Matrix-correlated random variables: A statistical physics and signal processing duet

Florian Angeletti

NITheP National Institute for Theoretical Physics

Work in collaboration with Hugo Touchette, Patrice Abry and Eric Bertin.

4 December 2015

Random vectors

Random vectors in signal processing

- Joint probability density : $\mathbb{P}(x_1, \ldots, x_n)$
- i.i.d.: $\mathbb{P}(x_1, \ldots, x_n) = f(x_1) \ldots f(x_n)$
- generalization to non-i.i.d. ?

Random vectors

Random vectors in signal processing

- Joint probability density : $\mathbb{P}(x_1, \ldots, x_n)$
- i.i.d.: $\mathbb{P}(x_1, \ldots, x_n) = f(x_1) \ldots f(x_n)$
- generalization to non-i.i.d. ?

Out-of-equilibrium physics

- Asymmetric Simple Exclusion Process model [Derrida et al., *J. Phys. A*, 1993]
- $p(x_1, \ldots, x_n) \propto \mathcal{R}(x_1) \ldots \mathcal{R}(x_n)$: f scalar $\Rightarrow \mathcal{R}$ matrix
- Preserved product structure.
- Signal processing application?

Matrix-correlated random variables

$$p(x_1, \ldots, x_n) = \frac{\mathcal{L}\left(\mathcal{R}(x_1)\ldots\mathcal{R}(x_n)\right)}{\mathcal{L}\left(\mathcal{E}^n\right)}$$

- linear form $\mathcal{L}$: $\mathcal{L}(M) = \mathrm{tr}(\mathcal{A}^T M)$
  - $\mathcal{A}$: $d \times d$ positive matrix
- $\mathcal{R}(x)$: $d \times d$ positive matrix function
  - structure matrix

$$\mathcal{E} = \int_{\mathbb{R}} \mathcal{R}(x)dx$$

  - probability density function matrix

$$\mathcal{R}_{i,j}(x) = \mathcal{E}_{i,j}\mathcal{P}_{i,j}(x)$$

- $d > 1$: Non-commutativity $\Longrightarrow$ Correlation

Objectives

Mathematical model

$$p(x_1, \ldots, x_n) \approx \mathcal{R}(x_1) \cdots \mathcal{R}(x_n)$$

Study the statistical properties of theses models

- Hidden Markov model representation
- Topology induces correlation

Signal processing applications?

- Large deviation functions
- Limit distributions for the sums
- Limit distributions for the extremes

Correlation

- Product structure: $p(x_1, \ldots, x_n) = \frac{\mathcal{L}(\mathcal{R}(x_1)\ldots\mathcal{R}(x_n))}{\mathcal{L}(\mathcal{E}^n)}$
- Moment matrix: $\mathcal{Q}(q) = \int x^q R(x) dx$

$$\mathbb{E}\left[X_k^p\right] = \frac{\mathcal{L}\left(\mathcal{E}^{k-1}\mathcal{Q}(p)\,\mathcal{E}^{n-k}\right)}{\mathcal{L}\left(\mathcal{E}^n\right)}$$

$$\mathbb{E}\left[X_k X_l\right] = \frac{\mathcal{L}\left(\mathcal{E}^{k-1}\mathcal{Q}(1)\,\mathcal{E}^{l-k-1}\mathcal{Q}(1)\,\mathcal{E}^{n-l}\right)}{\mathcal{L}\left(\mathcal{E}^n\right)}$$

$$\mathbb{E}\left[X_k X_l X_m\right] = \frac{\mathcal{L}\left(\mathcal{E}^{k-1}\mathcal{Q}(1)\,\mathcal{E}^{l-k-1}\mathcal{Q}(1)\,\mathcal{E}^{m-l-1}\mathcal{Q}(1)\,\mathcal{E}^{n-m}\right)}{\mathcal{L}\left(\mathcal{E}^n\right)}$$

$$\ldots$$

Stationarity

Translation invariance:
$p(X_{k_1} = x_1, \ldots, X_{k_l} = x_l) = p(X_{c+k_1} = x_1, \ldots, X_{c+k_l} = x_l)$

