

# MATRIX PRODUCTS FOR THE SYNTHESIS OF STATIONARY TIME SERIES WITH A PRIORI PRESCRIBED JOINT DISTRIBUTIONS

Florian Angeletti, Eric Bertin, Patrice Abry

Physics Dept., CNRS, UMR5672, ENS de Lyon, France.

ens-lyon.fr/PHYSIQUE/, firstname.lastname@ens-lyon.fr

## ABSTRACT

Inspired from non-equilibrium statistical physics models, a general framework enabling the definition and synthesis of stationary time series with a priori prescribed and controlled joint distributions is constructed. Its central feature consists of preserving for the joint distribution the *simple* product structure it has under independence while enabling to input controlled and prescribed dependencies amongst samples. To that end, it is based on products of  $d$ -dimensional matrices, whose entries consist of valid distributions. The statistical properties of the thus defined time series are studied in details. Having been able to recast this framework into that of Hidden Markov Models enabled us to obtain an efficient synthesis procedure. Pedagogical well-chosen examples (time series with the same marginal distribution, same covariance function, but different joint distributions) aim at illustrating the power and potential of the approach and at showing how targeted statistical properties can be actually prescribed.

**Index Terms**— Time Series Synthesis, Joint Distribution, A priori Prescription, Hidden Markov Model

## 1. INTRODUCTION

In modern signal processing, the need to produce numerically random vectors whose joint distributions are fully prescribed and controlled is more and more common. Typical examples are Bayesian estimation (with Monte Carlo Markov Chain or variational based resolution schemes, cf. e.g., [1]) or estimation performance assessment (where estimators are benchmarked on time series with prescribed distributions). For univariate Gaussian stationary time series, the so-called *Circulant Embedded Matrix* synthesis procedure [2] is considered as the state of the art solution. For non Gaussian stationary time series, various approaches were proposed that aim at controlling both the marginal distribution and the covariance structure of the process (cf. e.g., [3, 4] and references therein for reviews). Often, such methods suffer from two major limitations: The joint distribution cannot be prescribed a priori and is hence not controlled and results as a consequence of the details of the synthesis procedure; Distributions that consist of mixtures of elementary laws cannot always eas-

ily be combined with prescribed covariance. The very general framework of Markov Chain simulation 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.

Instead, in the present contribution, the focus is on the synthesis of stationary time series whose joint distributions are explicitly prescribed and chosen a priori. It is based on a construction inspired from out-of-equilibrium statistical physics models (cf. e.g., [5, 6, 7]). It is founded on the central idea that the joint distribution of a vector  $\underline{X}_N = x_1, x_2, \dots, x_N$  is still written as a (oriented) product:  $p_{\underline{X}} \propto \prod_{k=1}^N R_d(x_k)$ , where, however the  $R_d$ s are not single distributions, but instead  $d$ -dimensional matrices of distributions. The key results of the present contribution consist of first deriving the general properties of this construction and then to specify it to a very interesting case specifically suited for stationary time series with all joint distributions prescribed (cf. Section 2). An efficient synthesis procedure is devised by explicitly recasting this framework into that of Hidden Markov Models (cf. Section 3). Finally, explicit examples are studied and simulated numerically, aiming at illustrating the power of the proposed approach: They consist of time series sharing the same complicated (mixture) marginal distribution, the same covariance functions, but different joint distributions (cf. Section 4).

## 2. JOINT DISTRIBUTIONS AS MATRIX PRODUCT

**General Framework.** Let us first define a general formalism for the construction of joint distribution functions of random vectors  $\underline{X}_N$ , inspired from non equilibrium statistical physics models (cf. e.g. [5, 6, 7]) and based on products of matrices of distributions. Let  $R_d(x)$  denote a  $d$ -dimensional matrix with entries

$$R_d(x)_{i,j} = \mathcal{E}_{i,j} \mathcal{P}(x)_{i,j}, \quad (1)$$

where  $\mathcal{P}(x)_{i,j}$  are valid distribution functions and  $\mathcal{E}_{i,j}$  arbitrary positive numbers, forming the matrix  $\mathcal{E}$ . Let  $A$  denote an arbitrary, but fixed strictly positive and non-random, ma-

trix and  $\mathcal{L}(M)$  the linear form applied to matrix  $M$  defined as  $\mathcal{L}(M) = \text{tr}(A^T M)$ .

