. This section further addresses this issue by proposing a practical procedure that ensures positive definiteness of $r^X_{p,p'}$, at the price of yielding series $\tilde{Y}_p$ whose covariance $\tilde{r}^Y_{p,p'}$ constitutes an approximation of the targeted covariance $r^Y_{p,p'}$. This section first details a procedure, based on Circulant Matrix Embedding, for the synthesis of multivariate Gaussian time series (Section 4.1). Then, it details the approximation algorithm (Section 4.2.1), discusses the quality of the obtained approximation (Section 4.2.2) and addresses a form of *optimality* of this approximation (Section 4.2.3). Finally, the procedure is illustrated on an example (Section 4.2.4) and briefly discussed in respect to other available methods (Section 4.2.5).

Let us emphasize again that this procedure is new and original, even for the univariate case, $P = 1$.

4.1. Synthesis of multivariate stationary Gaussian series

Before describing the proposed procedure, we need to recall the key steps of the multivariate stationary Gaussian series synthesis procedure in [32, 33], using circulant matrix embedding.

For the synthesis of N samples $X[n]$, $n = 0, \ldots, N-1$, consider $N \times 1$ vectors $X_p = (X_p[0], \ldots, X_p[N-1])^T, p = 1, \ldots, P$, of component series of X. Let

$$\Sigma_{p,p'} = EX_p X_{p'}^T = (r_{p,p'}^X[n'-n])_{1 \le n,n' \le N}$$

which completely characterize the covariance structure of $X[n]$, $n = 0, \ldots, N-1$. The synthesis algorithm is based on circulant embedding of the matrices $\Sigma_{p,p'}$.

For $1 \le p \le p' \le P$, let $\widetilde{\Sigma}_{p,p'}$ be a circulant matrix with the first row

$$(r_{p,p'}^X[0] \cdots r_{p,p'}^X[N-1] \; r_{p,p'}^X[N] \; r_{p,p'}^X[-N+1] \cdots r_{p,p'}^X[-1]).$$

For $1 \le p' < p \le P$, set $\widetilde{\Sigma}_{p,p'} = \widetilde{\Sigma}_{p',p}^T$, which is also circulant. In all the cases, $\widetilde{\Sigma}_{p,p'}$ contain $\Sigma_{p,p'}$ in their upper-left corner. Another simpler embedding is described in Appendix A for the case where the covariance matrices $\Sigma_{p,p'}$ are symmetric.

Because the discrete Fourier basis diagonalizes circulant matrices, one can write

$$\widetilde{\Sigma}_{p,p'} = F^* \Lambda_{p,p'} F, \tag{25}$$

where $\Lambda_{p,p'} = diag(\lambda_{p,p'}[0], \ldots, \lambda_{p,p'}[2N-1])$, * indicates the conjugate transpose, and F^* is the $2N \times 2N$ Fourier matrix with k-th column

$$e_k = \frac{1}{\sqrt{2N}} \left(1, e^{-\imath \frac{2\pi k}{2N}}, \ldots, e^{-\imath \frac{2\pi k(2N-1)}{2N}}\right)^T.$$

In practice, the eigenvalues $\lambda_{p,p'}[0], \ldots, \lambda_{p,p'}[2N-1]$ are computed in an efficient manner by applying the FFT to the first row of $\widetilde{\Sigma}_{p,p'}$.

The next step in the construction is another formulation of the one given in [33] (see Appendix A for a brief review) and written in the spirit of [32]. For $p = 1, \ldots, P$, set

$$\widetilde{X}_p = F^* U_p, \tag{26}$$

where $U_p = (U_p[0], \ldots, U_p[2N-1])^T$ are specially chosen, complex-valued vectors. Let $\underline{U}_m = (U_1[m], \ldots, U_P[m])^T$, $m = 0, \ldots, 2N-1$, and

$$G_m = (2\Lambda_{p,p'}[m,m])_{1 \le p,p' \le P}, \quad m = 0, \ldots, 2N-1. \tag{27}$$

Note that, by the choice $\widetilde{\Sigma}_{p,p'} = \widetilde{\Sigma}^T_{p',p}$ above and using the decomposition (25), we have $\Lambda_{p,p'} = \Lambda^*_{p',p}$ for $p' < p$. This implies, in particular, that matrices G_m are always Hermitian symmetric.

Suppose now that the eigenvalues of G_m are nonnegative (G_m, being Hermitian symmetric, has real eigenvalues in general). As discussed extensively in [33], this is a mild assumption: it always holds for large enough N (and strictly speaking, at least for the so-called short-range dependent series) and it holds for any N for most time series models of interest.

If G_m has nonnegative eigenvalues, it admits the Cholesky factorization

$$G_m = B_m B_m^*.$$

Then $\underline{U}_m$ (or U_p) can be defined as

$$\underline{U}_m = B_m W_m,$$

where $W_m = W_m^0 + \imath W_m^1$ with two independent vectors W_m^0 and W_m^1 being $\mathcal{N}(0, I_P/2)$. This construction is carried out independently across $m =$

$0, \ldots, 2N-1$. Finally, the desired vectors X_p can be chosen as the first N entries of the real (or imaginary) parts of the vectors $\widetilde{X}_p$ given by (26). Note that if all the matrices G_m are nonnegative definite, this Gaussian synthesis method is exact. The total computational complexity for the procedure is only $O(P^2 N \log N)$ (see [33] for more detailed discussion about computational issues).

4.2. Approximation procedure

4.2.1. Algorithm

The procedure is simple to implement, natural when using circulant embedding, and can be summarized through the following steps:

Algorithm for synthesizing underlying Gaussian series.

(1) For given transformations f_p in (2) matching desired marginal distributions, consider a relationship (6) between targeted and underlying covariances. By inverting relation (6), solve for $r^X_{p,p'}[n]$ in terms of targeted $r^Y_{p,p'}[n]$.

(2) Even though $r^X_{p,p'}[n]$ above does not necessarily define a valid covariance structure, proceed with circulant embedding synthesis as described in Section 4.1, in the formulation following (26). Compute in particular, Hermitian symmetric matrices G_m in (27).

(3) Factorize G_m through Schur decomposition as

$$G_m = O_m S_m O_m^*, \tag{28}$$

where $S_m = diag(s_m[1], \ldots, s_m[P])$ is diagonal, and O_m is unitary. Define $\tilde{S}_m = diag(\tilde{s}_m[1], \ldots, \tilde{s}_m[P])$, where

$$\tilde{s}_m[p] = \begin{cases} s_m[p], & \text{if } s_m[p] \geq 0, \\ 0, & \text{if } s_m[p] < 0, \end{cases} \tag{29}$$

and consider nonnegative definite matrices

$$\tilde{G}_m = O_m \tilde{S}_m O_m^*. \tag{30}$$

(4) Finally, use nonnegative definite matrices $\tilde{G}_m$ instead of G_m in the circulant embedding method (Section 4.1) to generate underlying Gaussian series. Denote the approximating covariances by $\widetilde{r}^{Y}_{p,p'}[n]$ and the approximating series by $\widetilde{Y}[n]$.

The steps above need some further explanation. In Step (1), inversion of (6) is carried out explicitly or numerically as discussed in Section 3. We assume, in particular, that the targeted covariances $r^{Y}_{p,p'}[n]$ fall in the range where this relation can actually be inverted. How to determine this in practice is also discussed in Section 3. Regarding Step (2), we emphasize again that $r^{X}_{p,p'}[n]$ from Step (1) do not necessarily define a valid covariance structure, and that despite this fact, the matrices G_m will always be Hermitian symmetric. Step (3) above just makes the matrices G_m nonnegative definite by setting their eigenvalues to be nonnegative.

