Random vector and time series definition and synthesis from matrix product representations: From Statistical Physics to Hidden Markov Models

Florian Angeletti*, Eric Bertin, Patrice Abry, *IEEE Fellow*

Université de Lyon, Laboratoire de Physique, Ecole Normale Supérieure de Lyon, CNRS
46 allée d'Italie, F-69007 Lyon, France
ens-lyon.fr/PHYSIQUE/, firstname.lastname@ens-lyon.fr

***Abstract*—[1] [2] Inspired from modern out-of-equilibrium statistical physics models, a matrix product based framework is defined and studied, that permits the formal definition of random vectors and time series whose desired joint distributions are a priori prescribed. Its key feature consists of preserving the writing of the joint distribution as the *simple* product structure it has under independence, while inputing controlled dependencies amongst components: This is obtained by replacing the product of probability densities by a product of matrices of probability densities. It is first shown that this matrix product model can be remapped onto the framework of Hidden Markov Models. Second, combining this dual perspective enables us both to study the statistical properties of this model in terms of marginal distributions and dependencies (a stationarity condition is notably devised) and to devise an efficient and accurate numerical synthesis procedure. A design procedure is also described that permits the tuning of model parameters to attain targeted statistical properties. Pedagogical well-chosen examples of times series and multivariate vectors aim at illustrating the power and versatility of the proposed approach and at showing how targeted statistical properties can actually be prescribed.**

***Index Terms*—Random vectors and Time Series numerical synthesis, Joint Distribution, Matrix Product, Hidden Markov Model, Statistical Physics.**

EDICS Category: SSP-SNMD SSP-NGAU SSP-NSSP

I. Introduction

In modern signal processing, it is very often needed that synthetic data are produced numerically in fast and efficient manners with (some of) their statistical properties being a priori prescribed as closely as desired from selected targets (marginal distributions, covariance, spectrum, multivariate distribution, ...). This is for instance the case when the performance of newly developed statistical analysis need to be assessed. The theoretical derivation of such performance may turn too difficult to achieve, specially when it is intended to apply such tools to real-world data, whose properties are not well known. Instead, one can resort to Monte Carlo simulations: Performance are derived from averages over independent realizations of synthetic data, that are designed to resemble as close as possible the real-world data. Another example stems from Bayesian estimation procedures, that generally involve Monte Carlo Markov Chain [1] or variational based resolution schemes (cf. e.g., [2], [3]) to generate numerically independent copies of hyper-parameters drawn from (possibly complicated) distributions, derived from the Bayesian formalism.

For the numerical synthesis of Gaussian time series, with prescribed covariance function, the so-called *Circulant Embeded Matrix* synthesis procedure [4]–[6] is nowadays considered as the state-of-the-art solution and is currently widely used. It has then been extended to the synthesis of multivariate Gaussian time series, with prescribed auto- and cross-covariance, notably in [7]–[9] . The synthesis of non-Gaussian time series turns far more complicated as the prescription of the full joint distributions turns very difficult to achieve while that of the sole marginal distributions and covariance functions does not uniquely define the process. Several approaches were proposed to address such issues (cf. e.g., [10], [11] and references therein for reviews). Other strategies are based on surrogate techniques: Starting from the original real-world data of interest, surrogate copies are obtained by randomizing one attribute of the data (e.g., the phase of its Fourier transform) while maintaining another fixed (e.g., the amplitude of the Fourier transform) (cf. e.g., [12]). Recently, an *optimal transport* procedure was proposed aiming at iteratively modifying the joint distribution of random vectors or time series to attain a given target [13]. The general framework of Markov Chain models offers an alternative and broad class of solutions, focusing on the modeling of local dynamical properties, while not explicitly putting the emphasis on a direct prescription of the joint distributions of the process (cf. e.g., [14] for a review). Markov Chain schemes are also widely used to synthesize independent realizations of random vectors with prescribed properties (cf. e.g., [14] for a review).

The present contribution takes place in this long tradition of synthetic data design in signal processing, but is however rooted in a very different scientific field, that of statistical physics. Inspired from the exact solutions of stochastic out-of-equilibrium models describing particle diffusion on a one-

[2]J.-Y. Tourneret and N. Pustelnik are gratefully acknowledged for fruitful advices on time series design and convex optimization, respectively. Anonymous reviewers are also gratefully acknowledged for fruitful comments. Work partially supported by the *Young Research Team* Del Duca French Academy of Sciences Award, 2007.

dimensional lattice with volume exclusion, like the Asymmetric Simple Exclusion Processes (ASEP) [15]–[17], the design procedure proposed here is founded on the request that the joint distribution of the random vectors or time series $\underline{X}_N = (x_1, x_2, \dots, x_N)$ be written as a(n oriented) product $P_{\underline{X}} \propto \prod_{k=1}^{N} R_d(x_k)$, as in the case of independent components. However, the R_ds no longer consist of univariate distributions, but instead of d dimensional matrices of distributions. In statistical physics, this matrix product form of the joint distribution was envisaged as a practical ansatz to find exact solutions of the master equation associated e.g. to the ASEP model. In signal processing, this enables us to devise a theoretically powerful and practically versatile and efficient procedure to define and generate numerically random vectors or random time series, with a priori prescribed statistical properties. A preliminary and partial attempt relying on specific choices of matrices and focusing on stationary time series only was presented in [18]. It is here extended to a more general and versatile setting.

After definition, the first contribution of this article is to show (Section II) that this Statistical Physics matrix product framework can be remapped onto that of Hidden Markov Models [19], more classical in Signal Processing Theory. The statistical properties (partial and joint distributions and multivariate moments) of the model are then studied in depth in Section III, and a general condition ensuring stationarity for time series is further devised. In Section IV, the variety of dependence structures that can be achieved within this framework are analyzed and the tuning of the time-scales that enter these dependence structures is studied. Taking advantage of the combined matrix product and Hidden Markov Model perpectives, a fast and efficient practical synthesis procedure is devised in Section V and it is further explained, through a well chosen stationary time series example, how targeted marginal distributions and other statistical properties can be attained. Section VI consists in a set of pedagogical examples aiming at illustrating the efficiency and versatility of the tool for the synthesis of various random vectors with same prescribed marginal distributions and correlations but different joint distributions.

II. Matrix product and Hidden Markov models

A. Joint probability density function as a matrix product

Inspired from out-of-equilibrium statistical physics models (cf. e.g. [15]–[17]), a general framework for the design of joint probability density functions of random vectors is defined, based on density matrix product and relying on four key ingredients. First, let $d \in \mathbb{N}\backslash\{0\}$. Second, let $\mathcal{A}$ denote a fixed non-zero non-random $d \times d$ real matrix, with positive entries, referred to as the *projection matrix*, and $\mathcal{L}(M)$ an associated linear form defined (for any matrix M) by:

$$\mathcal{L}(M) = \mathsf{tr}\left(\mathcal{A}^T M\right). \tag{1}$$

Third, let $\mathcal{E}$ denote a fixed non-zero non-random $d \times d$ matrix, with positive entries $\mathcal{E}_{i,j}$, referred to as the *structure matrix* (for reasons made clear in Section IV-A). Fourth, let matrix $\mathcal{P}$ consists of a $d \times d$ set of valid probability density functions $(\mathcal{P}_{i,j}(x))_{i=1,\dots,d;j=1,\dots,d}$ on a sample space Ω. Further, let

$$R_d(x) = \mathcal{E} \odot \mathcal{P}(x), \tag{2}$$

where $\odot$ denotes the Hadamard (entry-wise) matrix product.

Let $\underline{X} \equiv (X_n)_{1 \le n \le N}$ denote the random vector defined by its joint probability density function:

$$\mathbb{P}_{\underline{X}}(x_1, \dots, x_N) = \frac{\mathcal{L}\left(\prod_{k=1}^{N} R_d(x_k)\right)}{\mathcal{L}\left(\mathcal{E}^N\right)}, \tag{3}$$

where $\prod_{k=1}^{N} R_d(x_k) = R_d(x_1) \dots R_d(x_N)$ denotes the oriented product (i.e., the order of the factors is fixed). We can show that the joint choice of positive entries for matrices $\mathcal{A}$ and $\mathcal{E}$, and of valid probability density functions in $\mathcal{P}$ is sufficient to ensure that Eq. (3) defines a valid joint probability density function under the condition that $\mathcal{L}\left(\mathcal{E}^N\right) \neq 0$. Though the model is defined for any arbitrary sample space Ω, this article is focused on the case $\Omega = \mathbb{R}$.

B. Hidden Markov Model remapping

The algebraic form of $\mathbb{P}_{\underline{X}}$ in Eq. (3) allows us to derive the statistical properties of $\underline{X}$, as Section III shall show. However, this form gives little hints on how to construct and synthesize numerically such random vectors. To overcome that drawback, the present section aims at showing that for any joint probability density function $\mathbb{P}_{\underline{X}}$ defined from Eq. (3), an equivalent Hidden Markov Model can be obtained.

First, generalizing the fact that the entries of the matrix product (ABC) reads $(ABC)_{i,j} = \sum_{k,l} a_{i,k} b_{k,l} c_{l,j}$, enables us to recast Eq. (3) into

$$\begin{aligned}\mathbb{P}_{\underline{X}}(x_1, \dots, x_N) &= \sum_{\underline{\Gamma}} \frac{\mathcal{A}_{\Gamma_0,\Gamma_N}}{\mathcal{L}\left(\mathcal{E}^N\right)} \prod_{k=1}^{N} R_d(x_k)_{\Gamma_{k-1},\Gamma_k} \\ &= \sum_{\underline{\Gamma}} \frac{\mathcal{A}_{\Gamma_0,\Gamma_N}}{\mathcal{L}\left(\mathcal{E}^N\right)} \prod_{k=1}^{N} \mathcal{E}_{\Gamma_{k-1},\Gamma_k} \mathcal{P}(x_k)_{\Gamma_{k-1},\Gamma_k}\end{aligned} \tag{4}$$

that can be re-written as

$$\mathbb{P}_{\underline{X}}(x_1, \dots, x_N) = \sum_{\underline{\Gamma}} \kappa(\underline{\Gamma}) \mathbb{P}_{\underline{X}}(x_1, \dots, x_N | \underline{\Gamma}), \tag{5}$$

$$\text{where } \mathbb{P}_{\underline{X}}(x_1, \dots, x_N | \underline{\Gamma}) = \prod_{k=1}^{N} \mathcal{P}(x_k)_{\Gamma_{k-1},\Gamma_k}, \tag{6}$$

$$\kappa(\underline{\Gamma}) = \frac{\mathcal{A}_{\Gamma_0,\Gamma_N}}{\mathcal{L}(\mathcal{E}^N)} \prod_{k=1}^{N} \mathcal{E}_{\Gamma_{k-1},\Gamma_k}, \tag{7}$$

with $\sum_{\underline{\Gamma}} \kappa(\underline{\Gamma}) = 1$. Therefore, $\underline{\Gamma} \equiv (\Gamma_0, \dots, \Gamma_k, \dots, \Gamma_N)$ can be interpreted as one realization of a random vector taking values in $\{1, \dots, d\}^{N+1}$, according to the multivariate probability distribution $\kappa(\underline{\Gamma})$. Further, $\mathbb{P}_{\underline{X}}(x_1, \dots, x_N | \underline{\Gamma})$ denotes distribution of $\underline{X}$, conditional to a realization of the random vector $\underline{\Gamma}$. Hence, $\mathbb{P}_{\underline{X}}$ can be read as a $\kappa(\underline{\Gamma})$-weighted mixture of laws, each defined as the product $\prod_{k=1}^{N} \mathcal{P}_{\Gamma_{k-1},\Gamma_k}(x_k)$.