Further, let  $\underline{X} \equiv \{X_n\}_{1 \leq n \leq N}$  denote a random vector, of chosen size  $N$ , explicitly defined via its joint distribution:

$$\mathbb{P}(x_1, \dots, x_N) = \frac{\mathcal{L}\left(\prod_k^N R_d(x_k)\right)}{\mathcal{L}(\mathcal{E}^N)}, \quad (2)$$

where  $\prod_k^N R_d(x_k) = R_d(x_1) \dots R_d(x_N)$  denotes the oriented product (i.e., the order of the factors is fixed and cannot be changed). It is straightforward to check that the joint choice of strictly positive entries for matrices  $A$  and  $\mathcal{E}$ , and of valid distribution functions for  $\mathcal{P}(x)_{i,j}$  is sufficient to ensure that Eq. (2) defines a valid joint distribution function.

From these definitions, calculations and matrix manipulations not reported here enabled us to derive a number of statistical properties of the vector  $\{X_n\}_{1 \leq n \leq N}$ . Its univariate (marginal) distributions and one-sample moments take explicit forms (with  $M(q) = \int_{\mathbb{R}} x^q R_d(x) dx$ ):

$$\mathbb{P}(X_k = x) = \frac{\mathcal{L}(\mathcal{E}^{k-1} R_d(x) \mathcal{E}^{N-k})}{\mathcal{L}(\mathcal{E}^N)}, \quad (3)$$

$$\mathbb{E}[X_k^q] = \frac{\mathcal{L}(\mathcal{E}^{k-1} M(q) \mathcal{E}^{N-k})}{\mathcal{L}(\mathcal{E}^N)}. \quad (4)$$

Furthermore, the joint  $p$ -sample moments read (with  $p \in N$ ,  $k_1 < \dots < k_p$  and  $q_r$  the order associated to the entry  $x_{k_r}$ ):

$$\mathbb{E}\left[\prod_{r=1}^p X_{k_r}^{q_r}\right] = \frac{\mathcal{L}\left(\mathcal{E}^{k_1-1} \left(\prod_r^{p-1} M(q_r) \mathcal{E}^{k_{r+1}-k_r-1}\right) M(q_s) \mathcal{E}^{N-k_s}\right)}{\mathcal{L}(\mathcal{E}^N)} \quad (5)$$

**Stationary time series.** Let us now focus on specific choices for matrices  $A$  and  $E$ , of interest here to construct stationary time series:

$$A_{i,j} = \frac{1}{d}, \quad \mathcal{E} = \alpha I_d + \beta J_d, \quad \text{with } \alpha + \beta = 1, \quad (6)$$

$I_d$  is the  $d$ -dimensional unity matrix and  $J_d \in M_d(\mathbb{R})$  defined as:

$$J_d = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix}$$

Such choices for  $A$  and  $\mathcal{E}$  enabled us to show (full calculations not reported here) that:

$$\forall \mathcal{E}, n \in \mathbb{N} \quad \mathcal{L}(\mathcal{E}^n) = 1, \quad (7)$$

$$\forall M, \mathcal{E}, (n, r) \in \mathbb{N}^2, \quad \frac{\mathcal{L}(M \mathcal{E}^r)}{\mathcal{L}(\mathcal{E}^n)} = \mathcal{L}(M). \quad (8)$$

This further enables us to obtain that the results above can be specified as:

$$\mathbb{P}(X_k = x) = \frac{1}{d} \sum_{i,j} \mathcal{E}_{i,j} \mathcal{P}_{i,j}(x), \quad (9)$$

$$\mathbb{E}[X_k^q] = \mathcal{L}(M(q)), \quad (10)$$

$$\mathbb{E}\left[\prod_{r=1}^p X_{k_r}^{q_r}\right] = \mathcal{L}\left(\left(\prod_r^{p-1} M(q_r) \mathcal{E}^{k_{r+1}-k_r-1}\right) M(q_p)\right). \quad (11)$$

These relations clearly indicate that the vector  $\{X_n\}_{1 \leq n \leq N}$  can now be regarded as a stationary time series: All joint statistics depend only on time differences,  $k_{r+1} - k_r$ . Both its marginal and joint probabilities are prescribed by the choices of  $\mathcal{E}$  and of the  $\mathcal{P}_{i,j}$ . Also, Eq. (11) constitutes a key result with respect to applications as it clearly shows that the joint statistics of order  $q$  of the time series can be prescribed by the sole selection of a suitable  $M(q)$  matrix. This will be explicitly used in Section 4.

