



On the generation of colored non-Gaussian time sequences

Jacques Lavergnat

► To cite this version:

| Jacques Lavergnat. On the generation of colored non-Gaussian time sequences. 2016. hal-01399446

HAL Id: hal-01399446

<https://hal.science/hal-01399446>

Preprint submitted on 18 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the generation of colored non-Gaussian time sequences

Jacques Lavergnat

Juillet 2016

RÉSUMÉ

Une revue des principales méthodes de génération d'une série temporelle stationnaire, colorée non-Gaussienne, avec leurs limitations, est présentée. Quelques compléments sont apportés selon le problème à résoudre.

ABSTRACT

In this text we examine several means for generating non-Gaussian stationary time series with prescribed autocovariance. For each of them we focus on their limitations and propose some new arrangements which depend upon the involved problem.

1 Introduction

Generating pseudo-random numbers which have both a prescribed marginal distribution and a non Dirac autocovariance becomes a need which appears as a milestone in the application of Monte Carlo methods of simulation. Indeed, many stationary time series observed in various contexts (engineering, economics, earth-sciences...) are clearly non-Gaussian and exhibit internal correlation. This correlation may have short range or long range extent. The main part of this work deals with the first, but some topics will be given on the last one.

1.1 Preliminaries : what is to be solved ?

The title of this paper, which is closed to the point of view of an observer, deserves explanations for clarifying the aim of this work.

Some classical definitions and results have to be recalled in order to well state our approach.

A real random sequence $X_n, n = -\infty, \dots, \infty$ is said *strictly stationary* if the distribution of the sequence $X_k, X_{k+1}, \dots, X_{k+m}$ does not depend upon $k \in \mathbb{Z}$ whatever $m \in \mathbb{Z}^+$. The stationarity implies the invariance with respect of time translation, or in other words the independance upon the time origin, a situation which is mostly encountered in practical situations.

A sequence which has finite second-moments is said a *second-order sequence*. In this case, if the mean $m = E(X_k)$ is independant upon k and the covariance matrix is such than $\forall i, j \text{ Cov}(X_j, X_i) = \Sigma_{j,i} = \Sigma_{0,|j-i|}$, the sequence is said *second-order stationary* or *stationary in wide sense* according to the scientific community.

A sequence is *Gaussian* if all of its finite-dimensional marginal distribution are Gaussian. A Gaussian sequence is stationary if and only if it is second-order stationary, then the sequence is completely characterized by m and $\Sigma_{j,i}$.

When the sequence is not Gaussian, it is thus clear that even if it is identically distributed, the knowledge of $\Sigma_{j,i}$ is insufficient to completely characterize the sequence. Therefore there exist many time series, with different finite-dimensional marginal distributions but the first-order, which exhibit the same $\Sigma_{j,i}$. The most evident consequence is (c.f. infra) that they look at different.

Three points may be drawn from these considerations

- i) Additional constraints are necessary in order to build an algorithm and according to the choice of them the result is various [1, 2, 3].
- ii) It must be well understood that the fact to succeed in generating a sequence which has both a prescribed marginal distribution and covariance does not grant the quality of the simulation of a practical situation. It is not the aim of this paper to go further in this direction, each practical case being different. A check of suitability has to be made before using any algorithm. Theses tests are not the main purpose of this paper. However, for example, a test of linearity of the process [4] may be very useful to confirm the so-called linear method (cf infra).
- iii) Nevertheless, the most numerous algorithms are suitable, the most we will have the chance to get a « good » time series, that is to say, which are close to the phenomena we would simulate.

To be complete, underline that the properties of non-negativity and symmetry are necessary and sufficient for a matrix being an autocovariance one[4].

In real situation, one only gets a few of realizations of the defined sequence that prevents estimating its autocovariance. The assumption of second-order ergodicity allows to avoid this drawback and one uses the mean times autocorrelation in place of autocovariance with the probabilistic equivalence

$$\lim_{n \rightarrow \infty} \sum_{-n}^n \frac{1}{2n} E(X_n X_{n+k}) = \Sigma_{0,|k|} + m^2 \quad \text{almost surely}$$

The relevance of this hypothesis is not the matter of this work.

This paper deals with different algorithms convenient for our goal. The presentation is a little bit didactic, even if our purpose is not to achieve a pure mathematical derivation, but only to emphasize some milestones in the various mechanisms we can imagine. In some sens, its are variations on a same subject. Many ideas were already pointed out, their authors are cited in the right place. One can find at the end of the reference book [5] general considerations related to the matter.

1.2 General considerations on the various methods

In the Gaussian case, the problem was deeply investigated for a long time. Almost all of the methods use the fact that the Gaussian character of a sequence is maintained throughout linear system and that moreover, a Gaussian sequence obeys a linear recurrence finite or infinite.

By the way, filtering a Gaussian white noise (a Gaussian sequence iid¹) is the basis of most of the technics.

Before entering in details, let us remind that filtering a discrete process relies upon output y and input x , by a convolution :

$$y = x \star h \tag{1}$$

and that consequently :

$$R_{yy} = R_{xx} \star R_{hh} \tag{2}$$

where the $R_{..}$ are respectively the autocorrelation of the output, the input, and the impulse response. The simple law (2) suggests the three first tracks for solving our problem.

- i) If x is a white noise, then the output has the color of the filter. It just remains to find the right pdf² for x , which

1. iid : independant and identically distributed
2. pdf :probability density function

leads to the wanted pdf of y . This is the essence of the linear method. For a Gaussian noise, as previously mentioned, only the color of the filter has to be determined.