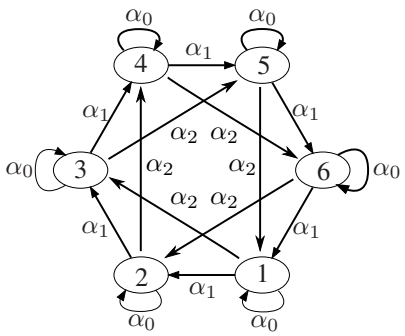

Fig. 1. **Transition graph.** Example for $d = 6$ and $\mathcal{E} = \alpha_0 I_d + \alpha_1 J_d + \alpha_2 J_d^2$ ($\alpha_0 + \alpha_1 + \alpha_2 = 1$ where J_d is defined as in Eq. (37)).

Second, let $\Gamma_{\geq t} \equiv (\Gamma_t, \ldots, \Gamma_{N-1})$ denote the set of chains starting at index t and stopping at index $N-1$. For a given pair (Γ_0, Γ_N), Eq. (7) enables us to show that:

$$\begin{aligned}\mathbb{P}(\Gamma_k = j | \Gamma_0 = \gamma_0, \ldots \Gamma_{k-1} = \gamma_{k-1}) = \\ \frac{\sum_{\Gamma_{\geq k+1}} \kappa\left(\gamma_0 \cdots \gamma_{k-1} \cdot j \cdot \Gamma_{\geq k+1}\right)}{\sum_{\Gamma_{\geq k}} \kappa\left(\gamma_0 \cdots \gamma_{k-1} \cdot \Gamma_{\geq k}\right)} \\ \frac{\mathcal{E}_{\gamma_0,\gamma_1} \ldots \mathcal{E}_{\gamma_{k-1},j} \sum_{\Gamma_{\geq k+1}} \mathcal{E}_{j,\Gamma_{k+1}} \ldots \mathcal{E}_{\Gamma_{N-1},\Gamma_N}}{\mathcal{E}_{\gamma_0,\gamma_1} \ldots \mathcal{E}_{\gamma_{k-2},\gamma_{k-1}} \sum_{\Gamma_{\geq k}} \mathcal{E}_{\gamma_{k-1},\Gamma_k} \ldots \mathcal{E}_{\Gamma_{N-1},\Gamma_N}} \\ = \mathcal{E}_{\gamma_{k-1},j} \frac{\left(\mathcal{E}^{N-k}\right)_{j,\Gamma_N}}{\left(\mathcal{E}^{N-k+1}\right)_{\gamma_{k-1},\Gamma_N}} = \mathbb{P}(\Gamma_k = j | \Gamma_{k-1} = \gamma_{k-1})\end{aligned} \tag{8}$$

and hence that $\underline{\Gamma}$ consists of an inhomogeneous d-state Markov chain, with transition probability matrix at step k reading:

$$\mathbb{P}(\Gamma_k = j | \Gamma_{k-1} = i) = \mathcal{E}_{i,j} \frac{\left(\mathcal{E}^{N-k}\right)_{j,\Gamma_N}}{\left(\mathcal{E}^{N-k+1}\right)_{i,\Gamma_N}}. \tag{9}$$

This shows that $\mathcal{E}$ can now be interpreted as (the basis for) the transition matrix underlying the Markov chain $\underline{\Gamma}$, cf. Fig. 1.

Third, Eq. (7) enables us to show that the initial distribution for (Γ_0, Γ_N) reads

$$\mathbb{P}(\Gamma_0 = i, \Gamma_N = j) = \frac{\mathcal{A}_{i,j} (\mathcal{E}^N)_{i,j}}{\mathcal{L}\left(\mathcal{E}^N\right)}. \tag{10}$$

We have thus shown that one can associate to any random vector $\underline{X}$ defined from its joint probability as in Eq. (3), a Hidden Markov Model with three hidden states: The current state Γ_k; The previous state Γ_{k-1}; and the final state Γ_N.

Such a dual perspective for $\underline{X}$ is fruitful: On the one hand, the matrix product is suited to the computation of statistical properties (cf. Sections III and IV); On the other hand, the Hidden Markov Model leads to an efficient and versatile synthesis procedure (cf. Sections V and VI).

III. Matrix product and statistical properties

The present section aims at devising the statistical properties of $\underline{X}$ as defined from Eq. (3): First, multivariate distributions and moments are established; Second, a sufficient condition to strong stationarity of $\underline{X}$, then seen as a time series, is derived; Third, achievable dependence structures are studied and shown to be controlled by the matrix $\mathcal{E}$.

A. Multivariate moments and distributions

From Eq. (3) and commutativity between integration and matrix product, an explicit expression for the univariate (or marginal) distributions $\mathbb{P}(X_k = x_k)$ can be derived:

$$\begin{aligned}\mathbb{P}(X_k = x_k) &= \frac{\int_{\mathbb{R}^{N-1}} \mathcal{L}\left(\prod_{i=1}^N R_d(x_i)\right) \prod_{i \neq k} dx_i}{\mathcal{L}\left(\mathcal{E}^N\right)} \\ &= \frac{\mathcal{L}\left(\int R_d(x_1) dx_1 \ldots R_d(x_k) \ldots \int R_d(x_N) dx_N\right)}{\mathcal{L}\left(\mathcal{E}^N\right)} \\ &= \frac{\mathcal{L}\left(\mathcal{E}^{k-1} R_d(x_k) \mathcal{E}^{N-k}\right)}{\mathcal{L}\left(\mathcal{E}^N\right)}.\end{aligned} \tag{11}$$

This shows that the marginal distributions necessarily consist of weighted linear combinations of the $\mathcal{P}_{i,j}$,

$$\mathbb{P}(X_k = x) = \sum_{i,j} c_{i,j,k} \mathcal{P}_{i,j}(x), \tag{12}$$

where the $c_{i,j,k}$ are functions of $\mathcal{A}$, $\mathcal{E}$ and N satisfying, $\forall k \in \{1, \ldots, N\}$, $\sum_{i,j} c_{i,j,k} = 1$:

$$c_{i,j,k} = \frac{1}{\mathcal{L}\left(\mathcal{E}^N\right)} \sum_{s,e} \mathcal{A}_{s,e} \left(\mathcal{E}^{k-1}\right)_{s,i} \mathcal{E}_{i,j} \left(\mathcal{E}^{N-k}\right)_{j,e}. \tag{13}$$

On condition that it exists, the univariate moment of order q can be shown to obey the following closed-form expressions:

$$\mathbb{E}\left[X_k^q\right] = \frac{\mathcal{L}(\mathcal{E}^{k-1} M(q)\, \mathcal{E}^{N-k})}{\mathcal{L}(\mathcal{E}^N)}, \tag{14}$$

with matrices $M(q) = \int_{\mathbb{R}} x^q R_d(x) dx = \mathcal{E} \odot \int_{\mathbb{R}} x^q \mathcal{P}(x) dx$, whose entries read

$$M(q)_{i,j} = \mathcal{E}_{i,j} \int_{\mathbb{R}} x^q \mathcal{P}_{i,j}(x) dx. \tag{15}$$

The univariate moment $\mathbb{E}\left[X_k^q\right]$ thus exists if and only if the moments of order q of every $\mathcal{P}_{i,j}(x)$ exist.

Further, The p-variate distributions (with $k_1 < \cdots < k_p$) can also be made explicit, $\mathbb{P}(X_{k_1} = x_{k_1}, \ldots, X_{k_p} = x_{k_p}) =$

$$\frac{\mathcal{L}\left(\mathcal{E}^{k_1-1} \left(\prod_{r=1}^{p-1} R_d(x_{k_r}) \mathcal{E}^{k_{r+1}-k_r-1}\right) R_d(x_{k_p}) \mathcal{E}^{N-k_p}\right)}{\mathcal{L}\left(\mathcal{E}^N\right)}, \tag{16}$$

as well as the p-variate moments (with q_r the order associated to entry k_r and $k_1 < \cdots < k_n$): $\mathbb{E}\left[\prod_{r=1}^p X_{k_r}^{q_r}\right] =$

$$\frac{\mathcal{L}\left(\mathcal{E}^{k_1-1} \left(\prod_{r=1}^{p-1} M(q_r)\, \mathcal{E}^{k_{r+1}-k_r-1}\right) M(q_p)\, \mathcal{E}^{N-k_p}\right)}{\mathcal{L}(\mathcal{E}^N)}. \tag{17}$$

A fundamental consequence of Eq. (16) consists of the fact that any subset of components $\{X_{k_1}, \ldots, X_{k_p}\}$ can still be associated to a matrix product based joint probability distribution, where, however, matrix R_d now varies with component index r according to $R_d^{(r)}(x) = R_d(x) \mathcal{E}^{k_{r+1}-k_r-1}$ (with the convention that $k_{p+1} = N$) for $r \neq 1$ and $R_d^{(1)}(x) = \mathcal{E}^{k_1-1} R_d(x) \mathcal{E}^{k_2-k_1-1}$.

This generalized setting where matrix R_d is not necessarily constant but may vary with component index k, $R_d^{(k)}$, is introduced and discussed in detail in Section VI-A.