The form of the covariance function can be further studied. The eigenvalues of  $\mathcal{E}$  read  $\lambda_k = \alpha + \beta e^{\frac{2\pi i k}{d}}$  and can be rewritten as  $\lambda_k = e^{-\frac{1}{\tau_k}} e^{\pm i \frac{2\pi}{T_k}}$ ,  $k = 1, \dots, \lfloor d/2 \rfloor$  (where  $\lfloor z \rfloor$  stands for the integer part of  $z$ ). For ease of notations, let  $(C_M)_k = \sum_l M_{l,k}$ ,  $(L_M)_k = \sum_c M_{k,c}$  and let  $\mathcal{F}$  denote the (non-normalized) discrete Fourier transform. Then, detailed calculations enabled us to show that, for any  $q$ ,

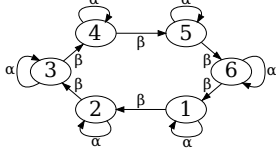
$$\mathbb{E}[X_0^q X_t^q] - \mathbb{E}[X_0^q] \mathbb{E}[X_t^q] = \sum_{k=1}^{\lfloor d/2 \rfloor} m_k \Re \left\{ \mathcal{F}(L_{M(q)})_k \overline{\mathcal{F}(C_{M(q)})_k} e^{-\frac{t-1}{\tau_k}} e^{i \frac{2\pi(t-1)}{T_k}} \right\}, \quad (12)$$

where  $m_k = 1$  if  $2k = d$  and 2 otherwise. This result shows that  $\tau_k$  and  $T_k$  are characteristic dependence lengths and periodicities that depend on the joint choices of  $d$  and  $\alpha$ . More precisely, from the definition of  $\tau_k = -1/\ln|\lambda_k|$ , we can show that  $\tau_k \approx_{\alpha\beta \rightarrow 0} [\alpha\beta (1 - \cos \frac{2\pi k}{d})]^{-1}$ . Enlarging the dimension  $d$  of the matrices  $R_d$  increases both the number of distinct dependence lengths as well as the ratio of the smallest to the largest such characteristic lengths, which varies asymptotically as  $(d/2\pi)^2/2$ .

Therefore, choosing  $d$ ,  $\alpha$ , hence  $\mathcal{E}$ , and the sequence of  $\mathcal{P}_{i,j}$  enables to select the marginal and joint distributions according to given targets. Section 4 will illustrate the potential of the method.

### 3. SYNTHESIS: HIDDEN MARKOV CHAIN

While the compact form of  $\mathbb{P}(\underline{X})$  in Eq. (2) constitutes a general framework for the analytical derivation of numerous statistical properties of  $\underline{X}$ , it provides few insights with respect to its numerical synthesis. To address this issue, a reformulation of Eq. (2) into the form of a Hidden Markov Model is now devised.



**Fig. 1. Transition graph.** For  $d = 6$  and  $\mathcal{E} = \alpha I_d + \beta J_d$

First, extending to product of  $N$  matrices the fact that the entries of the matrix product  $(ABC)$  reads  $(ABC)_{i,j} = \sum_{k,l} a_{i,k} b_{k,l} c_{l,j}$ , we have been able to recast Eq. (2) into:

$$\mathbb{P}(x_1, \dots, x_N) = \sum_{\underline{\Gamma}} \kappa(\underline{\Gamma}) \prod_{k=1}^N \mathcal{P}(x_k)_{\Gamma_{k-1}, \Gamma_k}, \quad (13)$$

with  $\underline{\Gamma} \equiv \{\Gamma_0, \dots, \Gamma_k, \dots, \Gamma_N\} \in [1, \dots, d]^{N+1}$  and

$$\kappa(\underline{\Gamma}) = \frac{A_{\Gamma_0, \Gamma_N}}{\mathcal{L}(\mathcal{E}^N)} \prod_{k=1}^N \mathcal{E}_{\Gamma_{k-1}, \Gamma_k}. \quad (14)$$