4.2.2. Quality of the approximation

When $r^{X}_{p,p'}[n]$ actually defines a valid covariance structure, it is known from [33] that under mild assumptions, the eigenvalues of G_m in (27) will be nonnegative and hence that the series X will be generated through the circulant embedding method. In this case, the synthesized series $\widetilde{Y}[n]$ will exactly correspond to the desired series $Y[n]$, i.e., their marginal and covariance will exactly match the targeted ones.

When $r^{X}_{p,p'}[n]$ does not define a valid covariance structure, it is similarly expected that some eigenvalues of G_m will be negative, and hence that only an approximation to the targeted covariance will be obtained. However, the approximating covariances $\widetilde{r}^{Y}_{p,p'}[n]$ can be calculated numerically using the covariance relationship (6), since the matrices $\tilde{G}_m$ together with decomposition (25) can give us the covariance of the synthesized multivariate Gaussian series. Moreover, when the targeted covariance is approximated, this information can be conveyed explicitly to practitioners, together with the actually synthetized covariance $\widetilde{r}^{Y}_{p,p'}[n]$. The approximating covariance is, hence, made available to practitioners, as illustrated in Figure 4 (see Section 4.2.4 below). Its quality can be checked a posteriori by practitioners using any distance that matches the proper needs of the applications being dealt with. Full knowledge of the resulting covariance is obviously very important for controlled numerical experiments making this a valuable feature of the method.

Finally, let us emphasize again that when the eigenvalues of G_m are nonnegative, the marginal distributions always match the targeted ones. Otherwise, one can always perfectly match the targeted marginals by rescaling the Gaussian series X appropriately; e.g., in the case of standard transforms as in (9), the variance of the series X needs to be equal to one. Note that a similar solution would be necessary to ensure exact match of marginals in the synthesis procedures reviewed in Section 1.3.

4.2.3. Optimality

We now turn to a form of optimality for the obtained approximation of the targeted covariance. The argument is akin to that found in [17], where a related problem is considered in the univariate case.

Suppose the covariance structure $R_Y[n]$ of Y is given. Let $T : \mathbb{R}^{P\times P} \to \mathbb{R}^{P\times P}$ be the transformation relating $R_X[n]$ and $R_Y[n]$ as in Eq. (6), that is, $R_Y[n] = T(R_X[n])$. (Here, $R_X[n]$ is not necessarily a valid covariance structure.) Then, proceeding as in [17], Section IV, consider

$$\begin{aligned} I &:= \min \sum_n \|R_Y[n] - R_{app,Y}[n]\|^2_{W_n} \\ &= \min \sum_n \|T(R_X[n]) - T(R_{app,X}[n])\|^2_{W_n}, \end{aligned}$$

where $\|x\|^2_W = vec(x)^T W vec(x)$ with a positive definite matrix W, W_n are positive definite weight matrices, and the minimum is over all covariance structures $R_{app,X}[n]$. We will argue informally that the minimum above is obtained by taking the series $X[n]$ as in the approximation algorithm of Section 4.2.1.

One expects in practice that $R_X[n] \approx R_{app,X}[n]$ for the optimal $R_{app,X}[n]$. Then, using the approximation $vec(T(R_X[n])-T(R_{app,X}[n])) \approx T'(R_X)vec(R_X[n]-$

$R_{app,X}[n]$) with $T'(R_X) = diag(vec(T'(R_X[n])))$ and T' denoting the derivative of the function T, we get that

$$I \approx \min \sum_n \|R_X[n] - R_{app,X}[n]\|^2_{\overline{W}_n},$$

where $\overline{W}_n = T'(R_X)^T W_n T'(R_X)$. Supposing now that W_n are such that $\overline{W} \approx I_P$, we further get that

$$I \approx \min \sum_n \|R_X[n] - R_{app,X}[n]\|^2,$$

where $\|\cdot\|$ stands for the Frobenius norm, i.e., $\|A\|^2 = \sum_{p=1}^P \sum_{p'=1}^P |a_{p,p'}|^2$ for a matrix $A = (a_{p,p'})_{1 \le p,p' \le P}$. By using Parseval's relation, we obtain that

$$I \approx \min \frac{1}{2\pi} \int_{-\pi}^{\pi} \|f_X(\lambda) - f_{app,X}(\lambda)\|^2 d\lambda,$$

where, for example, $f_X(\lambda) = \sum_{n=-\infty}^{\infty} R_X[n] e^{-i\lambda n}$ is the "spectral density" matrix of $R_X[n]$. (Note that $f_X(\lambda)$ here is not necessarily nonnegative definite.) For large $2N$, we further get that

$$I \approx \min \frac{1}{N} \sum_{m=0}^{2N-1} \left\| f_X\left(\frac{2\pi m}{2N}\right) - f_{app,X}\left(\frac{2\pi m}{2N}\right) \right\|^2.$$

As in the proof of Theorem 3.1 of [33], one expects that $2f_X(2\pi m/2N) \approx G_m$, where G_m is defined in (27). Then,

$$I \approx \min \frac{2}{N} \sum_{m=0}^{2N-1} \|G_m - G_{app,m}\|^2.$$

The optimal $G_{app,m}$ for the right-hand side above is now $\tilde{G}_m$ defined through (30). This follows from the next result which can be proved easily using the fact that the Frobenius norm is invariant under multiplication by unitary matrices.

Lemma 4.1. *Let G be a $P \times P$ Hermitian symmetric matrix such that $G = OSO^*$, where $S = diag(s[1], \ldots, s[P])$ and O is unitary. Then,*

$$\min \|G - G_{app}\| = \|G - \tilde{G}\|,$$

where the minimum is over nonnegative definite matrices G_{app}, and $\tilde{G} = O\tilde{S}O^*$ with $\tilde{S} = diag(\tilde{s}[1], \ldots, \tilde{s}[P])$ and $\tilde{s}[p] = \max\{s[p], 0\}$.

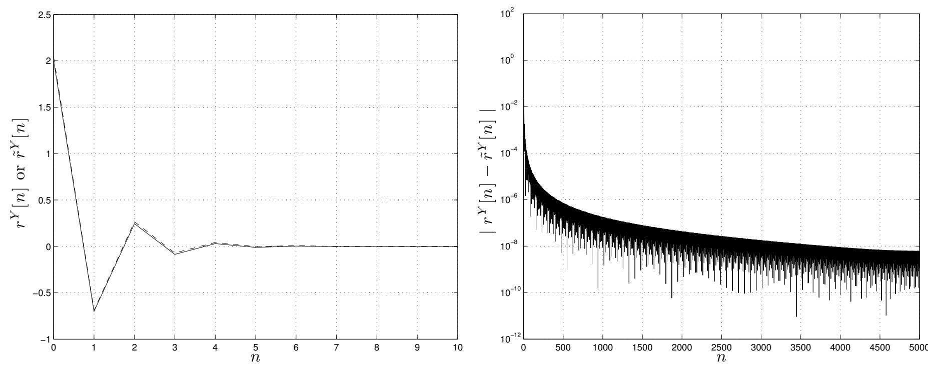

Figure 4: Plots of $r^Y[n]$ and its approximation $\widetilde{r}^Y[n]$ (broken line) for AR(1) covariance and the transformation (18).