B. Strong stationarity

1) Strong stationarity condition: Eq. (3) shows that the in general non-stationary nature of $\underline{X}_N$ stems from the non-commutativity of the matrix product. However, imposing the additional commutativity property for $\mathcal{A}^T$ and $\mathcal{E}$,

$$[\mathcal{A}^T, \mathcal{E}] \equiv \mathcal{A}^T\mathcal{E} - \mathcal{E}\mathcal{A}^T = 0, \tag{18}$$

we are able to force stationarity in $\underline{X}$. Indeed, this yields

$$\begin{aligned} \forall M, \mathcal{L}(\mathcal{E}M) &= \mathsf{tr}\left(\mathcal{A}^T\mathcal{E}M\right) \\ &= \mathsf{tr}\left(\mathcal{E}\mathcal{A}^TM\right) \\ &= \mathsf{tr}\left(\mathcal{A}^TM\mathcal{E}\right) = \mathcal{L}(M\mathcal{E}). \end{aligned} \tag{19}$$

Notably, for any x, we have $\mathcal{L}(\mathcal{E}R_d(x)) = \mathcal{L}(R_d(x)\mathcal{E})$.

Assuming Eq. (18) is satisfied, the marginal distributions in Eq. (12) simplify to a same distribution no longer depending on k:

$$\mathbb{P}(X_k = x) = \frac{\mathcal{L}\left(R_d(x)\mathcal{E}^{N-1}\right)}{\mathcal{L}\left(\mathcal{E}^N\right)} = \sum_{i,j} c_{i,j}\mathcal{P}_{i,j}(x), \tag{20}$$

with

$$c_{i,j} = \frac{1}{\mathcal{L}\left(\mathcal{E}^N\right)}\sum_e \mathcal{A}_{i,e}\mathcal{E}_{i,j}\left(\mathcal{E}^{N-1}\right)_{j,e} = \mathcal{E}_{i,j}\frac{\left(\mathcal{E}^{N-1}\mathcal{A}^T\right)_{j,i}}{\mathcal{L}\left(\mathcal{E}^N\right)}. \tag{21}$$

The p-variate distributions (cf. Eq. (16)) and moments (cf. Eq. (17)) reduce to $\mathbb{P}(X_{k_1} = x_{k_1}, \ldots, X_{k_p} = x_{k_p}) =$

$$\frac{\mathcal{L}\left(\left(\prod_{r=1}^{p-1} R_d(x_{k_r})\mathcal{E}^{k_{r+1}-k_r-1}\right)R_d(x_{k_p})\mathcal{E}^{N-(k_p-k_1)-1}\right)}{\mathcal{L}\left(\mathcal{E}^N\right)}, \tag{22}$$

$$\mathbb{E}\left[\prod_{r=1}^{p} X_{k_r}^{q_r}\right] = \frac{\mathcal{L}\left(\left(\prod_{r=1}^{p-1} M(q_r)\,\mathcal{E}^{k_{r+1}-k_r-1}\right)M(q_p)\,\mathcal{E}^{N-(k_p-k_1)-1}\right)}{\mathcal{L}\left(\mathcal{E}^N\right)}. \tag{23}$$

Notably, for $p = 1$: $\mathbb{E}[X_k] = \mathcal{L}\left(M(1)\,\mathcal{E}^{N-1}\right)/\mathcal{L}\left(\mathcal{E}^N\right)$. These relations clearly indicate that the vector $(X_n)_{1\le n\le N}$ can now be regarded as a time series with strong stationarity: All joint statistics depend only on time differences, $k_{r+1} - k_r$. The sole matrix $\mathcal{A}$ satisfying Eq. (18) for all matrices $\mathcal{E}$, is the identity matrix, in which case $\mathcal{L}$ consists of the trace operator. However, the trace operator induces automatically a circular correlation structure, $l \ge 2k$, $\mathbb{E}[X_kX_l] = \mathbb{E}[X_kX_{N+2k-l}]$, a highly undesirable consequence for application purposes.

2) Doubly stochastic matrices: Alternatively, we have found that Eq. (18) is automatically satisfied (and thus strong stationarity obtained) when selecting jointly specific pairs $(\mathcal{A}, \mathcal{E})$, such that

$$\mathcal{A}_{i,j} = (1/d) \tag{24}$$

and $\mathcal{E}$ is a so-called *doubly stochastic* matrix, defined as

$$\forall i, j, \quad \sum_k \mathcal{E}_{k,j} = \sum_k \mathcal{E}_{i,k} \equiv 1. \tag{25}$$

In addition, such choices yield

$$\forall k,\, M, \quad \mathcal{L}\left(M\mathcal{E}^k\right) = \mathcal{L}(M), \tag{26}$$

that further leads to a major simplification in Eq. (20) to Eq. (23): The marginal and p-variate distributions of $\underline{X}_N$, as well as the p-variate moments, no longer depend explicitly on the sample size N, and their expressions simplify to

$$\mathbb{P}(X_k = x) = \mathcal{L}(R_d(x)) = \sum_{i,j} c_{i,j}\mathcal{P}_{i,j}(x), \quad c_{i,j} = \frac{\mathcal{E}_{i,j}}{d}. \tag{27}$$

$$\mathbb{P}(X_{k_1} = x_{k_1}, \ldots, X_{k_p} = x_{k_p}) = \mathcal{L}\left(\left(\prod_{r=1}^{p-1} R_d(x_{k_r})\mathcal{E}^{k_{r+1}-k_r-1}\right)R_d(x_{k_p})\right), \tag{28}$$

$$\mathbb{E}\left[\prod_{r=1}^{p} X_{k_r}^{q_r}\right] = \mathcal{L}\left(\left(\prod_{r=1}^{p-1} M(q_r)\,\mathcal{E}^{k_{r+1}-k_r-1}\right)M(q_p)\right). \tag{29}$$

Eq. (29) constitutes a key result with respect to applications as it clearly shows that the joint statistics of order q of $\underline{X}$ can be prescribed by the sole selection of matrices $M(q)$. Elaborating on a preliminary work (cf. [18]), this particular setting will be used in Section V-B to efficiently design stationary time series.

IV. Dependence structure

Eqs. (16) and (17) indicate that the structure matrix $\mathcal{E}$ essentially controls the dependence structure within vector $\underline{X}_N$, hence its name, via the collection of its powers $\mathcal{E}^n$, $n = 1, \ldots, N$, while the matrices $M(q)$ fix the amplitudes of the dependencies at order q. Analyzing the forms possibly taken by $\mathcal{E}^n$ is hence crucial to understand the potential dependence structures of $\underline{X}_N$ achievable in this framework. Notably, the classical distinction between diagonalizable and non-diagonalizable $\mathcal{E}$ plays a crucial role. This issue is now studied in detail for the correlation structure.

A. Diagonalizable structure matrix

1) General case : Let us assume that $\mathcal{E}$ is diagonalizable, with r different eigenvalues $\lambda_1, \ldots, \lambda_r$. Then, $\mathcal{E}$ can be decomposed into r sub-matrices $E_1, \ldots, E_r$ such that

$$\mathcal{E}^k = \sum_{i=1}^{r} \lambda_i^k E_i, \tag{30}$$

and Eq. (11) can be rewritten as

$$\mathbb{P}(X_k = x) = \frac{\sum_{i,j} \lambda_i^{k-1}\lambda_j^{N-k}\mathcal{L}(E_iR_d(x)E_j)}{\sum_{i=1}^{r} \lambda_i^N \mathcal{L}(E_i)}, \tag{31}$$

which makes explicit the dependence in k and is reminiscent of Eq. (12) given that the $\mathcal{L}(E_iR_d(x)E_j)$ consist of linear combinations of the $\mathcal{P}_{l,m}(x)$. Further, for the 2-sample statistics, or covariance function, Eq. (17) applied to the points k and l (with $k < l$) becomes

$$\mathbb{E}[X_kX_l] = \frac{\sum_{i,j,m} \lambda_i^{k-1}\lambda_j^{N-l}\lambda_m^{l-k-1}\mathcal{L}(E_iM(1)\,E_mM(1)\,E_j)}{\sum_{i=1}^{r} \lambda_i^N \mathcal{L}(E_i)}. \tag{32}$$

Assuming stationarity, i.e., Eq. (18), the relation above further reduces to

$$\mathbb{E}\left[X_k X_l\right] = \frac{\sum_{j,m} \lambda_j^{N-2} \left(\frac{\lambda_m}{\lambda_j}\right)^{l-k-1} \mathcal{L}\left(M(1)\, E_m M(1)\, E_j\right)}{\sum_{i=1}^{r} \lambda_i^N \mathcal{L}\left(E_i\right)}, \tag{33}$$

which shows that the covariance function consists of the sum of weighted exponential functions $\exp(-(l-k-1)(\ln\lambda_j - \ln\lambda_m))$, with at most $r_M(r_M-1)/2 = \binom{r_M}{2}$ characteristic time scales, $\tau_{j,m} = (\ln|\lambda_j| - \ln|\lambda_m|)^{-1}$, where r_M stands for the number of eigenvalues of matrix $\mathcal{E}$ with different modulus. A preliminary study of the covariance function in the stationary and diagonalizable case has been devised in [18] and an example is worked out in Section V-B.

Without assuming stationarity, this exponential decrease of the covariance function still holds. Indeed, let us assume that the norm of λ_1 is strictly larger than the norm of the other eigenvalues and that $\mathcal{L}\left(E_1\right) \neq 0$. Then, the normalisation term $\mathcal{L}\left(\mathcal{E}^N\right)$ can be approximated in the limit $N \to +\infty$ as $\mathcal{L}\left(\mathcal{E}^N\right) \sim \lambda_1^N \mathcal{L}\left(E_1\right)$. Combining this asymptotic form with Eq. (31) yields

$$\begin{aligned}\mathbb{P}(X_k = x) &= \sum_{i,j} \lambda_i^{k-1} \lambda_j^{N-k} \frac{\mathcal{L}\left(E_i R_d(x) E_j\right)}{\mathcal{L}\left(\mathcal{E}^N\right)} \\ &\sim \frac{\mathcal{L}\left(E_1 R_d(x) E_1\right)}{\lambda_1 \mathcal{L}\left(E_1\right)},\end{aligned} \tag{34}$$

and combining this result with Eq. (32) leads to:

$$\mathbb{E}\left[X_k X_l\right] \sim \sum_i \left(\frac{\lambda_i}{\lambda_1}\right)^{l-k-1} \frac{\mathcal{L}\left(E_1 M(1)\, E_i M(1)\, E_1\right)}{\mathcal{L}\left(E_1\right) \lambda_1^2}. \tag{35}$$

These equations show that using a diagonalizable $\mathcal{E}$, with a dominant eigenvalue, implies that asymptotically, i.e., in the limit $N \to +\infty$, each component of the vector $\underline{X}_N$ shares the same univariate distribution, and that the autocovariance functions reads as a sum of exponential functions, depending only on $|k-l|$. In the limit $N \to +\infty$, vector $\underline{X}_N$ is hence asymptotically quasi-stationary.