It is straightforward to verify that  $\sum_{\underline{\Gamma}} \kappa(\underline{\Gamma}) = 1$ , hence  $\kappa(\underline{\Gamma})$  can be interpreted as the probability function of  $\underline{\Gamma}$ . Moreover, Eq. (13) shows that  $\mathbb{P}(\underline{X})$  can be read as a  $\kappa(\underline{\Gamma})$ -weighted mixture of laws, each defined as the product  $\prod_{k=1}^N \mathcal{P}(x_k)_{\Gamma_{k-1}, \Gamma_{k+1}}$ .

Second, from Eq. (14), using the special form of  $A$  and  $\mathcal{E}$ , it can be shown that:

$$\begin{aligned} \mathbb{P}(\Gamma_k = j | \Gamma_0 = \gamma_0, \dots, \Gamma_{k-1} = \gamma_{k-1}) &= \\ \frac{\mathcal{E}_{\gamma_0, \gamma_1} \dots \mathcal{E}_{\gamma_{k-1}, j} \sum_{\Gamma} \mathcal{E}_{j, \Gamma_{k+1}} \dots \mathcal{E}_{\Gamma_{N-1}, \Gamma_N}}{\mathcal{E}_{\gamma_0, \gamma_1} \dots \mathcal{E}_{\gamma_{k-2}, \gamma_{k-1}} \sum_{\Gamma} \mathcal{E}_{\gamma_{k-1}, \Gamma_k} \dots \mathcal{E}_{\Gamma_{N-1}, \Gamma_N}} & \quad (15) \\ = \mathcal{E}_{\gamma_{k-1}, j} = \mathbb{P}(\Gamma_k = j | \Gamma_{k-1} = \gamma_{k-1}) \end{aligned}$$

and hence that  $\underline{\Gamma}$  is a homogeneous  $d$ -state Markov chain, with transition probability matrix:

$$\mathbb{P}(\Gamma_{k+1} = j | \Gamma_k = i) = \mathcal{E}_{i,j}. \quad (16)$$

Hence,  $\mathcal{E}$  can be recast as the transition matrix of the Markov chain  $\underline{\Gamma}$ , as illustrated in Fig. 1.

Third, the final step required for numerical synthesis is to derive the initial distribution for  $\Gamma_0$ . Eq. (14) enabled us to show that it should follow a uniform distribution:

$$\mathbb{P}(\Gamma_0 = i) = \frac{1}{d}. \quad (17)$$

Therefore, the time series  $\underline{X}_k$ , as defined from Eq. (1), can be read and interpreted as a Hidden Markov Model, with 2 hidden states: the current state  $\Gamma_k$  and the previous state  $\Gamma_{k-1}$ .

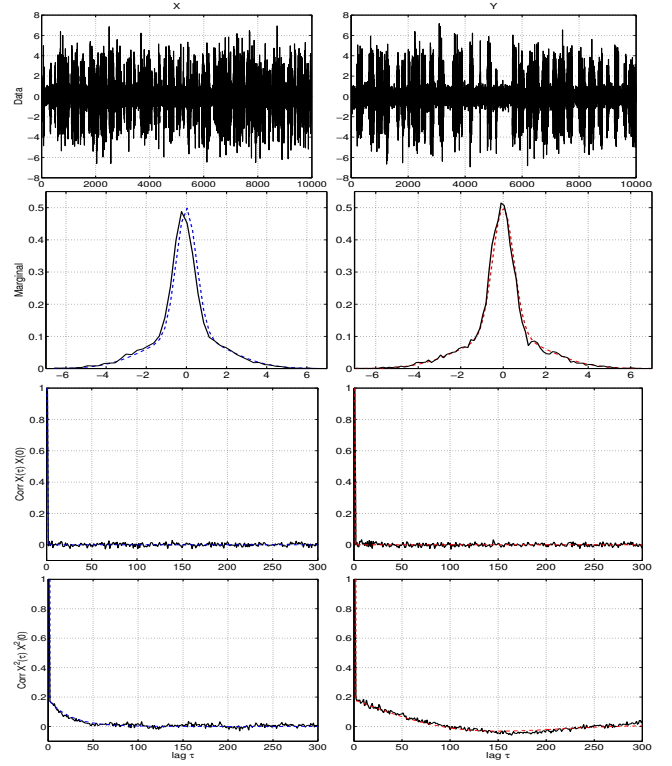
Combining Eqs. (16) and (17), a synthesis algorithm can be sketched as follows:

Step 1 : Initialization: Use Eq. (17) to generate the state  $\Gamma_0$ .