4.2.4. Numerical illustration

We briefly illustrate the algorithm above on a univariate ($P = 1$) example. Suppose a series $Y[n]$ with χ_1^2 marginal and covariance function $r^Y[n] = 2\phi^{|n|}$, $n \geq 0$, $|\phi| < 1$, of AR(1) series is targeted. Suppose also that negative correlations are sought, that is, $-1 < \phi < 0$. Consider the transformation (18) for which negative correlations are possible. Take $\phi = -0.35$ so that $\rho^Y[n] = r^Y[n]/r^Y[0]$, $n \geq 1$, fall in the domain $R_{0,g} = [g(-1), 1]$ where inversion is possible and $g(-1)$ is about -0.44 as discussed in Section 3.2.3, Example (a). Take the sample size $N = 5000$. In the top plot of Figure 4, for the first few lags n, we plot the targeted covariance $r^Y[n]$ and its approximation $\widetilde{r}^Y[n]$ obtained from the algorithm described in Section 4.2.1. In the bottom plot of the figure, we present the error of the approximation for a larger range of lags n.

4.2.5. Further discussion on the proposed and other methods

We first briefly compare the performance of our method to a one-dimensional version of the synthesis method described in [29]. The latter "spectral representation" method is very popular in a more applied literature, and was briefly sketched in Section 1.3.2 (see also [24]). The setting is as in Section 4.2.4. To enable comparison in the covariance (time) domain, we slightly modify the algorithm in [29] and use circulant embedding for exact synthesis of the underlying Gaussian series $X[n]$. Define the discrete spectrum $S[k] = \sum_{n=0}^{2N-1} \widetilde{r}[n]e^{-\imath 2\pi nk/(2N)}$, where $\widetilde{r}[n] = r[n]$ for $n = 0, \dots, N$, and $\widetilde{r}[n] = r[2N - n]$, $n = N+1, \dots, 2N-1$. Let S_X and S_Y be the discrete spectrums corresponding to X and Y, respectively. The one-dimensional version of the algorithm in [29] can be described as follows (using discrete spectrums):

1. Calculate S_Y from the embedding $\widetilde{r}^Y$ of r^Y. Set $S_X^{(1)} = S_Y$ and $i = 1$.

2. Generate Gaussian series $X^{(i)}[n]$, $n = 0, \dots, N-1$, with the spectrum $S_X^{(i)}$ and transform to get non-Gaussian series $Y^{(i)}[n] = f(X^{(i)}[n])$.

3. Estimate the spectral density of the generated non-Gaussian series based on the realization $Y^{(i)}[n]$ from Step 2. Denote the estimated spectrum $S_Y^{(i)}$.

4. If the relative error $\epsilon^{(i)} := \|S_Y - S_Y^{(i)}\|_{\ell_1}/\|S_Y\|_{\ell_1}$ is smaller than a threshold, stop. Otherwise, update the Gaussian spectrum according to

$$S_X^{(i+1)} = S_X^{(i)} \frac{S_Y}{S_Y^{(i)}}, \tag{31}$$

set $i = i + 1$, and return to Step 2.

In the procedure described in [29], the spectral estimation in Step 3 is based on the modulus of the DFT of $Y^{(i)}$ which will typically result in a "spiky" (irregular) estimate. Here we use instead a standard Welch's spectral estimation method, dividing the data into 20 windowed segments (using a Hamming window) with 50% overlap. This smoothens the spectral estimate and works in favor of the method due to the smooth spectrum of the AR(1) series.

Figure 5 (bottom) shows the behavior of the relative error $\epsilon^{(i)}$ in Step 4 for the first 100 iterations. The minimum value, about 0.07, is reached after 10 iterations. In comparison, the relative error of our proposed method is 0.02. The two top plots of Figure 5 compare the targeted covariance r^Y with the covariance $r^{Y,i}$ of the synthesized non-Gaussian series $Y^{(i)}$ corresponding to the iteration where $\epsilon^{(i)}$ reaches minimum (among the 100 iterations). The covariance $r^{Y,i}$

of $Y^{(i)}$ was calculated numerically by applying the IDFT to the spectrum $S_X^{(i)}$ to get the actual covariance of the underlying synthesized Gaussian series $X^{(i)}$, and using the covariance relationship (6) (hence no estimation required). When comparing these plots with those in Figure 4, we see that for large lags (where the targeted covariance is essentially zero), the absolute difference from the targeted covariance is of the same order. On the other hand, for AR models, the covariance is essentially determined by its values at small lags. As Figures 4 and 5 show, this is where the performance of the iterative (estimation-based) scheme is far inferior to the proposed method. It is worth mentioning that the updating rule (31) in the iterative method does not seem applicable to all nonstandard transformations f (this can, for example, be checked numerically for the targeted covariance $r^Y[n] = 2\phi^{|n|}$ with ϕ positive and $f(x) = x^2$).

We conclude this discussion by stressing the qualitative differences and advantages of our method to other available methods. As indicated in Section 4.2.3, our method is closest in spirit to the spectral truncation method of [17]. The performance of the latter method on the AR(1) example of Section 4.2.4 (not reported here) is comparable to our method. However, based on a popular approach of circulant matrix embedding, the proposed method is particularly easy to implement, is computationally efficient, and extends to the multivariate case. Two other distinct features of the proposed method are the following. First, the method is based on covariances, rather than spectral densities as, for example, in "spectral representation" methods (see Section 1.3). Second, in the Gaussian case, the proposed method is *exactly* that of the synthesis of Gaussian stationary series by using circulant matrix embedding. The latter method is the most popular and preferred method for the synthesis of such series (see, e.g., [33] in the multivariate context).

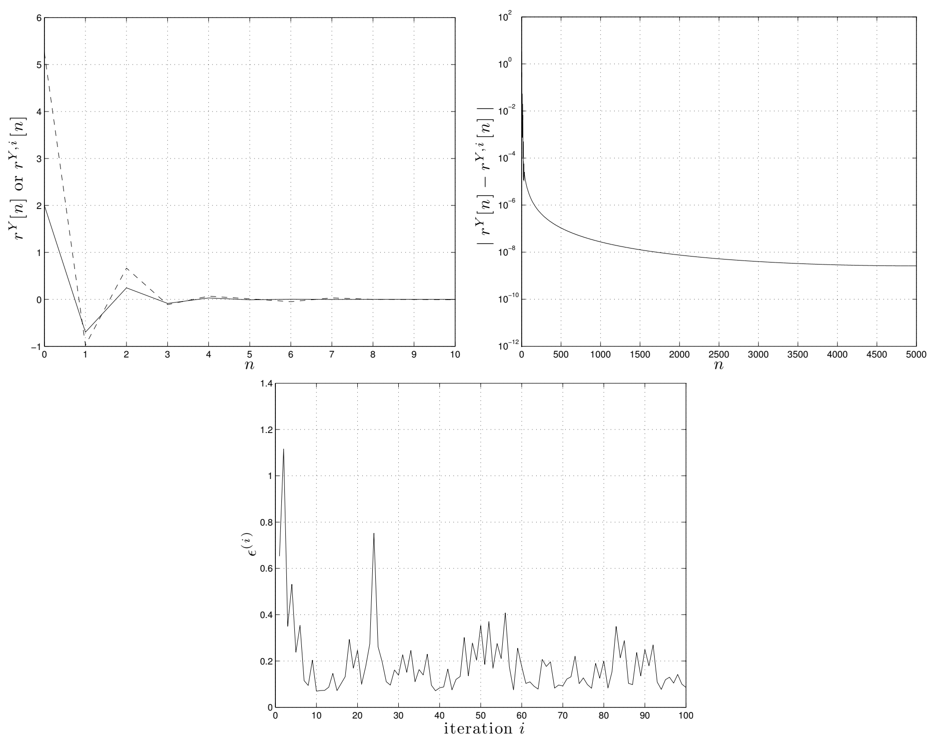

Figure 5: Plots of $r^Y[n]$ and $r^{Y,i}[n]$, the covariance of $Y^{(i)}$ where $\epsilon^{(i)}$ is minimum, (broken line) for the example of Section 4.2.5. Bottom plot shows relative error $\epsilon^{(i)}$ for subsequent iterations of the algorithm.