Sufficient condition

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

$$\forall M, \quad \mathcal{L}(M\mathcal{E}) = \mathcal{L}(\mathcal{E}M)$$

- $p(X_k = x) = \frac{\mathcal{L}\left(\mathcal{R}(x)\mathcal{E}^{n-1}\right)}{\mathcal{L}(\mathcal{E}^n)}$
- $p(X_k = x, X_l = y) = \frac{\mathcal{L}\left(\mathcal{R}(x)\mathcal{E}^{l-k-1}\mathcal{R}(y)\mathcal{E}^{n-|l-k|-1}\right)}{\mathcal{L}(\mathcal{E}^n)}$
- $p(X_k = x, X_l = y, X_m = z) = \frac{\mathcal{L}\left(\mathcal{R}(x)\mathcal{E}^{l-k-1}\mathcal{R}(y)\mathcal{E}^{m-l-1}\mathcal{R}(z)\mathcal{E}^{n-|m-k|-1}\right)}{\mathcal{L}(\mathcal{E}^n)}$

Numerical generation

$$p(x_1,\ldots,x_n) = \frac{\mathcal{L}\left(\mathcal{R}(x_1)\ldots\mathcal{R}(x_n)\right)}{\mathcal{L}\left(\mathcal{E}^n\right)}$$

How do we generate a random vector X for a given triple $(\mathcal{A},\mathcal{E},\mathcal{P})$?

- Expand the matrix product

$$\mathcal{L}\left(\mathcal{E}^n\right) p(x_1,\ldots,x_n) = \sum_{\Gamma\in\{1,\ldots,d\}^{n+1}} \mathcal{A}_{\Gamma_1,\Gamma_{n+1}} \mathcal{E}_{\Gamma_1,\Gamma_2} \mathcal{P}_{\Gamma_1,\Gamma_2}(x_1)\ldots\mathcal{E}_{\Gamma_n,\Gamma_{n+1}}\mathcal{P}_{\Gamma_n,\Gamma_{n+1}}(x_n)$$

$$p(x_1,\ldots,x_n) = \sum_{\Gamma} P(\Gamma)P(\underline{X}|\Gamma)$$

- Γ, hidden Markov chain

Hidden Markov Model representation

Hidden Markov Chain

$$p(\Gamma) = \frac{\mathcal{A}_{\Gamma_1,\Gamma_{n+1}}}{\mathcal{L}\left(\mathcal{E}^n\right)} \prod_k \mathcal{E}_{\Gamma_k,\Gamma_{k+1}}$$

Conditional pdf $(X|\Gamma)$

$$p(X_k = x|\Gamma) = \mathcal{P}_{\Gamma_k,\Gamma_{k+1}}(x)$$

- $\mathcal{E}$ non-stochastic $\implies$ Non-homogeneous markov chain
- Specific non-homogeneous Hidden Markov model:
  - Hidden Markov Model $\not\Rightarrow$ Matrix representation

Dual representation

Matrix representation

- Algebraic properties
- Statistical properties computation

Hidden Markov Model

- 2-layer model: correlated layer + independant layer
- Efficient synthesis

Correlation and Jordan decomposition

$$\mathbb{E}\left[X_k X_l\right] = \frac{\mathcal{L}\left(\mathcal{E}^{k-1}\mathcal{Q}\left(1\right)\mathcal{E}^{l-k-1}\mathcal{Q}\left(1\right)\mathcal{E}^{n-l}\right)}{\mathcal{L}\left(\mathcal{E}^{n}\right)}$$

- The dependency structure of X depends on the behavior of $\mathcal{E}^n$
- λ_c eigenvalues of $\mathcal{E}$ ordered by their real parts $\Re(\lambda_1)\Re(\lambda_2) > \cdots > \lambda_r$
- $J_{m,p}$ Jordan block associated with eigeivalue λ_m

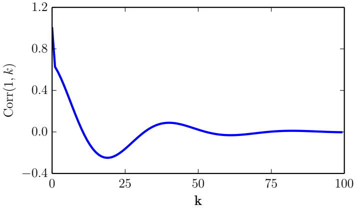
$$\mathcal{E} = B^{-1}\begin{pmatrix} J_{1,1} & 0 & 0 \\ & \ddots & \\ 0 & & J_{m,p} \end{pmatrix} B, \quad J_{m,s} = \begin{pmatrix} \lambda_m & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_m \end{pmatrix}$$