2) Circular structure matrix: An interesting subset of diagonalisable matrices is that of circulant matrices $\mathcal{E}$:

$$\mathcal{E} = \sum_{k=0}^{d-1} \alpha_k J_d^k, \;\; \sum_{k=0}^{d-1} \alpha_k \equiv 1, \tag{36}$$

where J_d is a $d \times d$ real matrix defined as:

$$J_d = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & & & & \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix} \tag{37}$$

and $J_d^0 = I_d$, the Identity matrix. Circulant matrices are also doubly stochastic, so that $\mathcal{A}_{i,j} = (1/d)$ is a natural choice. The matrix B of eigenvectors, and the eigenvalues $\underline{\lambda} = \{\lambda_k\}$ of the circulant matrix $\mathcal{E}$ can be written as:

$$B_{i,j\in\{0,\ldots,d-1\}^2} = \omega_d^{ij}, \quad \text{with } \omega_d = \exp(2\imath\pi/d) \tag{38}$$

$$\text{and } \underline{\lambda} = B\underline{\alpha}. \tag{39}$$

From these definitions, using a change of basis

$$\tilde{M}(q) = B^{-1} M(q)\, B \tag{40}$$

leads to a simple expression of the dependencies

$$\mathbb{E}\left[X_k^{q_1} X_l^{q_2}\right] = \sum_i \lambda_i^{|k-l|} \tilde{M}(q_1)_{d,i} \tilde{M}(q_2)_{i,d}. \tag{41}$$

For ease of notations, let us further define the two vectors $(C_{M(q)})_k = \sum_i M(q)_{i,k}$, $(R_{M(q)})_k = \sum_j M(q)_{k,j}$ and let $\mathcal{F}$ denote the (normalized) Discrete Fourier Transform:

$$\mathcal{F}(\underline{v})_k = \frac{1}{d} \sum_{l=1}^{d} v_l \omega_d^{kl}, \quad k = 1, \ldots, d. \tag{42}$$

This enables to rewrite $\tilde{M}(q)_{i,d}$ and $\tilde{M}(q)_{d,i}$ as:

$$\begin{aligned}\tilde{M}(q)_{i,d} &= \frac{1}{N} \sum_{k,l} \omega_d^{ik-ld} M(q)_{k,l} = \overline{\mathcal{F}(\underline{C}_{M(q)})_k} \\ \tilde{M}(q)_{d,i} &= \mathcal{F}(\underline{R}_{M(q)})_k\end{aligned}$$

Because $\lambda_{d-j} = \overline{\lambda_j}$, $j = 1, \ldots, d-1$, conjugate eigenvalues λ_k can be grouped by pairs, leading to an explicit form of the higher order dependencies:

$$\begin{aligned}&\mathbb{E}\left[X_0^{q_1} X_t^{q_2}\right] - \mathbb{E}\left[X_0^{q_1}\right] \mathbb{E}\left[X_t^{q_2}\right] = \\ &\sum_{k=1}^{\lfloor d/2 \rfloor} m_k \mathfrak{Re}\left\{\overline{\mathcal{F}(\underline{C}_{M(q_1)})_k} \mathcal{F}(\underline{R}_{M(q_2)})_k \lambda_k^{t-1}\right\},\end{aligned} \tag{43}$$

where $\lfloor z \rfloor$ stands for the integer part of z, $m_k = 1$ if $2k = d$ and $m_k = 2$ otherwise, $\mathfrak{Re}\{\}$ denotes the real part, and the overbar designates the complex conjugate. Rewriting

$$\lambda_k = e^{-\frac{1}{\tau_k}} e^{\pm\imath\frac{2\pi}{T_k}}, k = 1, \ldots \lfloor d/2 \rfloor, \tag{44}$$

Eq. (41) and Eq. (43) above show that dependencies of any order in $\underline{X}$ consist of the weighted sums of d exponential decreasing functions, with at most $\lfloor d/2 \rfloor$ distinct characteristic time scales for the dependencies in $\underline{X}$, $\tau_k = -1/\ln|\lambda_k|$ and $\lfloor d/2 \rfloor$ distinct oscillating periods: $T_k = 2\pi/\arg(\lambda_k)$. The λ_k (and thus the τ_k) depend explicitly on the joint choices of d and $\{\alpha_k, k = 0, \ldots, d-1\}$.

B. Non-diagonalizable structure matrix

Non-diagonalizable matrices $\mathcal{E}$ draw a very different landscape. For illustration, let us consider the case where $\mathcal{E} = I_d + H$ with I_d the Identity matrix and H any nilpotent matrix of order $p+1$ (i.e., $H^{p+1} \equiv 0$ while $H^k \neq 0$, for $1 \leq k \leq p$), chosen as a standard example of non-diagonalizable matrix. Then, for any $k \geq p$, one has

$$\mathcal{E}^k = \sum_{j=0}^{p} \binom{k}{j} H^j,$$

which combined with Eq. (31) yields

$$\mathbb{P}(X_k = x_k) = \frac{\sum_{i,j\leq p} \binom{k-1}{i} \binom{N-k}{j} \mathcal{L}\left(H^i R_d(x) H^j\right)}{\sum_{i=1}^{p} \binom{N}{i} \mathcal{L}\left(H^i\right)} \tag{45}$$

and, for the covariance function

$$\mathbb{E}\left[X_k X_l\right] = \frac{\sum_{i,j,m\leq p}\binom{k-1}{i}\binom{l-k-1}{m}\binom{N-l}{j}\mathcal{L}\left(H^i M(1)\, H^m M(1)\, H^j\right)}{\sum_{i=1}^{p}\binom{N}{i}\mathcal{L}\left(H^i\right)}. \tag{46}$$

To gain a better grasp of what these equations imply, let us study their asymptotic behavior as $N \to +\infty$. Using

$$\binom{N}{p} \sim \frac{N^p}{p!}, \quad N \to +\infty$$

and the assumption that $\mathcal{L}\left(H^p\right) \neq 0$ leads to

$$\mathcal{L}\left(\mathcal{E}^N\right) \sim \frac{N^p}{p!}\mathcal{L}\left(H^p\right)$$

and, assuming that k diverges with N, to

$$\begin{aligned}\mathbb{P}(X_k = x) &= \sum_{i,j\leq p}\binom{k-1}{i}\binom{N-k}{j}\frac{\mathcal{L}\left(H^i R_d(x) H^j\right)}{\mathcal{L}\left(\mathcal{E}^N\right)} \\ &\sim \sum_{i+j=p}\binom{p}{i}\left(\frac{k}{N}\right)^i\left(1-\frac{k}{N}\right)^j\frac{\mathcal{L}\left(H^i R_d(x) H^j\right)}{\mathcal{L}\left(H^p\right)}.\end{aligned} \tag{47}$$

Compared to Eq. (31), this indicates that the marginal distribution of component k reads as a mixture of distributions, with weights depending on the relative position k/N, rather than the absolute position k, as is the case for diagonalizable $\mathcal{E}$.

The covariance function also turns quite different from the diagonalizable case, for large N,

$$\mathbb{E}\left[X_k X_l\right] \sim \sum_{i+j+m=p}\left[\frac{p!}{i!m!j!}\left(\frac{k}{N}\right)^i\left(\frac{l-k}{N}\right)^j\left(1-\frac{l}{N}\right)^m\frac{\mathcal{L}\left(H^i M(1)\, H^j M(1)\, H^m\right)}{\mathcal{L}\left(H^p\right)}\right] \tag{48}$$

with the occurrence of algebraic terms, $\left(\frac{l-k}{N}\right)^j$, that indicate long-range correlations developing on a range that is as large as the size of $\underline{X}_N$.

This specific choice for non-diagonalizable $\mathcal{E}$ enables us to figure out that, combining a block diagonal Jordan reduction of $\mathcal{E}$ and results obtained in the diagonalizable case, the covariance function consists, in the very general case, of a weighted sum of algebraic and exponential decreases.

V. Synthesis and Design

A. Synthesis procedure from Hidden Markov Model

While the matrix product structure of the probability density function allowed us to derive the statistical properties of $\underline{X}$ (see Section III), this algebraic structure turns quite useless to derive an actual synthesis procedure. In constrast, exploiting the Hidden Markov Model remapping enables us to devise an efficient and versatile numerical synthesis procedure.

1) General case: Combining Eqs. (9) and (10), the synthesis algorithm can be sketched as follows:

Step 1: Initialization:
  Use Eq. (10) to generate the states Γ_0 and Γ_N.
Step 2: Iteration on $k = 1, \ldots, N$:
  Step 2.1: Choose at random state Γ_k, according to the transition probability given in Eq. (9);
  Step 2.2: Generate X_k according to $\mathcal{P}_{\Gamma_{k-1},\Gamma_k}$.

2) Stationary doubly stochastic case: Under the particular choices $\mathcal{A}_{i,j} = 1/d$ (cf. Eq. (24)) and $\mathcal{E}$ doubly stochastic (cf. Eq. (25)), it can be shown that there is no need to set a priori the final state Γ_N and that the Markov chain becomes homogeneous, so that Eqs. (9) and (10) can be rewritten as (cf. [18]):

$$\mathbb{P}(\Gamma_k = j | \Gamma_{k-1} = i) = \mathcal{E}_{i,j} \text{ and } \mathbb{P}(\Gamma_0 = i) = \frac{1}{d}. \tag{49}$$

Such simplifications show that $\mathcal{E}$ exactly defines the transition matrix for the Markov chain Γ and, moreover, that the initial state of the Markov chain Γ_0 follows a uniform distribution.