5. Component-Wise Transformations of Independent Copies of Gaussian Multivariate Series

5.1. Framework

In the previous sections, a non-Gaussian series Y was obtained by transformation of a single auxiliary Gaussian series X. Here we briefly address the useful case where $Y[n] = (Y_1, \dots, Y_P[n])^T$ is taken of the form

$$Y_p[n] = f_p(X_p^{(1)}[n], \dots, X_p^{(K)}[n]), \tag{32}$$

where $X^{(k)}[n] = (X_1^{(k)}[n], \dots, X_P^{(k)}[n])^T$, $k = 1, \dots, K$, are i.i.d. copies of a stationary Gaussian series $X[n]$ (which is as in the previous sections). Variants of this case were considered in [21, 22, 8].

A reason of interest in constructions (32) is that many marginals of practical use can be matched through such explicit constructions when $K \geq 2$. In particular, this can lead to explicit (direct and inverted) relationships between the covariances of Y and X, thus, numerical methods are not needed to deal with these relationships. A number of examples are given in Section 5.2 below. Though the main interest is in explicit relationships, the covariances of Y and X can also be related in terms of Hermite polynomials akin to (6). This and some other issues are discussed in Sections 5.3 and 5.4.

5.2. Transformations and exact inversion for some classical marginals

This section only considers the case where $f_1 = \cdots = f_P$.

5.2.1. Linear combinations of χ_1^2 variables

Consider the transform

$$f_p(x) = f_p(x_1, \ldots, x_K) = \sum_{k=1}^{K} b_k x_k^2, \quad x \in \mathbb{R}^K,$$

where $\{b_k, k = 1, \ldots, K\}$ is a deterministic sequence. Many important distributions result from such linear combinations of independent χ_1^2 random variables; see Table 1. Using the decomposition $x_k^2 = H_0(x_k) + H_2(x_k)$, as in Section 2.5, and the fact that $X^{(k)}$ and $X^{(k')}$ are independent for $k \neq k'$, one gets

$$r^Y_{p,p'}[n] = \sum_{k=1}^{K} 2b_k^2 (r^X_{p,p'}[n])^2.$$

This relationship can be inverted if and only if $r^Y_{p,p'}[n] \geq 0$, in which case

$$r^X_{p,p'}[n] = \sqrt{\frac{r^Y_{p,p'}[n]}{2\sum_{k=1}^{K} b_k^2}}, \tag{33}$$

when enforcing $r^X_{p,p'}[n]$ to be nonnegative as well.

5.2.2. Products of Log-χ^2 variables

Consider the transform

$$f_p(x) = f_p(x_1, \ldots, x_K) = \exp\left(\sum_{k=1}^{K} b_k x_k^2\right), \tag{34}$$

Table 1: Important marginals resulting from $f_p(x) = \sum_{k=1}^{K} b_k x_k^2$.

Marginal Type	Parameter Values
Exponential distribution with mean $a > 0$	$K = 2$, $b_1 = b_2 = a/2$
Laplace$(0, a)$ distribution with zero mean and variance $2a^2$	$K = 4$, $b_1 = b_3 = a/2$, $b_2 = b_4 = -a/2$
Chi-square distribution with ν degrees of freedom	$K = \nu$, $b_k = 1, \forall k$
Erlang(α, β) distribution, i.e., sum of α exponential variables, each with mean $\beta > 0$	$K = 2\alpha$, $b_k = \beta/2, \forall k$

with $\{b_k, k = 1, \ldots, K\}$ being deterministic. Then $\log Y_p[n]$ is a linear combination of independent χ_1^2 variables. Some important examples of marginals resulting from this type of transformation are listed in Table 2. Note that to get a Pareto distribution with a general minimum $\beta > 0$ (i.e., $P(Y_p < y) = 1-(\beta/y)^\alpha$ for $y \geq \beta$), one can take $f_p(x) = \beta \exp((x_1^2 + x_2^2)/(2\alpha))$ and replace the covariance sequences $r_{p,p'}^Y$ in the formulas below by $r_{p,p'}^Y/\beta^2$.

Table 2: Important marginals resulting from $f_p(x) = \exp(\sum_{k=1}^{K} b_k x_k^2)$.

Marginal Type	Parameter Values
Pareto distibution with minimum 1 and index $\alpha > 0$	$K = 2$, $b_1 = b_2 = \frac{1}{2\alpha}$
Uniform$(0, 1)$ distibution	$K = 2$, $b_1 = b_2 = -1/2$

To find a closed form for the relationship of the covariances we will use the classical formula for the moment-generating function of χ_1^2 variable

$$Ee^{tX^2} = (1 - 2t)^{-1/2}, \quad t < 1/2, \quad \text{where } X \sim \mathcal{N}(0, 1), \tag{35}$$

and the following lemma (proved in Appendix C):

Lemma 5.1. *Assume* $X_j \sim \mathcal{N}(0, \sigma_j^2)$, $j = 1, 2$, *and* $EX_1X_2 = \sigma_{12}$, *with* $\sigma_{12}^2 < \sigma_1^2\sigma_2^2$. *Then*

$$Ee^{t(X_1^2+X_2^2)} = \left((1 - 2t\sigma_1^2)(1 - 2t\sigma_2^2) - 4t^2\sigma_{12}^2\right)^{-1/2} \tag{36}$$

provided that $t\lambda < 1/2$, *where*

$$\lambda = \frac{\sigma_1^2 + \sigma_2^2}{2} + \frac{1}{2}\sqrt{(\sigma_1^2 - \sigma_2^2)^2 + 4\sigma_{12}^2}.$$

Using Lemma 5.1 with $\sigma_1^2 = \sigma_2^2 = r_{p,p}^X[0] = 1$ and $\sigma_{12} = r_{p,p'}^X[n]$, one gets

$$\begin{aligned} EY_p[n']Y_{p'}[n'+n] &= \prod_{k=1}^{K} Ee^{b_k(X_p^{(k)}[n']^2 + X_{p'}^{(k)}[n'+n]^2)} \\ &= \prod_{k=1}^{K} \left((1 - 2b_k)^2 - 4b_k^2 r_{p,p'}^X[n]^2\right)^{-1/2}, \end{aligned}$$

provided that $b_k < 1/4$, $k = 1, \ldots, K$; the conditions of the lemma become $b_k < (2(1 + |r_{p,p'}^X[n]|))^{-1}$, where the right-hand side takes the minimum value $(2(1 + |r_{p,p}^X[0]|))^{-1} = 1/4$. The mean of $Y_p[n]$ is

$$EY_p[n] = \prod_{k=1}^{K} Ee^{b_k(X_p^{(k)}[n])^2} = \prod_{k=1}^{K} (1 - 2b_k)^{-1/2}.$$

(The condition $b_k < 1/2$ for this formula to hold is weaker than the earlier condition.) These calculations give us a closed formula for $r_{p,p'}^Y[n] = EY_p[n']Y_{p'}[n'+n] - EY_p[n']EY_{p'}[n'+n]$ in terms of $r_{p,p'}^X[n]$.