Dependency structure

Case 1: Short-range correlation

- λ_2 exists:

$\mathbb{E}[X_k X_l] - \mathbb{E}[X_k]\mathbb{E}[X_l] \approx \sum \alpha_m \frac{\lambda_m}{\lambda_1}^{|k-l|}$

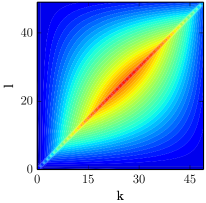


Case 2: Constant correlation

- More than one block $J_{1,s}$: Constant correlation term

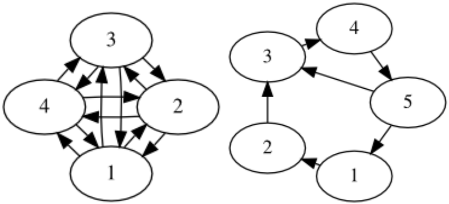
Case 3: Long-range correlation

- $J_{1,s}$ with size $p > 1$:
  - $\mathbb{E}[X_k X_l] - \mathbb{E}[X_k]\mathbb{E}[X_l] \approx P(\frac{k}{n}, \frac{k-l}{n}, \frac{l}{n}), \quad P \in \mathbb{R}[X, Y, Z]$



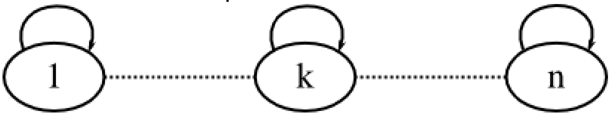
Short-range correlation: Ergodic chain

- $\mathcal{E}$ irreducible $\iff$ Γ ergodic Irreducible matrix $\mathcal{E}$ $\iff$ $G(\mathcal{E})$ is strongly connected
- Short-range correlation



Constant correlation: Identity $\mathcal{E}$

- Disconnected components:

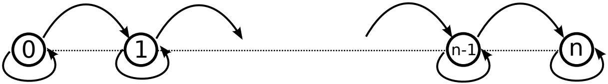


$$\mathcal{E} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

- The chain Γ is trapped inside its starting state
- Constant correlation:
  - $\mathbb{E}[X_k X_l] - \mathbb{E}[X_k]\,\mathbb{E}[X_l] = \frac{\mathcal{L}\left(\mathcal{Q}(1)^2\right) - \mathcal{L}(\mathcal{Q}(1))^2}{\mathcal{L}(\mathcal{E})}$

Long-range correlation: Linear irreducible $\mathcal{E}$

- Irreversible transitions:



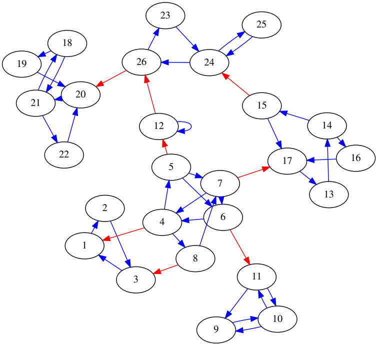
$$\mathcal{E} = \begin{pmatrix} 1 & \epsilon & 0 & \\ & \ddots & \ddots & \\ 0 & & 1 & \end{pmatrix}$$

- The chain Γ can only stay in its current state or jump to the next
- All chains with a non-zero probability and the same starting and ending points are equiprobable
- Polynomial correlation:

$$\mathbb{E}\left[X_k X_l\right] \approx \sum_{r+s+t=d-1} c_{r,s,t} \left(\frac{k}{n}\right)^r \left(\frac{l-k}{n}\right)^s \left(1-\frac{l}{n}\right)^t$$

General shape of $\mathcal{E}$

$$\mathcal{E} = \begin{pmatrix} I_1 & * & T_{k,l} \\ & \ddots & * \\ 0 & & I_r \end{pmatrix}$$



- Irreducible blocks I_k
- Irreversibles transitions $T_{k,l}$
- Correlation: Mixture of short-range, constant and long-range correlations

Summary

- Short-range correlation $\implies$ Strongly connected component of size $s > 1$
- More than one weakly connected component $\implies$ Constant correlation
- Polynomial correlation $\implies$ More than one strongly connected component

Necessary but non sufficient conditions

Synthesis

$$p(x_1, \ldots, x_n) = \frac{\mathcal{L}\left(\mathcal{R}(x_1) \ldots \mathcal{R}(x_n)\right)}{\mathcal{L}\left(\mathcal{E}^n\right)}$$

How to choose

- d ?
- $\mathcal{E}$?
- $\mathcal{P}$?
- $\mathcal{A}$?