Step 2 : Iteration on  $k$ :

- i) Choose at random state  $\Gamma_k$ , using Eq. (16),
- ii) Generate  $X_k$  according to  $\mathcal{P}_{\Gamma_{k-1}, \Gamma_k}$ .

## 4. ILLUSTRATIONS



**Fig. 2. Numerical Synthesis.** Two different time series with the same marginal distribution (mixture of two Gaussians,  $\sigma_1 = 0.5$ ,  $\sigma_2 = 2$ ), the same covariance functions (chosen as  $\delta$  functions) but different covariance functions for their squares, hence different prescribed joint distributions ( $\alpha = 0.98$ ). Left side  $\underline{X}$ , right side  $\underline{Y}$ . First line, one realization of the time series. Then, from top to bottom, estimated (solid black lines) and theoretical (dashed colored lines) marginals, correlation functions and correlation functions for the squared time series.

To illustrate the potential of the proposed time series theoretical construction and synthesis procedure, a pedagogical example is proposed and the construction of the targeted statistical properties is devised: It consists of a pair  $(\underline{X}, \underline{Y})$  of processes sharing the same marginal distributions (a mixture of two Gaussian laws), the same autocovariance functions (a  $\delta$  function, i.e., no correlation), but different joint distributions, hence different higher order statistic dependences. In this example, the dependence of the 4–th order statistics will be prescribed.

We select  $d = 6$  and  $\mathcal{E} = \alpha I_d + \beta J_d$ . For the sake of simplicity, the univariate (or marginal) distributions is set to be a mixture of two Gaussian distributions:  $p(x) = \frac{1}{2} (\mathcal{N}_{0, \sigma_1}(x) + \mathcal{N}_{0, \sigma_2}(x))$ . For that case, it is natural to envisage that the matrix  $\mathcal{P}$  consists only of Gaussian distributions.

Note however that this is not the sole possibility. This restriction implies that  $\mathcal{N}_{0,\sigma_1}$  and  $\mathcal{N}_{0,\sigma_2}$  must appear exactly three times both in the principal and upper circular diagonals of  $\mathcal{P}$ . In other words, selecting the univariate distribution fixes the number of occurrences of each law in matrix  $\mathcal{P}$ , but not their position in the matrix.

Moreover, to make the example convincing, we chose the covariance functions of both  $\underline{X}$  and  $\underline{Y}$  to be  $\delta$ -functions. These two time series will hence have no autocorrelation but (higher-order) dependencies. To achieve this, it is sufficient that matrix  $M(1)$  is a zero matrix (cf. Eq. (11)). In our example, this is automatically obtained due to the choice of zero-mean distributions:  $\mathcal{N}_{0,\sigma_1}$  and  $\mathcal{N}_{0,\sigma_2}$ .