3) Computational cost: For the general case where $\mathcal{E}$ is not an invertible matrix, the algorithmic complexity is dominated by the cost of the computation of matrix powers, and hence a total complexity in $\mathcal{O}(d^3 N \ln N)$ (with a $\mathcal{O}(d^3)$ matrix multiplication cost and a $\mathcal{O}(\ln N)$ matrix power computation cost). When $\mathcal{E}$ is invertible, the global cost reduces to $\mathcal{O}(d^3 N)$ (as matrix multiplications only are required). With the choice of doubly stochastic matrices $\mathcal{E}$ and $\mathcal{A}_{i,j} = 1/d$, the computational cost is further reduced to $\mathcal{O}(N \ln d)$. Therefore, the synthesis procedure scales efficiently for large sample size (large N) and large complex dependence structure (large d) (with numerous further potential optimizations for specific cases such as sparse structure matrix).

B. Design of a stationary time series

When all four ingredients, $d, \mathcal{A}, \mathcal{E}$ and $\mathcal{P}$ are selected, the synthesis procedures above enable to produce numerically independent copies of the desired $\underline{X}$. Let us now explain how to tune these parameters to reach targeted properties of $\underline{X}$. For sake of simplicity, it is first detailed in the stationary time series set up. Non stationary vectors are further addressed in Section VI.

1) Problem formulation: In real-world applications, it is seldom the case that the joint marginal distribution of a time series is chosen a priori. Instead one rather prefers to target first the marginal distribution, second, the covariance function, and then, potentially, higher-order dependencies.

To satisfy the stationarity request, and for sake of simplicity, design is envisaged in the framework of Section III-B, where $\mathcal{A}$ is chosen as in Eq. (24) and $\mathcal{E}$ as a circulant matrix (cf. Eq. (36)). With these choices, not only is the constraint $[\mathcal{A}^T, \mathcal{E}]$ satisfied, but the associated Hidden Markov Model is homogeneous. The goal is now to explain how to select the parameters of the model that remain free, namely d, $\underline{\alpha} = (\alpha_0, \ldots, \alpha_{d-1})$ and matrix $\mathcal{P}(x)$, to reach i) a targeted marginal distribution $\mathbb{P}(x)$, ii) a targeted covariance function and potentially iii) targeted higher order dependencies of order

q_1 and q_2. Following Eq. (43), reachable correlation functions are necessarily sums of (possibly oscillating) decreasing exponential functions. We thus choose a set of target covariance function $c_{q_1,q_2}(t) \equiv \mathbb{E}\left[X_0^{q_1} X_t^{q_2}\right] - \mathbb{E}\left[X_0^{q_1}\right]\mathbb{E}\left[X_t^{q_2}\right]$ of the form

$$c_{q_1,q_2}(t) = \sum_{k=1}^{r} \mathfrak{Re}\left\{\beta\left(q_1,q_2\right)_k \theta_k^t\right\} \quad (t>0), \tag{50}$$

for s distinct couples (q_1,q_2), which generically include $q_1 = q_2 = 1$. Eq. (50) defines a set of r characteristic time scales θ_k, and of sr amplitudes $\beta\left(q_1,q_2\right)_k$. The goal is thus, given a targeted marginal distribution $\mathbb{P}(x)$, targeted time scales θ_k and targeted dependence amplitudes $\beta\left(q_1,q_2\right)_k$, to select d, $\underline{\alpha}$ and $\mathcal{P}(x)$ so that the generated signal $\underline{X}_N$ satisfies the targeted properties. The constraints to be satisfied are for the marginal (cf. Eq. (27))

$$\frac{1}{d}\sum_{i,j} \mathcal{P}_{i,j}(x) = \mathbb{P}(x) \tag{51}$$

and for the covariances

$$\sum_{k=1}^{\lfloor d/2 \rfloor} m_k \mathfrak{Re}\left\{\overline{\mathcal{F}(\underline{C}_{M(q_1)})_k}\mathcal{F}(\underline{R}_{M(q_2)})_k \lambda_k^{t-1}\right\} = c_{q_1,q_2}(t) \tag{52}$$

for all $t>0$, and for the s couples (q_1,q_2). In order to satisfy these constraints, we proceed through three successive steps.

2) Time scale selection: For a given d, the number of characteristic time scales $\lambda_k = e^{-\frac{1}{\tau_k}} e^{\pm\imath\frac{2\pi}{T_k}}$ available in the synthesized process is $\lfloor d/2 \rfloor$, as seen from Eq. (43). The validity of Eq. (52) for all t implies that $\lfloor d/2 \rfloor = r$ and that $\lambda_k = \theta_k$, $k = 1,\ldots,r$, which defines the full vector $(\lambda_0,\ldots,\lambda_{d-1})$ using the constraints $\lambda_0 = 1$ and $\lambda_{d-j} = \overline{\lambda_j}$ $(j = 1,\ldots,d-1)$ required by the circulant matrix framework. Note that it is also possible to choose d such that $\lfloor d/2 \rfloor > r$, by adding arbitrary time scales θ_k, $k = r+1,\ldots,\lfloor d/2 \rfloor$ with zero amplitude $\beta\left(q_1,q_2\right)_k$. Then, using Eq. (39), $\underline{\alpha}$ is determined from the relation

$$\underline{\alpha} = \mathcal{F}(\underline{\lambda}). \tag{53}$$

The condition $\sum_{i=0}^{d-1}\alpha_i = 1$ is automatically satisfied since we imposed $\lambda_0 = 1$. The positivity constraint for the α_js is harder to characterize and is satisfied when $\underline{\lambda}$ belongs to the convex polytope generated by the vertices $v_i^{(k)} = \omega_d^{ki}$ for $k \in \{1,\ldots,d-1\}$ (where ω_d is defined in Eq. (38). Following [20], a convex optimization procedure, based on a projected gradient descent, has been implemented by ourselves[3], that outputs the best ℓ_2 approximate solution $\underline{\tilde{\alpha}}$ to Eq. (53) that simultaneously satisfies the range constraint $\underline{\tilde{\alpha}} \in [0,1]^d$ and the summation to 1 constraint (cf. Eq. (36)).

3) Controlling the amplitudes of the dependency: Now that d and $\underline{\alpha}$ have been set, the only remaining degree of freedom is the matrix $\mathcal{P}(x)$, which should be chosen to meet both the targeted dependency amplitudes $\beta\left(q_1,q_2\right)_k$, and the marginal distribution $\mathbb{P}(x)$. Eq. (52) shows that the constraint on $\beta\left(q_1,q_2\right)_k$ only involves the matrix $M(q)$, for all values of q appearing in the s couples (q_1,q_2). It is thus convenient to first determine the $M(q)$s that match the constraints on $\beta\left(q_1,q_2\right)_k$, and in a second stage to determine $\mathcal{P}(x)$ satisfying the resulting $M(q)$s, together with the marginal distribution $\mathbb{P}(x)$. To make effective the search for $M(q)$ in the potentially very large solution space, and consistently with the chosen form of the matrix $\mathcal{E}$, we propose an efficient parameterization of the matrices $M(q)$, in terms of a zero trace diagonal matrix $\tilde{D}\left(q\right) = \mathrm{diag}(\underline{D}\left(q\right))$ (i.e., $\mathsf{tr}(\tilde{D}\left(q\right)) = 0$):

$$M(q) = \sum_{k=0}^{d-1} \alpha_k \tilde{D}\left(q\right) J_d^k + \mathbb{E}\left[X^q\right]\mathcal{E}, \tag{54}$$

with $\mathbb{E}\left[X^q\right] \equiv \int x^q \mathbb{P}(x)\,\mathrm{dx}$. This parameterization results, using the properties of J_d^k, from the assumption that the deviation of $\mathcal{P}_{i,j}(x)$ from the marginal $\mathbb{P}(x)$ depends only on i, and satisfies

$$\forall i,j,\ \int x^q\left(\mathcal{P}_{i,j}(x) - \mathbb{P}(x)\right)dx = D\left(q\right)_i. \tag{55}$$

We can then show that the parameterization (54) simplifies dependencies in Eq. (43) into:

$$\begin{aligned}\mathbb{E}\left[X_0^{q_1}X_t^{q_2}\right] - \mathbb{E}\left[X_0^{q_1}\right]\mathbb{E}\left[X_t^{q_2}\right] = \\ \sum_{k=1}^{\lfloor d/2 \rfloor} m_k \mathfrak{Re}\left\{\overline{\mathcal{F}\left(\underline{D}\left(q_1\right)\right)}_k \mathcal{F}\left(\underline{D}\left(q_2\right)\right)_k \lambda_k^t\right\}.\end{aligned} \tag{56}$$

Combining Eq. (52) and Eq. (56), one sees that $\underline{D}\left(q\right)$ can be determined from $\beta\left(q_1,q_2\right)_k$ according to the relation

$$\beta\left(q_1,q_2\right)_k = m_k\overline{\mathcal{F}\left(\underline{D}\left(q_1\right)\right)}_k \mathcal{F}\left(\underline{D}\left(q_2\right)\right)_k, \quad k = 1,\ldots,r. \tag{57}$$

From the targeted $\beta\left(q_1,q_2\right)_k$, the $\underline{D}\left(q\right)$s can be obtained by inverting the set of equations (57).

4) Construction of $\mathcal{P}_{i,j}$: Now that d, $\mathcal{A}$, $\mathcal{E}$ and $M(q)$ (or $\underline{D}\left(q\right)$) have been chosen, it remains to select the matrix $\mathcal{P}(x)$ satisfying the constraints on the marginal $\mathbb{P}(x)$ (cf. Eq. (51)) and on the matrices $M(q)$, as expressed from Eq. (55). The difficulty in choosing the matrix $\mathcal{P}(x)$ thus consists in disentangling both groups of constraints.

To this end, we have devised a constructive procedure that relies on parameterizing $\mathcal{P}_{i,j}(x)$ using a set of shape functions $h = \{h_{i,j}\} \in \mathbb{R} \to \mathbb{R}^+\backslash\{0\}$, and a set of amplitudes $\nu = \{\nu_{i,j}\} \in \mathbb{R}^+\backslash\{0\}$:

$$\mathcal{P}_{i,j}(x;\nu,h) = \frac{1}{w(\nu,h)_{i,j}}\ \frac{\nu_{i,j}h_{i,j}(x)}{\sum_{k,l}\nu_{k,l}h_{k,l}(x)}\mathbb{P}(x) \tag{58}$$

$$\text{with } w(\nu,h)_{i,j} = \int_x \frac{\nu_{i,j}h_{i,j}(x)}{\sum_{k,l}\nu_{k,l}h_{k,l}(x)}\mathbb{P}(x)dx. \tag{59}$$

This parametric form satisfies by construction the relation

$$\sum_{i,j} w(\nu,h)_{i,j}\mathcal{P}_{i,j}(x;\nu,h) = \mathbb{P}(x). \tag{60}$$

Therefore, the marginal constraint in Eq. (51) is satisfied when

$$w(\nu,h)_{i,j} = \frac{\mathcal{E}_{i,j}}{d}. \tag{61}$$

For a given set of functions h, Eq. (61) corresponds to a set of non linear equations in ν. Furthermore, it can be shown that, for any matrix c such that $\sum_{i,j} c_{i,j} = 1$, Equation $w(\nu,h) = c$ always admits at least one solution $\nu^*(h)$. Setting $\mathcal{P}(x;h)_{i,j} \equiv$

[3]with the gratefully acknowledged assistance of N. Pustelnik.