Constraints

Classical constraints

- Marginal distribution: $\mathbb{P}_S$
- Autocovariance: $c_{1,1} \equiv \mathbb{E}[X_0 X_t] - \mathbb{E}[X_0]\,\mathbb{E}[X_t]$
- Higher-order dependencies:
  $c_{q_1,q_2}(t) \equiv \mathbb{E}\left[X_0^{q_1} X_t^{q_2}\right] - \mathbb{E}\left[X_0^{q_1}\right]\mathbb{E}\left[X_t^{q_2}\right]$
- Limitations: sum of r expoinential time scales θ_k with amplitudes $\beta(q_1, q_2)$

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

Choice of d, $\mathcal{A}$, $\mathcal{E}$, $\mathcal{P}$

$$\mathcal{A} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}, \quad J_d = \begin{pmatrix} 0 & 1 & \\ & \ddots & \ddots \\ 1 & & 0 \end{pmatrix}, \quad \mathcal{E} = \sum_k \alpha_k J_d^k$$

- Stationnarity: $[\mathcal{A}^T, \mathcal{E}] = 0$
- Dependencies: $\underline{\alpha} = \mathcal{F}(\underline{\tilde{\theta}})$

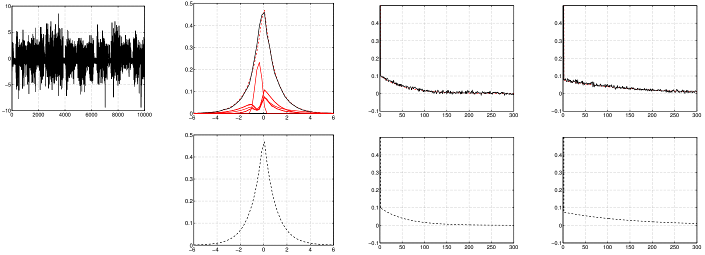
Objectives $\Rightarrow$ Free parameters

- $r \Rightarrow d$
- $\theta \Rightarrow \alpha$
- $\beta \Rightarrow M(q)$
- $\mathbb{P}_S \Rightarrow \mathcal{P}$

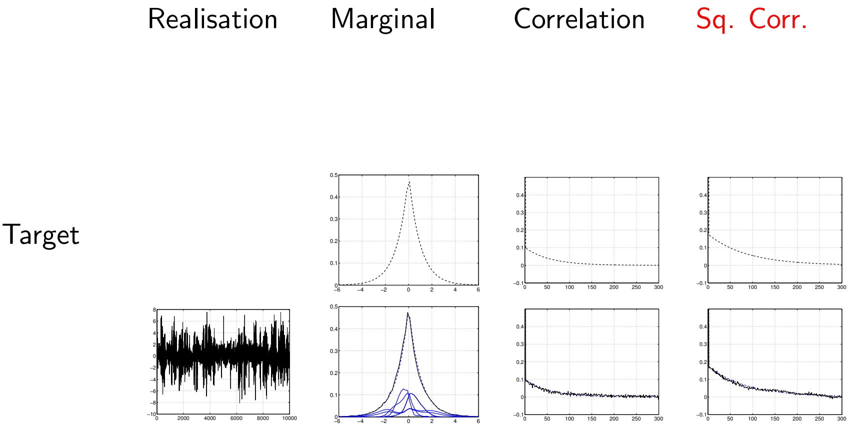
Stationary time series examples

Realisation Marginal Correlation Sq. Corr.