These two time series have hence the same marginal distributions and covariance functions. They are yet different by construction as we now impose that they have different joint distributions. Taking into account the circularity of the transition graph in Fig. 1, there exist 80 distinct choices of  $\mathcal{P}$ , consisting of exactly 6 entries set to  $\mathcal{N}_{0,\sigma_1}$  and 6 entries set to  $\mathcal{N}_{0,\sigma_2}$ , leading to different joint distributions. Making use of Eq. (12) and of the explicit form of the eigenvalues of  $\mathcal{E}$  for  $d = 6$  enables us to show that  $\mathbb{E}[X_0^q X_t^q] - \mathbb{E}[X_0^q] \mathbb{E}[X_t^q]$ , for any  $q$ , stems from the superimposition of three distinct exponential terms whose characteristic lengths, we derived asymptotically:  $\tau_1 \underset{\beta \rightarrow 0}{\approx} \frac{2}{\beta}$ ,  $\tau_2 \underset{\beta \rightarrow 0}{\approx} \frac{3}{2\beta}$  and  $\tau_3 \underset{\beta \rightarrow 0}{\approx} \frac{1}{2\beta}$ . The actual choice for the entries of  $\mathcal{P}$  fixes the complex valued coefficients, associated to each of these correlation lengths (cf. Eq. 12). As demonstrative examples, we selected for  $X$ :  $\mathcal{P}_{i,j}^X = \mathcal{N}_{0,\sigma_1}$  if  $i$  even, and  $\mathcal{P}_{i,j}^X = \mathcal{N}_{0,\sigma_2}$  if  $i$  odd. Applying Eq. (12) to that choice, shows that the sole  $\tau_3$ , i.e., the shortest correlation length, appears in the auto-covariance of  $\underline{X}^2$  (the complex-valued coefficients associated to the two other correlation lengths are forced to 0 by the specific choice made here):  $\mathbb{E}[X_0^2 X_t^2] - \mathbb{E}[X_0^2] \mathbb{E}[X_t^2] = \frac{(\sigma_1 - \sigma_2)^2}{4} (\alpha - \beta)^t$ . Conversely, choosing  $\mathcal{P}_{i,j}^Y = \mathcal{N}_{0,\sigma_1}$  if  $1 \leq i \leq 3$  and  $\mathcal{P}_{i,j}^Y = \mathcal{N}_{0,\sigma_2}$  if  $4 \leq i \leq 6$  leads to  $\mathbb{E}[X_0^2 X_t^2] - \mathbb{E}[X_0^2] \mathbb{E}[X_t^2] = \frac{(\sigma_1 - \sigma_2)^2}{d^2} [(\alpha - \beta)^t + 4\Re\{ (2 - \beta - \sqrt{3i}\beta)\lambda_1^{t-1} \}]$ . In that case, both  $\tau_1$  and  $\tau_3$  contribute to the autocovariance of  $\underline{Y}^2$ . However,  $\tau_1 \approx 4\tau_3$  is dominant at large  $t$  and hence constitutes the leading term. Using the synthesis procedure devised in Section 3 for this pair of examples yields Fig. 2, showing for times series  $\underline{X}$  (left column) and  $\underline{Y}$  (right column), a particular realization, the estimated and targeted univariate distributions, covariance functions and covariance functions for the squared time series (from top to bottom). It clearly shows that  $\underline{X}$  and  $\underline{Y}$  have the same marginal and covariance but different joint distributions (as targeted).

Using the same construction procedure, other pairs of examples with the same marginals, same (non  $\delta$ -) autocovariance functions but different joint distributions could as easily be devised. Mixture of Gaussians are used here by convenience, but mixtures of any other valid distributions could just as easily be reached.

## 5. CONCLUSION AND PERSPECTIVES

Inspired from *Statistical Physics* models, a general framework enabling us to define the joint distributions of a random vector  $\underline{X}$  has been described. It has then been specified to the definition of stationary time series, with control of their joint distributions and explicit derivation of numerous of their statistical properties. A remapping onto a Hidden Markov Model enabled us to devise an efficient synthesis procedure, available upon request (in MATLAB). Constructive examples aiming at showing the potential of the tools were proposed. This general framework will be further explored by departing from the restrictive choices made here for the matrices  $A$  and  $\mathcal{E}$ . This should notably enable us to define (and synthesize numerically) vectors of dependent variables with different marginal distributions and intricate dependencies. Notably, the intriguing and promising special cases of non-diagonalisable matrices  $\mathcal{E}$  are expected to offer a much larger versatility in the form of dependencies that can actually be reached. This is under current investigations.

J.-Y. Tournet and N. Dobigeon are gratefully acknowledged for fruitful discussions.

## 6. REFERENCES

- [1] 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.
- [2] 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.
- [3] M. Grigoriu, *Applied Non-Gaussian Processes*, Prentice-Hall, 1995.
- [4] 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.
- [5] 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.
- [6] K. Mallick and S. Sandow, “Finite dimensional representations of the quadratic algebra: Applications to the exclusion process,” *J. Phys. A*, vol. 30, pp. 4513, 1997.
- [7] 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.