$\mathcal{P}(x;\nu^*(h),h)_{i,j}$ thus produces a parameterization of $\mathcal{P}_{i,j}$, that automatically satisfies the marginal constraint.

The last step is to find the $h_{i,j}(x)$ such that $\mathcal{P}(x;h)$ satisfies the moment constraints listed in Eq. (55) above. As there is now only a finite number of constraints (as opposed to the functional constraint on the marginal), it is possible restrict the functions $h_{i,j}(x)$ to a parameterized kernel family $K_{\underline{m}}$: $h_{i,j}(x) = K_{\underline{m}_{i,j}}(x)$. To ensure the existence of solutions, it is natural to use kernels $K_{\underline{m}}$ with a parameter vector $\underline{m}$ whose size equals or exceeds the number of targeted moments constraints in Eq. (55). Note that the kernel parameters cannot consist of pure amplitude factors, as amplitudes are already accounted for by parameters $\nu_{i,j}$s. Typical kernels used here read

$$K_{m,\gamma}(x) = \frac{1}{\sqrt{2\pi\gamma^2}}\,\exp[-(x-m)^2/(2\gamma^2)] \tag{62}$$

and

$$K_{m,\gamma}(x) = \frac{1}{1+(x-m)^2/\gamma}. \tag{63}$$

Once the kernel chosen, classical non linear solvers [21] are used to solve Eq. (55).

5) Summary of the procedure: In summary, to design stationary time series $\underline{X}$ with prescribed univariate distributions and dependence structure, the following procedure is to be followed:

1) Select d and $\underline{\alpha}$ in agreement with the targeted dependence structure time scales (cf. Eq. (53) and Section V-B2).
2) Fix the moment matrices $M(q)$ or $\tilde{D}(q)$ in agreement with the targeted dependence amplitudes, at chosen order q. (cf. Eq. (57) and Section V-B3).
3) Select a kernel family $K_{\underline{m}}$ (cf. Eq. (62) and Eq. (63)).
4) Compute the distribution matrix $\mathcal{P}$ in agreement with the marginal distribution and moment constraints (cf. Equations (58) to (61) and Section V-B4).

A MATLAB implementation of this procedure is available on https://github.com/FAngeletti/MatrixAlea.

C. Illustration

To illustrate the potential and versatility of the design procedure described in previous section, let us assume that we intend to design two independent stationary time series, $\underline{X}$ and $\underline{Y}$, with same targeted marginal distribution $p_S(x)$ (here chosen to be a Laplace distribution $p_S(x) \equiv \sigma^{-1}\exp(-|x|/\sigma)$, with $\sigma = 2$), same targeted covariance function $c_{1,1}(t) = C_1\mathfrak{Re}\{\theta_1^t\}$ (for $t > 0$), but different targeted covariance of squares: $c^{\underline{X}}_{2,2} = C_{X,2}\mathfrak{Re}\{\theta_2^t\}$ and $c^{\underline{Y}}_{2,2} = C_{Y,2}\mathfrak{Re}\{\theta_3^t\}$ (for $t > 0$). Given that the joint definition of both $\underline{X}$ and $\underline{Y}$ requires three time constants (θ_1, θ_2 and θ_3) and three amplitudes ($C_1, C_{X,2}, C_{Y,2}$), $d = 6$ is sufficient to construct two different solutions. We arbitrarily choose:

$$\theta_k = e^{1/\tau_k}, \quad \tau_1 = 164,\ \tau_2 = 90,\ \tau_3 = 56 \tag{64}$$

and obtain $\underline{\alpha}$ from Eq. (53).

Because $\underline{X}$ and $\underline{Y}$ share the same covariance function, a simple choice is that they share the same vector $\underline{D}(1)$. To obtain a single time scale, $\lambda_1 = \theta_1$, with non-zero amplitudes entering the covariance function, Eq. (57) yields, with $q_1 = q_2 = 1$:

$$D(1)_k = \sqrt{2C_1}(\cos(k\pi)) \tag{65}$$

To obtain different covariance functions for the squares, each with a single non zero amplitude time scale, Eq. (57) yields, with $q_1 = q_2 = 2$:

$$D(2)^{(\underline{X})}{}_k = \sqrt{C_{2,\underline{X}}}\left(\cos\left(k\frac{\pi}{3}\right)\right) \tag{66}$$

and

$$D(2)^{(\underline{Y})}{}_k = \sqrt{C_{2,\underline{Y}}}\left(\cos\left(k\frac{2\pi}{3}\right)\right). \tag{67}$$

With this the values of $D(1)$ and $D(2)$ fixed, the procedure described in subsection V-B4 can be applied to obtain the $\mathcal{P}_{i,j}$, using here Kernel in Eq. (62). The $d = 6$ resulting $\mathcal{P}_{i,j}$ are shown in Fig. 2, second row.

Once all parameters are set, the Hidden Markov Model synthesis procedure described in Section V-A is used to synthesize the sample paths shown in Fig. 2 for $\underline{X}$ (left column) and $\underline{Y}$ (right column), together with the comparisons of their estimated and targeted marginal distributions (2nd row), covariance functions (3rd row) and covariance of the squares (bottom row). It clearly shows that $\underline{X}$ and $\underline{Y}$ have the same marginal and covariance but different covariance of the squares, hence different joint distributions (as targeted).

VI. Random vector definition and synthesis

This section now aims at showing how this matrix product framework can be further generalized and tailored to the definition of random vectors, whose components have (possibly) different prescribed marginal distributions as well as prescribed joint distributions. As illustration and exercise, we will show, on different trivariate vectors, how to vary the joint probability density functions while maintaining fixed the three different marginal distributions as well as pair-wise correlations.

A. Multivariate design

The matrix product formalism defined in Section II-A could be used directly to design non stationary random vectors (e.g., with possibly different marginals) by increasing the dimension of matrix R_d to $d \propto N$. It proves, however, more fruitful theoretically to generalize the matrix product formalism defined in Eq. (3) by letting matrix R_d depend on the component index k:

$$\mathbb{P}_{\underline{X}}(x_1,\ldots,x_N) = \frac{\mathcal{L}\left(\prod_{k=1}^N R_d^{(k)}(x_k)\right)}{\mathcal{L}\left(F(1,N)\right)}, \tag{68}$$

where for any $l \geq k$

$$F(k,l) = \prod_{j=k}^{l}\mathcal{E}^{(j)} \text{ with } \mathcal{E}^{(j)} = \int_{\mathbb{R}} R_d^{(j)}(x)dx. \tag{69}$$

and for $l < k$, $F(k,l) = \mathsf{Id}$.

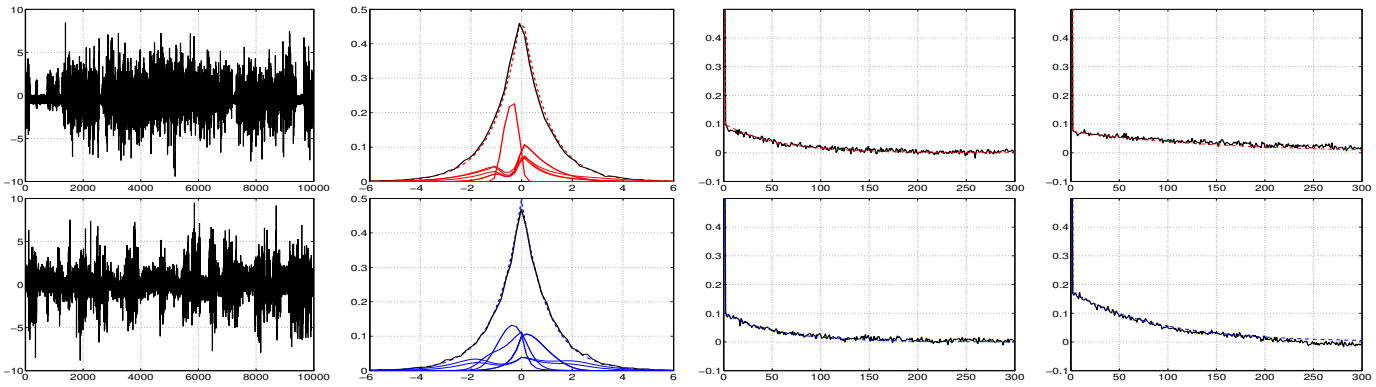

Fig. 2. **Numerical Synthesis.** Each row correspond to two different time series with the same marginal distribution (a Laplace distribution with $\lambda = 1$), with the same correlation function but different correlation functions for their squares, hence different prescribed joint distributions. First column, example of one sample path. Then, from left to right, estimated (solid black lines) and theoretical (dashed-colored lines) marginals (together with constructed $\mathcal{P}_{i,j}$, solid-colored lines), correlation functions and correlation functions for squared time series.