Target



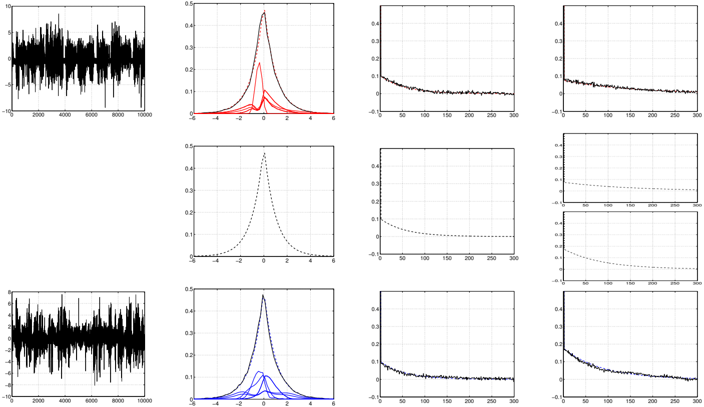
Stationary time series examples



Stationary time series examples

	Realisation	Marginal	Correlation	Sq. Corr.

Target



Random vector examples

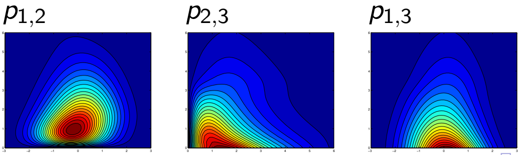
- $d = 3$
- Generalization de la structure en produit de matrice

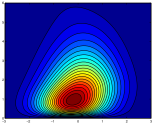
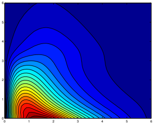
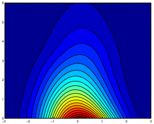
$$p(x_1, x_2, x_3) = \frac{\mathcal{L}\left(\mathcal{R}_1(x_1)\mathcal{R}_2(x_2)\mathcal{R}_3(x_3)\right)}{\mathcal{L}\left(\mathcal{E}^n\right)}$$

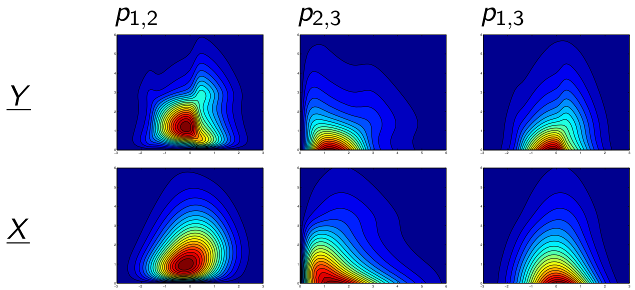
- Distribution marginale choisies a priori :
  - p_1 gaussian
  - p_2 gamma distribution $\alpha = 2$
  - p_3 exponential distribution
- Inter-covariance $\mathbb{E}\left[X_i X_j\right]$ prescribed

Distribution partielle bivariée :

$\underline{X}$

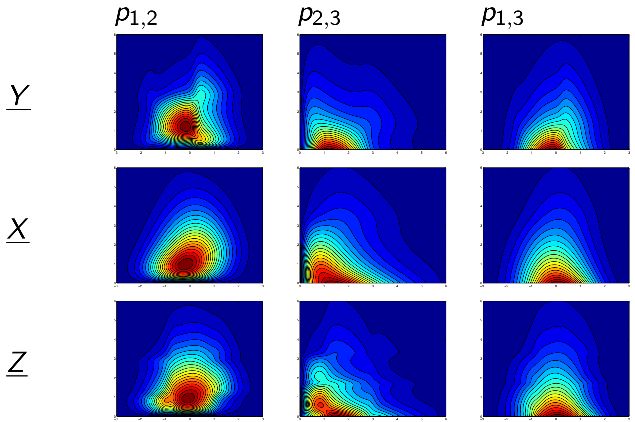


	$p_{1,2}$	$p_{2,3}$	$p_{1,3}$
$\underline{X}$			



$\underline{X}$, $\underline{Y}$

- Same marginal
- Same correlation
- Sq. corr. distinct



$\underline{X}$, $\underline{Y}$

- Same marginal
- Same correlation
- Sq. corr. distinct
- (joint pdf distinct)

$\underline{X}$, $\underline{Z}$

- Same marginals
- Same correlations
- Same sq. corr.
- joint pdf distincts

- Precise control over dependencies structure
- Synthesis: both stationary time series and random vectors

Random vector sum

Sum

$$S(\underline{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

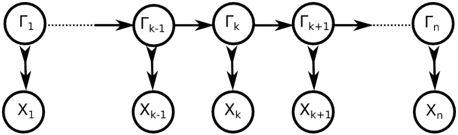
Correlated random variables

- Law of large numbers?
- Central limit theorem?
- Large deviations?