Consider the special case when $b = b_1 = \cdots = b_K$ and write $\mu := EY_p[n] = (1 - 2b)^{-K/2}$. Then

$$r_{p,p'}^Y[n] = \left(\mu^{-4/K} - 4b^2 r_{p,p'}^X[n]^2\right)^{-K/2} - \mu^2, \tag{37}$$

which can be rewritten as

$$r_{p,p'}^X[n]^2 = \frac{1}{4b^2}\left(\mu^{-4/K} - \left(r_{p,p'}^Y[n] + \mu^2\right)^{-2/K}\right). \tag{38}$$

Note that (37) can be inverted only if $r_{p,p'}^Y[n] \geq 0$ (this follows also from the fact that f_p is even, see Section 2.4). These formulas are useful when generating series with the marginals listed in Table 2.

5.3. Relationship between covariances via Hermite expansions

In the more general framework (32), the covariances $r^Y_{p,p'}$ and $r^X_{p,p'}$ can also be related by using multivariate Hermite expansions,

$$f_p(x) = \sum_{l_1,\ldots,l_K=0}^{\infty} c^{(p)}_{l_1,\ldots,l_K} \prod_{k=1}^{K} H_{l_k}(x_k), \tag{39}$$

where for $l_1, \ldots, l_K \geq 0$,

$$c^{(p)}_{l_1,\ldots,l_K} = E\left[f_p(X^{(1)}_p[n], \ldots, X^{(K)}_p[n]) \prod_{k=1}^{K} H_{l_k}(X^{(k)}_p) \right]$$

(see [39]). As shown in Appendix D, this leads to

$$r^Y_{p,p'}[n] = \sum_{m=1}^{\infty} a^{(p,p')}_m (r^X_{p,p'}[n])^m, \tag{40}$$

where

$$a^{(p,p')}_m = \sum_{l_1+\cdots+l_K=m} c^{(p)}_{l_1,\ldots,l_K} c^{(p')}_{l_1,\ldots,l_K} \prod_{k=1}^{K} l_k!. \tag{41}$$

5.4. Matching targeted covariances and synthesis algorithm

All the issues addressed in Sections 3 and 4 can be revisited under the more general framework (32). For example, the relationship of the covariances $r^Y_{p,p'}[n]$ in terms of $r^X_{p,p'}[n]$ can be inverted either explicitly as in (33), (38), or using reversion of the series (40). For targeted covariances $r^Y_{p,p'}[n]$, we again expect that inversion does not necessarily lead to a valid covariance structure $r^X_{p,p'}[n]$. The latter could be made valid through the algorithm described in Section 4.2.1 and is optimal in the sense of Section 4.2.3.

6. Discussion on Applications to Internet Traffic Modeling

It is now widely accepted that aggregated traffic time series, formed counting the number of Pkts or Bts in a time interval $[n\Delta, (n+1)\Delta)$ where Δ is an aggregation level, are characterized by two major statistical features: long-range dependence and significant departures from Gaussianity [40, 7]. In particular, it has been shown [41] that such time series are well modeled by $\Gamma_{\alpha,\beta}$ marginal distributions and FARIMA(p, d, q) covariance structures.

In Internet traffic engineering (maintenance, security, ...), the actual deployment of security strategies and algorithms in real networks is a challenging task involving severe security protocols and monetary issues. Therefore, the possibility of testing the efficiency, performance and robustness of various algorithms on synthetic data that mimics as closely as possible real traffic is crucial, hence the great interest of versatile and efficient synthesis procedures. Let us detail an example showing the importance of varying f.

It is regarded as essential to correctly estimate the long-range dependence parameter, referred to as d. While estimation of d for Gaussian time series is now well-documented ([42] and references therein), this is far less the case for non-Gaussian time series. Estimation performance can hence be tested on synthetic time series with $\Gamma_{\alpha,\beta}$ marginal distributions and FARIMA(p, d, q) covariance. However, the analysis and examples presented in Section 2 clearly indicate that for a given joint choice of marginal and covariance, there are multiple ways to synthesize time series. More precisely, $\Gamma_{\alpha,\beta}$ distribution could be reached both through the transformation $f(x) = F^{-1}_{\Gamma_{\alpha,\beta}}(\Phi(x))$ and $f(x) = F^{-1}_{\Gamma_{\alpha,\beta}}(2(\Phi(|x|) - 1/2)$. On the condition that the desired covariance can be attained by both transformations, the marginal distribution and covariance of the two non-Gaussian series will be identical in both cases, yet their full joint distributions will be different. For such two series, the analysis reported in [43] clearly reveals that while the estimation of d appears almost as *easy* as in the Gaussian case for the choice $f(x) = F^{-1}_{\Gamma_{\alpha,\beta}}(\Phi(x))$, this is no longer the case for the choice $f(x) = F^{-1}_{\Gamma_{\alpha,\beta}}(2(\Phi(|x|) - 1/2)$. At this stage, it is an open question whether Internet traffic time series are better modeled by either choice of f. To investigate this question, it is of major practical importance to have at disposition tools that offer the possibility of varying the choice of f to design and explore a large variety of different non-Gaussian times series that yet have the same marginal distribution and covariance function.

As an illustration, using the tools proposed in this work, we synthesized

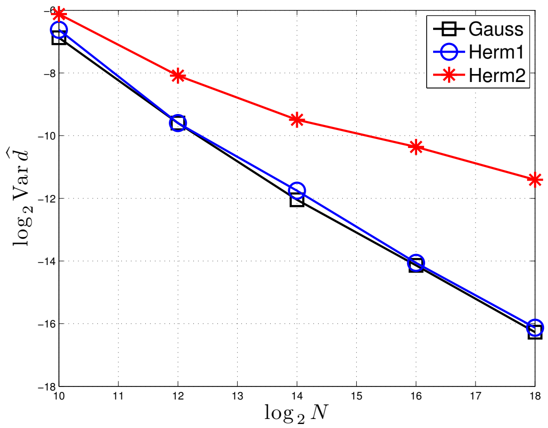

Figure 6: Log of the variance of $\widehat{d}$, the wavelet-based estimator of d, versus the log of series length, for time series with the same FARIMA$(0, 0.3, 0)$ covariance, for a Gaussian times series, time series with a $\Gamma_{3,2}$ marginal obtained from $f(x) = F^{-1}_{\Gamma_{3,2}}(\Phi(x))$ (Hermite rank 1) and time series with $\Gamma_{3,2}$ marginal obtained from $f(x) = F^{-1}_{\Gamma_{3,2}}(2(\Phi(|x|) - 1/2)$ (Hermite rank 2).

numerically 500 realizations of lengths $N = 2^J$, for $J = 10, 12, \ldots, 18$, of non-Gaussian times series with marginal distribution $\Gamma_{3,2}$ and FARIMA$(0, 0.3, 0)$ covariance, these parameters being chosen so as to match what is commonly observed on Internet traffic times series from the Japanese MAWI database [41, 44]. Two transformations,$f(x) = F^{-1}_{\Gamma_{3,2}}(\Phi(x))$ and $f(x) = F^{-1}_{\Gamma_{3,2}}(2(\Phi(|x|) - 1/2)$, are used and in both cases, marginal distributions and covariance functions are exactly the same. The parameter d is estimated using a classical wavelet-based method described in, for example, [7]. For both synthesized times series, the estimator of d is found unbiased (not shown here). The variances of the estimates of d are compared (as a function of N) in Figure 6, which shows clearly that for data synthesized using $f(x) = F^{-1}_{\Gamma_{3,2}}(\Phi(x))$, the variance decreases with N as quickly as in the Gaussian case, while the decrease is far more slow for data synthesized using $f(x) = F^{-1}_{\Gamma_{3,2}}(2(\Phi(|x|)-1/2))$, a fact with major practical consequences. (The Hermite rank referred to in the caption of Figure 6 is the index of the first non-zero coefficient in the Hermite expansion.)