Following the reasoning in Section II-B , this generalization can also be extended to a hidden Markov model whose transition matrix, initial state distribution and conditional probability density function read:

$$R_d^{(k)}(x) = \mathcal{E}^{(k)} \odot \mathcal{P}^{(k)}(x), \tag{70}$$

$$\mathbb{P}\left(\Gamma_k = j | \Gamma_{k-1} = i\right) = \mathcal{E}_{i,j}^{(k)} \frac{F(k+1,N)_{j,f}}{F(k,N)_{i,f}}, \tag{71}$$

$$\mathbb{P}\left(\Gamma_0 = s, \Gamma_N = f\right) = A_{s,f} \frac{F(1,N)_{s,f}}{\mathcal{L}\left(F(1,N)\right)}, \tag{72}$$

$$\mathbb{P}\left(\underline{X} | \Gamma\right) = \prod_{i=1}^{N} \mathcal{P}_{\Gamma_{i-1},\Gamma_i}^{(k)}(x_i). \tag{73}$$

Within this generalized framework, the counterparts of Eqs. (11), (16) and (17) can also be derived:

$$\mathbb{P}(X_k = x_k) = \frac{\mathcal{L}\left(F(1,k-1) R_d^{(k)}(x_k) F(k+1,N)\right)}{\mathcal{L}\left(F(1,N)\right)}, \tag{74}$$

$$\mathbb{P}(X_{k_1} = x_{k_1}, \ldots, X_{k_p} = x_{k_p}) = \frac{1}{\mathcal{L}\left(F(1,N)\right)} \mathcal{L}\Bigg[F(1,k_1-1) \left(\prod_{r=1}^{p} R_d^{(k_r)}(x_{k_r}) F(k_r+1, k_{r+1}-1) \right) \Bigg], \tag{75}$$

$$\mathbb{E}\left[\prod_{r=1}^{p} X_{k_r}^{q_r} \right] = \frac{1}{\mathcal{L}(F(1,N))} \mathcal{L}\Bigg(F(1,k_1-1) \left(\prod_{r=1}^{p} M^{(k_r)}(q_r) F(k_r+1, k_{r+1}-1) \right) \Bigg), \tag{76}$$

with $M^{(k)}(q) = \int_{\mathbb{R}} x^q R_d^{(k)}(x) dx$ and with, by convention, $k_{p+1} = N+1$.

This general framework turns however uneasy to handle for practical purposes, as it implies the tuning of far too many degrees of freedom (e.g., the joint selection of the structure matrices $\mathcal{E}^{(k)}$ and the density probability matrices $\mathcal{P}^{(k)}$, since $R_d^{(k)}(x) = \mathcal{E}^{(k)} \odot \mathcal{P}^{(k)}(x)$). A fruitful trade-off consists in letting $\mathcal{P}^{(k)}$ depend on k while retaining structure matrices that do not depend on k:

$$\mathcal{E}^{(k)} = \mathcal{E}^*. \tag{77}$$

With that choice, one obtains that F reads, for any $l \geq k-1$,

$$F(k,l) = \left(\mathcal{E}^*\right)^{l-k+1}, \tag{78}$$

which implies that achievable dependence structures are of the exact same form as those obtained for the constant matrix $R_d^{(k)} \equiv R_d$ case and detailed in Section IV.

This trade-off setting permits to define each component of the random vector $\underline{X}_N$ with a relative independence. Notably, the univariate distribution of X_k depends only on k, $\mathcal{P}^{(k)}$ and $\mathcal{E}^*$.

It is furthermore possible to eliminate the explicit dependency in k by imposing

$$[\mathcal{A}^T, \mathcal{E}^*] = 0. \tag{79}$$

That condition, which no longer implies stationarity along k as the probability density matrix $\mathcal{P}^{(k)}$ still depends on k, yields a further simplification for the univariate distributions (compared to Eq. (74)):

$$\mathbb{P}(X_k = x_k) = \frac{1}{\mathcal{L}\left(\mathcal{E}^{*N}\right)} \mathcal{L}\left(R_d^{(k)}(x_k) \mathcal{E}^{*N-1}\right). \tag{80}$$

Corresponding expressions can be computed for partial joint density functions and sample moments.

B. Trivariate Example

As a pedagogical example, Let us now develop the construction of two different trivariate random vectors $\underline{X}_N = (X_1, X_2, X_3)$ and $\underline{Y}_N = (Y_1, Y_2, Y_3)$, with marginal distributions set to a Gaussian distribution $\mathcal{N}_{0,1}$, for X_1 and Y_1, to a Gamma distribution, with shape and scale parameters $(\alpha, \beta) = (2, 3)$ for X_2 and Y_2, and to a Gamma distribution with $(\alpha, \beta) = (1, 2)$ for X_3 and Y_3. To illustrate the potential of the tool, $\underline{X}_N$ and $\underline{Y}_N$ also have the same correlations, but different joint distributions.

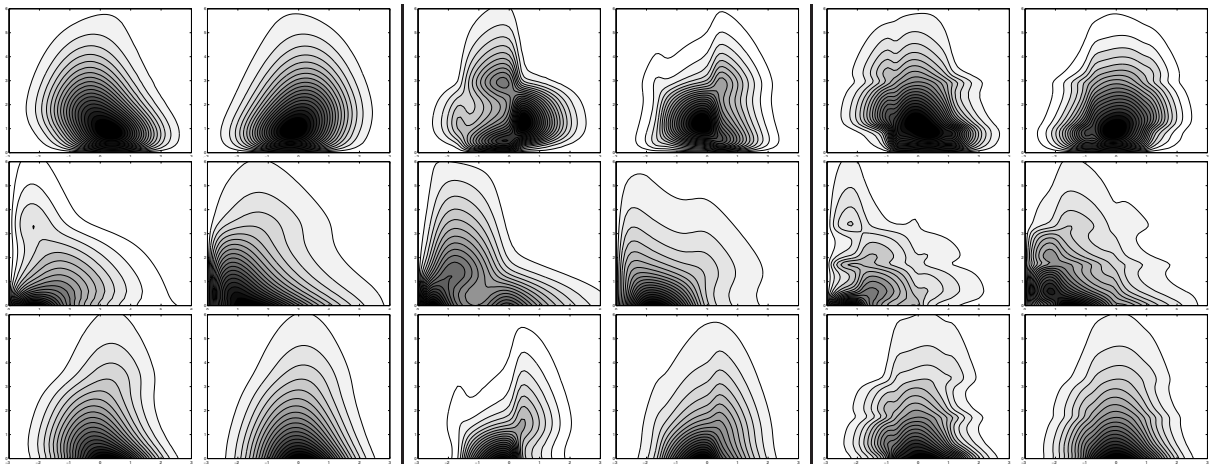

Fig. 3. Bivariate probability distributions of $\underline{X}_N$ (Left) and $\underline{Y}_N$ (Middle) and $\underline{Z}_N$(Right), for pairs 1-2 (top), pairs 2-3 (middle), pairs 1-3 (bottom). For each vector: Left, $\alpha_0 = 0.1$; Right, $\alpha_0 = 0.8$ correspond respectively to negative and positive correlation.

First, $d = 2$, $\mathcal{E}^* = \alpha_0 I_d + \alpha_1 J_d$ and $\mathcal{A}_{i,j} = (1/d)$ are selected ($\alpha_0 + \alpha_1 = 1$). Second, to control correlations, the moment matrix $M^{(k)}(1)$ is set, for both vectors, to ($k = 1, 2, 3$)

$$M^{(k)}(1) = \mathcal{E}^* \odot \begin{pmatrix} m_{k,1} & m_{k,1} \\ m_{k,2} & m_{k,2} \end{pmatrix}, \tag{81}$$

with the constraints $(m_{k,1} + m_{k,2})/2 = \mathbb{E}[X_k]$. Defining $\Delta_k = (m_{k,1} - m_{k,2})/2$, the covariance reads, for both $\underline{X}_N$ and $\underline{Y}_N$:

$$\begin{aligned} \mathsf{Cov}[X_1, X_2] &= \mathsf{Cov}[Y_1, Y_2] = (2\alpha_0 - 1)\Delta_1\Delta_2, \\ \mathsf{Cov}[X_2, X_3] &= \mathsf{Cov}[Y_2, Y_3] = (2\alpha_0 - 1)\Delta_2\Delta_3, \\ \mathsf{Cov}[X_1, X_3] &= \mathsf{Cov}[Y_1, Y_3] = (2\alpha_0 - 1)^2\Delta_1\Delta_3. \end{aligned}$$

Therefore, the covariance for any two consecutive components depends linearly on α_0, and, when $\Delta_k\Delta_l > 0$, is maximum for $\alpha_0 = 1$, vanishes at $\alpha_0 = 0.5$ and is minimum for $\alpha_0 = 0$.

The two trivariate joint distributions can now be made different via their moment matrices, $M^{(k)}(2)$, which, to adjust the targeted variances of each component is, for $\underline{X}_N$, set to:

$$\begin{aligned} &M^{(1)}(2) = \mathcal{E}^* \odot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, M^{(3)}(2) = \mathcal{E}^* \odot \begin{pmatrix} 4.5 & 4.5 \\ 11.5 & 11.5 \end{pmatrix}, \\ &M^{(2)}(2) = \mathcal{E}^* \odot \begin{pmatrix} 4.5 & 4.5 \\ 7.5 & 7.5 \end{pmatrix}; \end{aligned} \tag{82}$$

and for $\underline{Y}_N$ to

$$\begin{aligned} &M^{(1)}(2) = \mathcal{E}^* \odot \begin{pmatrix} 0.5 & 0.5 \\ 1.5 & 1.5 \end{pmatrix}, \quad M^{(3)}(2) = \mathcal{E}^* \odot \begin{pmatrix} 8 & 8 \\ 8 & 8 \end{pmatrix}. \\ &M^{(2)}(2) = \mathcal{E}^* \odot \begin{pmatrix} 2.75 & 2.75 \\ 9.25 & 9.25 \end{pmatrix}. \end{aligned} \tag{83}$$

To construct the $\mathcal{P}^{(k)}{}_{i,j}$ from the procedure developed in Section V-B4, the kernel $K_{m,\sigma}$ is chosen as defined in Eq. (62). The resulting probability matrices $\mathcal{P}^{(k)}$ have only two distinct components.

Furthermore, changing the kernel K parametrizing $\mathcal{P}^{(k)}_{i,j}$, also constitutes an efficient way to further vary the joint distributions, while maintaining fixed the marginal distributions, hence introducing further versatility in the procedure. For example, K could be chosen as the Gamma distribution family, or the union of the Gaussian and Gamma families. To illustrate this, let us now construct a third trivariate random vector $\underline{Z}_N$, sharing the same marginals and covariance $\underline{X}_N$ and $\underline{Y}_N$ (it actually shares exactly the same matrices $M^{(k)}(1)$ and $M^{(k)}(2)$ as those of $\underline{X}_N$), though obtained from kernel $K_{m,\sigma}(x) = (0.1 + ((x-m)/\sigma)^2)\exp(-((x-m)/\sigma)^2)$.