Two paths:

- Hidden Markov chain representation
- Matrix representation

Hidden Markov path



- $\nu_{i,j}$ fraction of $(i \to j)$-transition:

$$\underline{\nu} = \left(\frac{\mathrm{card}\{k / \Gamma_k = i, \Gamma_{k+1} = j\}}{n} \right)_{i,j}$$

- $S(\underline{X}|\Gamma)$: sum of sums of $\mathrm{i.i.d.}$ random variables:

$$S(\underline{X}|\Gamma) = \sum_{i,j} \sum_{k=1}^{n\nu_{i,j}} (X_k | i,j) \equiv S(\underline{X}|\underline{\nu})$$

- Standard convergence theorem (law of large numbers or central limit theorem)

Layer combination

$$p(S(\underline{X}) = s) = \sum_{\underline{\nu}} p(\underline{\nu}) p(S(\underline{X}|\underline{\nu}) = s)$$

Limit distribution for $S(\underline{X}|\underline{\nu})$

$$p(S(\underline{X}|\underline{\nu}) = s)$$

$+$

Limit distribution for $\underline{\nu}$

$$p(\underline{\nu})$$

$$\Downarrow$$

Limit distribution for $S(\underline{X})$

$$p(S(\underline{X}) = s)$$

Distribution of ν

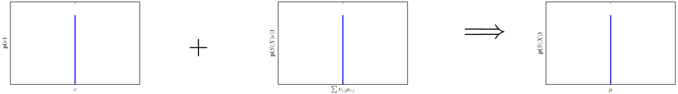
- How $\underline{\nu}$ is distributed?
- Difficulty: Γ non-homogeneous Markov chain.

Three important subclasses:

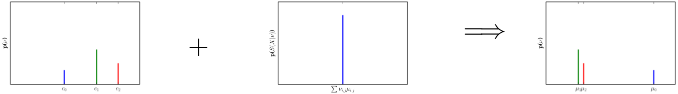
- Irreducible $\mathcal{E}$ (short-range correlation)
  - ν converges towards a dirac distribution
- Identity $\mathcal{E}$ (constant correlation)
  - ν converges towards a discrete mixture of dirac distributions
- Linear irreversible $\mathcal{E}$ (long-range polynomial correlation)
  - ν converges towards a uniform distribution on a d-simplex

Limits laws for core examples

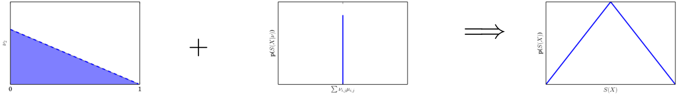
- Irreducible $\mathcal{E}$ (short-range correlations):
  - Standard limit laws



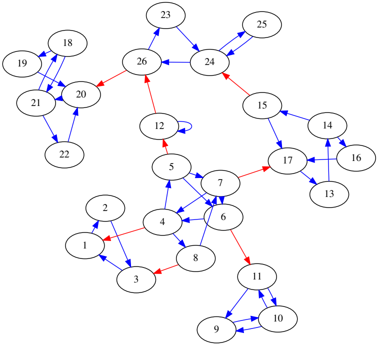
- Identity $\mathcal{E}$ (constant correlation):
  - Discrete mixture of standard limit laws:



- Linear irreversible $\mathcal{E}$ (long-range correlation):
  - Continuous mixture of limit laws



General case



- Combinations of three core behaviors
  - Irreducible blocks: Fast convergence to the stationary state : dirac distribution
  - Separated componentes : discrete mixture
  - Irreversible transitions: continuous mixture
- Limit laws :
  - Discrete mixture of continuous mixture of standard limits distributions

Conclusion

- Three kind of correlation:
  - Exponential short-range correlation
  - Constant correlation
  - Polynomial long-range correlation
- Synthesis of stationary time series with controlled correlations.
- Extension of the law of large numbers and the central limit theorems:
  - Long-range correlation : Continuous and discrete mixture of standard limit laws

Perspective

- Extreme statistics
- Physical model
- Infinite dimension, higher tensor order