This example illustrates that the versatility in varying f to synthesize times series with the same marginal distribution and covariance function is crucial, both when addressing theoretical issues (such as estimation) or to meet practical requirements (fast synthesis of very long time series).

7. Conclusions and Further Directions

We have presented a novel practical methodology for synthesizing non-Gaussian series where the underlying algorithm is simple and very fast. The method can be applied for general transforms f_p, although we also provide a set of important and practical examples where one has simple analytic formulae for the relationship between covariances R_Y and R_X and its inversion. In cases where the targeted covariance needs to be approximated to impose nonnegative definiteness of R_X, the resulting covariance approximates it in an optimal sense.

Since non-Gaussian series are not fully determined by their mean and covariance, it remains to investigate the effect of different transforms f_p, giving the same marginal and covariance, on higher-order moments, e.g., $Cov(Y_p[k]^q, Y_{p'}[n+k]^q)$, $q \geq 2$. Indeed, we have emphasized that although different transforms f_p can give the same marginal distribution for $Y_p[n]$, these transforms can have different properties and restrictions. Would it be possible to extend the proposed methodology and gain control of higher-order moments in synthesis?

Another open theoretical issue to address would be to provide sufficient conditions for the inversion of the covariance relation to give nonnegative definite covariance R_X ([22] addresses this for the univariate case for some specific covariance models and marginals).

8. Software

MATLAB codes implementing the methods and demos are publicly available at `http://www.hermir.org`.

Acknowledgements

We thank the Associate Editor and referees for their many helpful comments.

Appendix A. Further details about the synthesis of Gaussian series

This section provides some additional information relating to the multivariate Gaussian synthesis procedure described in Section 4.1.

Appendix A.1. Circulant embedding for time-reversible series

As pointed out in [33], other circulant embeddings than the one provided in Section 4.1 could work. For example, a slightly simpler embedding is possible if the Gaussian series is time-reversible, which is equivalent to

$$R_X[n] = R_X[-n] = R_X[n]^T, \quad n \geq 1, \tag{A.1}$$

that is, covariance matrices being symmetric. In this case, one can take $\widetilde{\Sigma}_{p,p'}$ to be circulant matrices with the first row being

$$(r^X_{p,p'}[0] \cdots r^X_{p,p'}[N-1] \; r^X_{p,p'}[N-2] \cdots r^X_{p,p'}[1]).$$

Note that for this case, the size of $\widetilde{\Sigma}_{p,p'}$ becomes $(2N-2) \times (2N-2)$ which would require the construction in Section 4.1 to be changed accordingly.

Appendix A.2. Another construction for synthesis

The construction given in [33] replaces (26) by

$$\widetilde{X}_p = F^* \Lambda_{p,p}^{1/2} Z_p. \tag{A.2}$$

Here, $Z_p = (Z_p[0], \dots, Z_p[2N-1])^T$ are complex-valued, column vectors of size $2N$, specially chosen as follows. Instead of dealing with Z_p directly, set $\underline{Z}_m = (Z_1[m], \dots, Z_p[m])^T$, $m = 0, \dots, 2M-1$. Consider $2N \times 2N$ diagonal matrices

$$D_{p,p'} = 2\Lambda_{p,p}^{-1/2} \Lambda_{p,p'} \Lambda_{p',p'}^{-1/2}, \quad 1 \leq p, p' \leq P,$$

and $P \times P$ matrices

$$C_m = (D_{p,p'}[m,m])_{1 \leq p,p' \leq P}, \quad m = 0, \dots, 2N-1.$$

Factor out C_m as

$$C_m = A_m A_m^*,$$

and define $\underline{Z}_m$ (or Z_p) as

$$\underline{Z}_m = A_m W_m,$$

where W_m are the same as in Section 4.1. This construction is carried out independently across $m = 0, \dots, 2N-1$. One can easily check that this formulation is equivalent to the one given in Section 4.1 (where basically $U_p = \Lambda_{p,p}^{1/2} Z_p$).

Appendix B. Using Price theorem to derive covariance relationships

For demonstration purposes, we include simple applications of Price theorem to derive covariance relationships for the χ_1^2 and log-normal marginals in Section 2.5.

Appendix B.1. Price theorem for χ_1^2 marginals

Consider $f_p(x) = x^2$. Taking $k = 1$ in (8), one gets

$$\frac{\partial r_{p,p'}^Y[n]}{\partial r_{p,p'}^X[n]} = 4EX_p[0]X_{p'}[n] = 4r_{p,p'}^X[n].$$

The relation (17) follows by integrating and using the fact that $r_{p,p'}^X[n] = 0$ (that is, $X_p[0]$ and $X_{p'}[n]$ are independent) implies $r_{p,p'}^Y[n] = 0$.

Appendix B.2. Price theorem for log-normal marginals

Here we consider $f_p(x) = e^{\sigma x + \mu}$. Using $\partial f_p(x)/\partial x = \sigma f_p(x)$, Price theorem gives

$$\frac{\partial r_{p,p'}^Y[n]}{\partial r_{p,p'}^X[n]} = \sigma^2 EY_p[n]Y_{p'}[n] = \sigma^2(r_{p,p'}^Y[n] + e^{2\mu+\sigma^2}),$$

since $r_{p,p'}^Y[n] = EY_p[n]Y_{p'}[n] - E(Y_p[n])E(Y_{p'}[n])$ and $EY_p[n] = e^{\mu+\sigma^2/2}$. The solution to this standard ordinary differential equation is

$$r_{p,p'}^Y[n] = e^{2\mu+\sigma^2}(e^{\sigma^2 r_{p,p'}^X[n]} - 1),$$

where the constant term is determined by the fact that $r_{p,p'}^X[n] = 0$ implies $r_{p,p'}^Y[n] = 0$.

Appendix C. Proof of Lemma 5.1

Using matrix notation, write $X = (X_1\ X_2)^T$ and $EXX^T = Q\Lambda Q^T$, where Q is orthogonal and $\Lambda = diag(\lambda_+, \lambda_-)$. This allows one to write $X = Q\Lambda^{1/2}Z$, where $Z = (Z_1\ Z_2)^T$, Z_1 and Z_2 are i.i.d. $\mathcal{N}(0,1)$ variables, yielding $X_1^2 + X_2^2 = X^TX = \lambda_+ Z_1^2 + \lambda_- Z_2^2$. Then, by independence and using the formula for the moment-generating function of χ_1^2 variables, one gets $E\exp\big(t(X_1^2 + X_2^2)\big) = E\exp(t\lambda_+ Z_1^2) E\exp(t\lambda_- Z_2^2) = (1 - 2t\lambda_+)^{-1/2}(1 - 2t\lambda_-)^{-1/2}$, provided that $t\lambda_\pm < 1/2$; this condition is equivalent to $t\lambda_+ < 1/2$, since $\lambda_+ \geq \lambda_- > 0$. The result follows from elementary algebra.