Eq. (75) is used to compute the bivariate partial distributions (rather than trivariate joint distributions, for clarity) which are shown in Fig. 3 for the pairs (X_1, X_2) (top row), (X_2, X_3) (middle row) and (X_1, X_3) (bottom row) for $\underline{X}_N$ (left block), $\underline{Y}_N$ (middle block) and $\underline{Z}_N$ (right block), for negative (left column in each block) and positive (right column) correlations. These plots clearly show that, for any two pairs, the bivariate distributions (hence a fortiori the joint trivariate distributions) are different for the three vectors, though their components have same univariate marginals and same correlations. Furthermore, Fig. 4 shows how the targeted marginal distributions $\mathbb{P}(x_k)$ are obtained by summation of the $c_{i,j,k}\mathcal{P}^{(k)}_{i,j}$ which were produced by the algorithmic procedure described in sub-section V-B4. Note that with the chosen setting of the proposed examples, the summation has been a priori limited (for simplicity) to only two different $\mathcal{P}^{(k)}_{i,j}$, while $d \times d = 4$ distinct distributions would have been available in general. Interestingly, comparisons for a given component (i.e., for a given row) amongst the three vectors illustrates that a same marginal distribution $\mathbb{P}(x_k)$ is attained from different summation of the $c_{i,j,k}\mathcal{P}^{(k)}_{i,j}$ for the three vectors, which constitutes a signature of the fact that the joint distributions of $\underline{X}_N$, $\underline{Y}_N$ and $\underline{Z}_N$ are different.

Here, the example was chosen trivariate, as a trade-off between ease of exposition (three partial distributions of pairs of components remain easy to display) and demonstration of the potential and richness of the tool. Multivariate examples are however just as easy to construct.

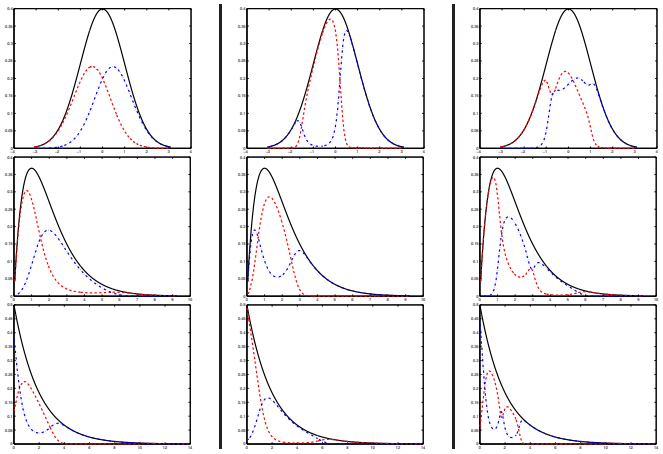

Fig. 4. Marginal distributions $\mathbb{P}(X_k)$ (solid lines) for $\underline{X}_N$ (left), $\underline{Y}_N$ (middle) and $\underline{Z}_N$ (right) and the designed $c_{i,j,k}\mathcal{P}_{i,j}$ (dashed lines); for components X_1, Y_1, Z_1 (top); Components X_2, Y_2, Z_2 (middle); Component X_3, Y_3, Z_3 (bottom).

VII. Conclusion and Perspectives

Inspired from out-of-equilibrium statistical physics models, a general framework for the definition of random vectors, based on a matrix product formalism, has first been shown to possess an equivalent Hidden Markov Model reformulation, a highly non trivial result. Second, playing jointly with the matrix product and Hidden Markov Model perspectives, we have been able to study in depth the statistical properties of the model (achievable dependence structures have been characterized in details, a stationarity condition has been obtained and fully exploited) as well as to devise a practical and effective numerical synthesis procedure. Further, we have explored theoretically the ability of the model to shape jointly marginal distributions and dependence structure and have devised a constructive procedure enabling to effectively tune the model parameters to reach targeted marginal distributions and dependence structures. Pedagogical examples, for both stationary time series and random vectors, have been detailed to illustrate the versatility and richness of the procedure in actually attaining targeted properties. Both for targeted property design and numerical synthesis, MATLAB implementation of the proposed procedures are available upon request.

Comparisons of the proposed approach to other numerical synthesis frameworks in terms of potential, versatility, efficiency, precision and implementation are under current investigations but are beyond the scope of the present contribution. Further, the benefits of using other matrices $\mathcal{A}$ and $\mathcal{E}$ will be explored. Moreover, having maintained the writing of the multivariate distributions as a product, as is the case for independent components, leads to possible computations of the distribution of the maximum $W = \max X_i$ or the sum $S = \sum X_i$, of the components of $\underline{X}_N$. Such results are of premier importance for the use of such models in statistical physics applications as well as in signal processing for problems involving statistical properties of extremes or time-averages as ensemble average estimators. This is being investigated. Finally, the potential use of this synthesis tool to generate independent copies of sets of hyper-parameters in Monte Carlo Markov Chain numerical schemes constitutes a natural track to investigate.

References

[1] B. A. Berg, *Markov chain Monte Carlo simulations and their statistical analysis*, World Scientific, 2004.

[2] D.G. Tzikas, A.C. Likas, and N.P. Galatsanos, "The variational approximation for bayesian inference," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 131–146, 2008.

[3] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[4] C.R. Dietrich and G.N. Newsam, "Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix," *SIAM J. Sci. Comput.*, vol. 18, no. 4, pp. 1088–1107, 1997.

[5] A. Dembo, C.L. Mallows, and L.A. Shepp, "Embedding nonnegative definite Toeplitz matrices in nonnegative definite circulant matrices, with applications to covariance estimation," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1206–1212, 1989.

[6] R.B. Davies and D.S. Harte, "Tests for Hurst effect," *Biometrika*, vol. 74, no. 4, pp. 95–101, 1987.

[7] A.T.A Wood and G. Chan, "Simulation of stationary Gaussian processes in $[0,1]^d$," *J. of Comput. and Graph. Stat.*, vol. 3, no. 4, pp. 409–432, 1994.

[8] G. Chan and A.T.A. Wood, "Simulation of stationary Gaussian vector fields," *Statistics and Computing*, vol. 9, no. 4, pp. 265–268, 1999.

[9] H. Helgason, V. Pipiras, and P. Abry, "Fast and exact synthesis of stationary multivariate Gaussian time series using circulant embedding," *Signal Processing*, vol. 91, no. 5, pp. 1123–1133, 2011.

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

[11] H. Helgason, V. Pipiras, and P. Abry, "Synthesis of multivariate stationary series with prescribed marginal distributions and covariance using circulant embedding," *Signal Processing*, vol. 91, no. 8, pp. 1741–1758, 2011.

[12] T. Schreiber and A. Schmitz, "Surrogate time series," *Physica D*, vol. 142, no. 3-4, pp. 346–382, 2000.

[13] P. Borgnat, P. Abry, and P. Flandrin, "Using surrogates and optimal transport for synthesis of stationary multivariate series with prescribed covariance function and non-gaussian joint-distribution," in *Proceeding of the IEEE Int. Conf. on Acoust. Speech and Sig. Proc. (ICASSP)*, Kyoto, 2012.

[14] M.-H. Chen, Q.-M. Shao, and J. G. Ibrahim, *Monte Carlo methods in Bayesian computation. Springer Series in Statistics*, Springer-Verlag, New York, 2000.

[15] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, "Exact solution of a 1d asymmetric exclusion model using a matrix formulation," *J Phys. A*, vol. 26, pp. 1493–1517, 1993.

[16] K. Mallick and S. Sandow, "Finite dimensional representations of the quadratic algebra: Applications to the exclusion process," *J. Phys. A: Math. Gen.*, vol. 30, pp. 4513, 1997.

[17] R A Blythe and M R Evans, "Nonequilibrium steady states of matrix-product form: a solver's guide," *Journal of Physics A: Mathematical and Theoretical*, vol. 40, no. 46, pp. R333–R441, 2007.

[18] F. Angeletti, E. Bertin, and P. Abry, "Matrix products for the synthesis of stationary time series with a priori prescribed joint distributions," in *Proceeding of the IEEE Int. Conf. on Acoust. Speech and Sig. Proc. (ICASSP)*, 2012.

[19] O. Cappé, E. Moulines, and T. Ryden, *Inference in Hidden Markov Models (Springer Series in Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[20] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

[21] J.A. Nocedal and S.J. Wright, *Numerical optimization*, Springer Series in Operations Research Series. Springer-Verlag New York, 1999.

Florian Angeletti

Florian Angeletti was born in Marseille, France in 1986. He received a Ph.D in theoretical physics from the École Normale Supérieure of Lyon in 2012. He is presently a postdoctoral research fellow at the NITheP, Stellenbosch University. His current research interests include the applications of statistical physics to signal processing and the study of exactly solvable models for disordered or non-equilibrium physical systems.

Eric Bertin

Eric Bertin was born in Chambéry, France in 1974. He received a degree from the Ecole Centrale Paris in 1998, a DEA (equivalent of a Master degree) from University Paris 7 in 1999 and a Ph.D degree in theoretical physics from University Paris 7 in 2003. He is presently a CNRS Research Associate at Physics dept. of Ecole Normale Supérieure de Lyon. He is the author of an introductory booklet on the statistical physics of complex systems. His current research interests include fundamental issues on the statistical characterization of non-equilibrium physical systems, as well as applications of statistical physics concepts and methods to other fields like statistical signal processing and modelling of social systems.

Patrice Abry

Patrice Abry was born in Bourg-en-Bresse, France in 1966. He received the degree of Professeur-Agrégé de Sciences Physiques, in 1989 at Ecole Normale Supérieure de Cachan and completed a PhD in Physics and Signal Processing, at Université Claude-Bernard University in Lyon in 1994. He is a CNRS Senior Scientist, at the Physics dept. of Ecole Normale Supérieure de Lyon, where he is in charge of the *Signal, systems and Physics* research team. Patrice Abry received the AFCET-MESR-CNRS prize for best PhD in Signal Processing for the years 93-94 and has been elected IEEE Fellow in 2011. He is the author of a book in French dedicated to wavelet, scale invariance and hydrodynamic turbulence and is also the coeditor of a book entitled ”Scaling, Fractals and Wavelets” (ISTE). His current research interests include wavelet-based analysis and modeling of scaling phenomena and related topics (self-similarity, stable processes, multifractal, 1/f processes, long-range dependence, local regularity of processes, infinitely divisible cascades, departures from exact scale invariance…). Beyond theoretical developments and contributions in multifractal analysis and stochastic process design, Patrice Abry shows a strong interest into real-world applications, ranging from hydrodynamic turbulence to the analysis and modeling of computer network teletraffic. He is also involved in the study of Heart Rate Variability both for adults and fetuses.