Appendix D. Derivation of covariance relationship based on multivariate Hermite expansions

Here we derive relationship (40). By using (39) and since $X_p^{(k)}$ and $X_{p'}^{(k')}$ are independent for $k \neq k'$, one has

$$
\begin{aligned}
EY_p[0]Y_{p'}[n] &= E \sum_{l_1,\dots,l_K=0}^{\infty} c^{(p)}_{l_1,\dots,l_K} \prod_{k=1}^{K} H_{l_k}(X_p^{(k)}[0]) \sum_{m_1,\dots,m_K=0}^{\infty} c^{(p')}_{l_1,\dots,l_K} \prod_{k=1}^{K} H_{m_k}(X_{p'}^{(k)}[n]) \\
&= E \sum_{l_1,\dots,l_K=0}^{\infty} \sum_{m_1,\dots,m_K=0}^{\infty} c^{(p)}_{l_1,\dots,l_K} c^{(p')}_{l_1,\dots,l_K} \prod_{k=1}^{K} H_{l_k}(X_p^{(k)}[0]) H_{m_k}(X_{p'}^{(k)}[n]) \\
&= \sum_{l_1,\dots,l_K=0}^{\infty} \sum_{m_1,\dots,m_K=0}^{\infty} c^{(p)}_{l_1,\dots,l_K} c^{(p')}_{l_1,\dots,l_K} \\
&\quad \cdot E\Big[H_{l_1}(X_p^{(1)}[0]) H_{m_1}(X_{p'}^{(1)}[n])\Big] \cdots E\Big[H_{l_K}(X_p^{(K)}[0]) H_{m_K}(X_{p'}^{(K)}[n])\Big] \\
&= \sum_{l_1,\dots,l_K=0}^{\infty} c^{(p)}_{l_1,\dots,l_K} c^{(p')}_{l_1,\dots,l_K} \\
&\quad \cdot (l_1!) \Big(EX_p^{(1)}[0] X_{p'}^{(1)}[n]\Big)^{l_1} \cdots (l_K!) \Big(EX_p^{(K)}[0] X_{p'}^{(K)}[n]\Big)^{l_K} \\
&= \sum_{l_1,\dots,l_K=0}^{\infty} c^{(p)}_{l_1,\dots,l_K} c^{(p')}_{l_1,\dots,l_K} \left(\prod_{k=1}^{K} l_k!\right) (EX_p[0]X_{p'}[n]))^{l_1+\dots+l_K} \\
&= \sum_{l_1,\dots,l_K=0}^{\infty} c^{(p)}_{l_1,\dots,l_K} c^{(p')}_{l_1,\dots,l_K} \left(\prod_{k=1}^{K} l_k!\right) \big(r^X_{p,p'}[n]\big)^{l_1+\dots+l_K}.
\end{aligned}
$$

By the independence of the series $X_p^{(1)}[n], \dots, X_p^{(K)}[n]$ and the property $EH_m(X) = 0$ for $m \geq 1$, $X \sim \mathcal{N}(0,1)$, we have $EY_p[n] = c^{(p)}_{0,\dots,0}$. Then the first term corresponding to $m = l_1 + \cdots + l_K = 0$ in the summation above is $EY_p[0]EY_{p'}[n]$. With the definition of $a_m^{(p,p')}$ in Eq. (41) and by subtracting the term $EY_p[0]EY_{p'}[n]$ from $EY_p[0]Y_{p'}[n]$, to obtain the covariance, the relationship (40) follows.

Appendix E. List of popular marginals and transformations

Table E.3 gathers together transformations for various important and popular marginals where the covariance relationship, for the case $f_1 = \cdots = f_P$, is given by a simple analytic formula which is easily invertible. Note that all these transformations, except in the log-normal case, are even functions and therefore the targeted covariance has to be non-negative (see Section 2.4.1). Pareto(α, β) stands for the distribution defined by $P(Y < y) = 1 - (\beta/y)^\alpha$ for $y \geq \beta$, and zero if $y < \beta$. The probability density function for Laplace$(0, a)$ variables is given by $f(y) = \exp(-|y|/a)/(2a)$, $a > 0$.

Appendix F. Summary of the proposed synthesis algorithm

Below is a step-by-step description of the methodology for a typical synthesis task. It is assumed that the user wants to target some specific marginals F_p and covariance R_Y.

Recipe for synthesizing multivariate non-Gaussian series:

1. *Choose transform functions* f_p: The simplest choice is $f_p(x) = F_p^{-1}(\Phi(x))$ but depending on the marginals, other natural choices might be available.

2. *Find tentative Gaussian "covariance sequence"* $R_X[n]$: If available, use analytic plug-in formulas to invert the relationship between the covariances. Otherwise compute numerically the Hermite coefficients of f_p, up to a sufficiently large order (see [18] for discussion about such truncation in the univariate case). Then, for each pair (p, p'):

Table E.3: Covariance relationships for important marginals (for the case $f_1 = \cdots = f_P$).

Marginal Transform	Covariance relationship
$LN(\mu, \sigma^2)$ $f_p(x) = e^{\sigma x + \mu}$	$r^Y_{p,p'}[n] = e^{2\mu+\sigma^2}(e^{\sigma^2 r^X_{p,p'}[n]} - 1)$
χ^2_ν $f_p(x) = \sum_{k=1}^{\nu} x_k^2$	$r^Y_{p,p'}[n] = 2\nu(r^X_{p,p'}[n])^2$
Exp. with mean $\mu > 0$ $f_p(x) = \frac{\mu}{2}(x_1^2 + x_2^2)$	$r^Y_{p,p'}[n] = \mu^2(r^X_{p,p'}[n])^2$
Erlang(α, β) $f_p(x) = \frac{\beta}{2}\sum_{k=1}^{2\alpha} x_k^2$	$r^Y_{p,p'}[n] = \alpha\beta^2(r^X_{p,p'}[n])^2$
Laplace$(0, a)$ $f_p(x) = \frac{a}{2}(x_1^2 - x_2^2 + x_3^2 - x_4^2)$	$r^Y_{p,p'}[n] = 2a^2(r^X_{p,p'}[n])^2$
Unif(0,1) $f_p(x) = e^{-(x_1^2+x_2^2)/2}$	$r^Y_{p,p'}[n] = \frac{r^X_{p,p'}[n]^2}{16-4r^X_{p,p'}[n]^2}$
Pareto(α, β) $f_p(x) = \beta e^{(x_1^2+x_2^2)/(2\alpha)}$	$r^Y_{p,p'}[n] = \frac{\beta^2\alpha^2}{(\alpha-1)^2} \cdot \frac{r^X_{p,p'}[n]^2}{(\alpha-1)^2 - r^X_{p,p'}[n]^2}$

(a) Form (truncated) series $g_{p,p'}(z)$ as in (13) and plot $(z, g_{p,p'}(z))$ for $-1 \le z \le 1$. The range for each targeted correlation sequence has to fall in the range of $g_{p,p'}(z)$, otherwise the targeted covariance is not suitable and needs to be adjusted. The choice of f_p can affect this range.

(b) Invert (6) numerically for each pair (p, p'). If using the series reversion procedure detailed in Section 3.2, investigate the convergence region of the (truncated) series g^{-1} by plotting $(g^{-1}_{p,p'}(w), w)$ for w in the range of $g_{p,p'}$ found previously.

3. *Synthesize multivariate Gaussian series:* Use the algorithm for synthesiz-

ing multivariate Gaussian series X for the tentative covariance sequence R_X from last step (see Section 4.2.1). If R_X is not nonnegative definite and approximation is needed, compare the resulting approximating covariances $\tilde{r}^Y_{p,p'}$ to the targeted one. A perfect match of the targeted marginal can be ensured by renormalizing the Gaussian series X so that $r^X_{p,p}[n] = 1$.

4. *Transform Gaussian series:* Synthesize component p of the non-Gaussian series by the transform $f_p(X_p)$. The covariance sequences for this multivariate series are given by $\tilde{r}^Y_{p,p'}$.

References

[1] G. Barrenetxea, F. Ingelrest, G. Schaefer, M. Vetterli, The hitchhiker's guide to successful wireless sensor network deployments, in: Proc. of the ACM SenSys Conference, 2008.

[2] Z.-Q. Luo, M. Gastpar, J. Liu, A. Swami, Distributed signal processing in sensor networks, IEEE Signal Proc. Mag. 23 (4).

[3] P. Ciuciu, P. Abry, C. Rabrait, H. Wendt, Log wavelet leaders cumulant based multifractal analysis of EVI fMRI time series: evidence of scaling in ongoing and evoked brain activity, IEEE J. of Selected Topics in Signal Proces. 2 (6) (2008) 929–943.

[4] S. Achard, D. Bassett, A. Meyer-Lindenberg, E. Bullmore, Fractal connectivity of long-memory networks, Phys. Rev. E 77 (3) (2008) 036104.

[5] J. Goldberger, Sympathovagal balance: how should we measure it?, Am. J. Physiol., 45 (1999) 1273–1280.

[6] B. Krishnamurthy, S. Sen, Y. Zhang, Y. Chen, Sketch-based change detection: Methods, evaluation, and applications, in: IMC'03, 2003, pp. 234–247.

[7] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, D. Veitch, Multiscale nature of network traffic, IEEE Signal Proc. Mag. 19 (3) (2002) 28–46.

[8] A. Scherrer, P. Abry, Synthèse de processus bivariés non Gaussiens à mémoires longues, in: 22nd GRETSI Symposium on Signal and Image Processing, Dijon, 2009.

[9] H. Lütkepohl, New Introduction to Multiple Time Series Analysis, Springer-Verlag, Berlin, 2005.

[10] G. Teyssière, P. Abry, Wavelet analysis of nonlinear long range dependent processes. applications to financial time series, in: Long Memory in Econometrics, G. Teyssière and A. P. Kirman, editors, Springer, Berlin, 2007.

[11] M. Grigoriu, Existence and construction of translation models for stationary non-Gaussian processes, Probabilistic Engineering Mechanics 24 (4) (2009) 545 – 551.

[12] M. Grigoriu, Applied Non-Gaussian Processes, Prentice-Hall, 1995.

[13] O. Barndorff-Nielsen, N. Shephard, Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics, Journal of the Royal Statistical Society Series B 63 (2) (2001) 167–241.

[14] N. Leonenko, E. Taufer, Convergence of integrated superpositions of Ornstein-Uhlenbeck processes to fractional Brownian motion, Stochastics: An International Journal of Probability and Stochastic Processes 77 (6) (2005) 477–499.

[15] R. Price, A useful theorem for nonlinear devices having Gaussian inputs, IRE Trans. Inform. Theory IT-4 (1958) 69–72.

[16] U. G. Gujar, R. J. Kavanagh, Generation of random signals with specified probability density functions and power density spectra, IEEE Transactions on Automatic Control 13 (6) (1968) 716–719.

[17] B. Liu, D. Munson, Jr., Generation of a random sequence having a jointly specified marginal distribution and autocovariance, Acoustics, Speech and Signal Processing, IEEE Transactions on 30 (6) (1982) 973–983.

[18] B. Puig, F. Poirion, C. Soize, Non-Gaussian simulation using Hermite polynomial expansion: convergences and algorithms, Probabilistic Engineering Mechanics 17 (3) (2002) 253 – 264.

[19] B. Puig, J.-L. Akian, Non-Gaussian simulation using Hermite polynomials expansion and maximum entropy principle, Probabilistic Engineering Mechanics 19 (2004) 293–305.

[20] S. Lowen, S. Cash, M. Poo, M. Teich, Quantal neurotransmitter secretion rate exhibits fractal behavior, The Journal of Neuroscience 17(15) (1997) 5666–5677.

[21] S. Lowen, M. Teich, Fractal-Based Point Processes, Wiley Series in Probability and Statistics, 2005.

[22] M. Crouse, R. Baraniuk, Fast, exact synthesis of Gaussian and nonGaussian long-range-dependent processes, Internal research report, Rice University, Houston, Texas, USA (2004).

[23] S. Cambanis, E. Masry, On the reconstruction of the covariance of stationary Gaussian process observed through zero-memory nonlinearities, IEEE Trans. Inform. Theory IT-24 (4) (1978) 485–494.

[24] P. Bocchinia, G. Deodatis, Critical review and latest developments of a class of simulation algorithms for strongly non-Gaussian random fields, Probabilistic Engineering Mechanics 23 (2008) 393–407.

[25] T. Schreiber, A. Schmitz, Surrogate time series, Physica D 142 (3-4) (2000) 346–382.

[26] J. C. S. S. Filho, M. D. Yacoub, Coloring non-Gaussian sequences, IEEE Transactions on Signal Processing 56 (12) (2008) 5817–5822.

[27] J. M. Nichols, C. C. Olson, J. V. Michalowicz, F. Bucholtz, A simple algorithm for generating spectrally colored, non-Gaussian signals, Probabilistic Engineering Mechanics 25 (3) (2010) 315–322.

[28] K. R. Gurley, A. Kareem, A conditional simulation of non-normal velocity/pressure fields, Journal of Wind Engineering and Industrial Aerodynamics 77–78 (1998) 39–51.

[29] R. Popescu, G. Deodatis, J. Prevost, Simulation of homogeneous nonGaussian stochastic vector fields, Probabilistic Engineering Mechanics 13 (1998) 1–13(13).

[30] R. J. A. Tough, K. D. Ward, The correlation properties of gamma and other non-Gaussian processes generated by memoryless nonlinear transformation, Journal of Physics D: Applied Physics 32 (1999) 3075–3084.

[31] M. Chambers, The simulation of random vector time series with given spectrum, Mathl. Comput. Modelling 22 (2) (1995) 1–6.

[32] G. Chan, A. Wood, Simulation of stationary Gaussian vector fields, Statistics and Computing 9 (4) (1999) 265–268.

[33] H. Helgason, V. Pipiras, P. Abry, Fast and exact synthesis of stationary multivariate Gaussian time series using circulant embedding, Signal Processing, In Press. doi:10.1016/j.sigpro.2010.10.014.

[34] D. Nualart, The Mallavian Calculus and Related Topics, 2nd Edition, Probability and its Applications, Springer-Verlag, Berlin, 2006.

[35] H. Lancaster, Some properties of the bivariate normal distribution considered in the form of a contingency table, Biometrika 44 (1957) 289–292.

[36] P. Brockwell, R. Davis, Introduction to Time Series and Forecasting, 2nd Edition, Springer, 2002.

[37] U. Grenander, G. Szegő, Toeplitz Forms and their Applications, Univ. of California Press, 1958.

[38] P. Henrici, Applied and Computational Complex Analysis, vol. 1, Wiley, 1974.

[39] M. Arcones, Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors, Ann. Probab. 22 (4) (1994) 2242–2274.

[40] K. Park, W. Willinger, Self-similar network traffic: An overview, in: K. Park, W. Willinger (Eds.), Self-Similar Network Traffic and Performance Evaluation, Wiley, New York, 2000, pp. 1–38.

[41] A. Scherrer, N. Larrieu, P. Owezarski, P. Borgnat, P. Abry, Non Gaussian and long memory statistical characterisations for Internet traffic with anomalies, IEEE Transaction on Dependable and Secure Computing 4 (1) (2007) 56–70.

[42] J.-M. Bardet, G. Lang, G. Oppenheim, A. Philippe, S. Stoev, M. Taqqu, Semi-parametric estimation of the long-range dependence parameter: A survey, in: P. Doukhan, G. Oppenheim, M. S. Taqqu (Eds.), Theory and applications of Long-range dependence, Birkhäuser, Boston, 2003, pp. 557–577.

[43] P. Abry, H. Helgason, V. Pipiras, Wavelet-based analysis of non-gaussian long-range dependent processes and estimation of the hurst parameter, submitted.

[44] MAWI Traffic Archive, WIDE Project, `http://mawi.wide.ad.jp/mawi/`.