- ii) In a dual way, assuming that the input signal has the wanted color, we can therefore filter by a white random signal. Now, we have to choose a good pdf for h in order to get the wanted pdf for the output. The difficulty is the same as previously. As moreover the algorithms for random filtering are much more complex, this method is not attractive and therefore will not be analyzed more further.
- iii) Any transform $g(x)$ of a time series changes both the autocorrelation and the statistic. First one determines $g(x)$ for getting the desired output pdf from a Gaussian one. By a backward computation, one can deduce the autocorrelation of this Gaussian process which gives, throughout the transform $g(x)$, the wanted R_{yy} . We are thus brought back to a classical problem. We refer this method as the non-linear one. The pioneer works are those of [6, 7].
- iv) When two variables are said to be correlated in common sense, this means that their values are « close » together. In other words, the correlation is concerned more by the relative order than by the absolute values. This observation leads to the method which consists in transferring the order of a time series having the good color on to a well distributed one. We refer this method as « decalcomania ». This idea was first explored by [8] and later by [9].
- v) A complete heuristic approach belonging to stochastic-swap methods has been proposed by [10], is very efficient and robust.

After presenting these methods we study additional tools allowing some combining of autocorrelation functions. Using these tools, one can decompose an autocorrelation function in simpler functions for which the previous methods are easier to implement.

1.3 The red thread

In order to compare the different methods an example will be used throughout this work, it acts as a « red thread ». Originating in the study of rain properties [11], it deals with the successive diameters of rain drops falling on a small defined area.

Figure 1 shows the observed (near Paris) histogram of the diameters which can be considered as issue from a shifted gamma law, the probability density of which being

$$p(x) = \frac{b^{-a}}{\Gamma(a)} (x - S)^{a-1} \exp\left(-\frac{x - S}{b}\right) \quad (3)$$

with $a = 0.69; b = 0.41; S = 0.2$ for x in mm

Figure 2 shows the autocorrelation which is quite a weighted sum of pulse and triangle. In order to achieve our goal we use the technic of combination already mentioned. Therefore we firstly deal with a triangular shape for the autocorrelation.

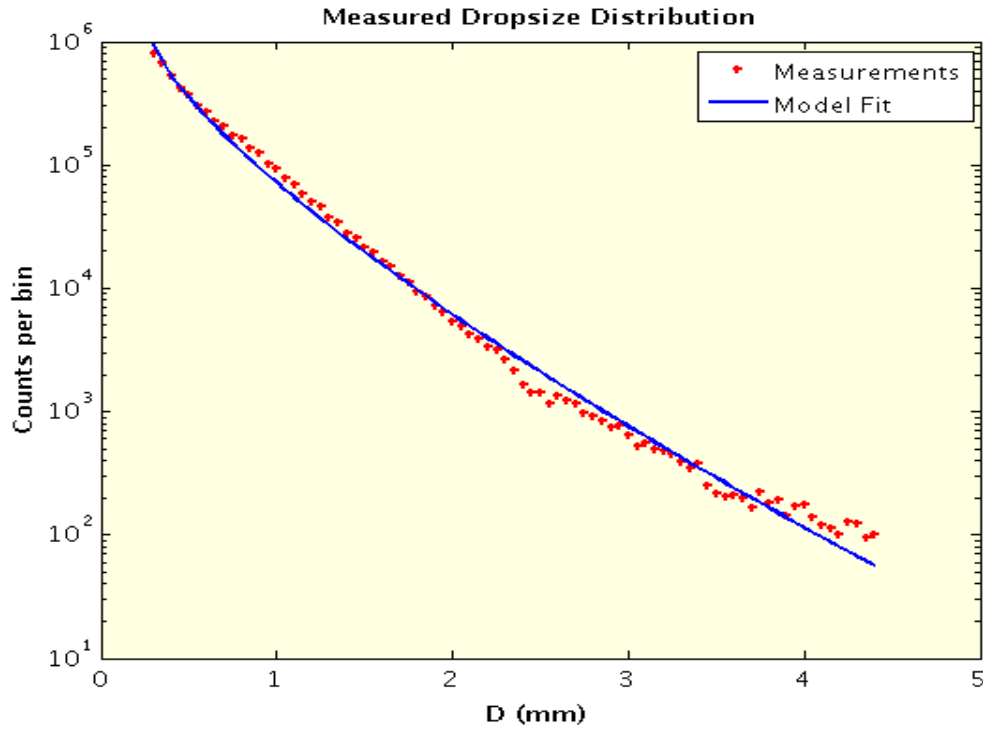


FIGURE 1 – Histogram of raindrop size observed near Paris(France) in 1993

2 The non-linear method

2.1 Principles

Let us take X a sequence of iid random variables, the pdf of which being $f_1(x)$. Let $g(x)$ be a regular transformation ($g'(x) \neq 0$). The new sequence $Y = g(X)$ has the following pdf :

$$p(y) = \frac{f_1(x)}{g'(x)} \quad (4)$$

The inversion of (4) is easy to obtained, analytically or numerically.

The autocorrelation of Y reads :

$$\begin{aligned} R_{yy}(k) &= E[y(i)y(i+k)] \\ &= \iint g(x(i))g(x(i+k))f_2[x(i), x(i+k)]dx(i)dx(i+k) \quad (5) \end{aligned}$$

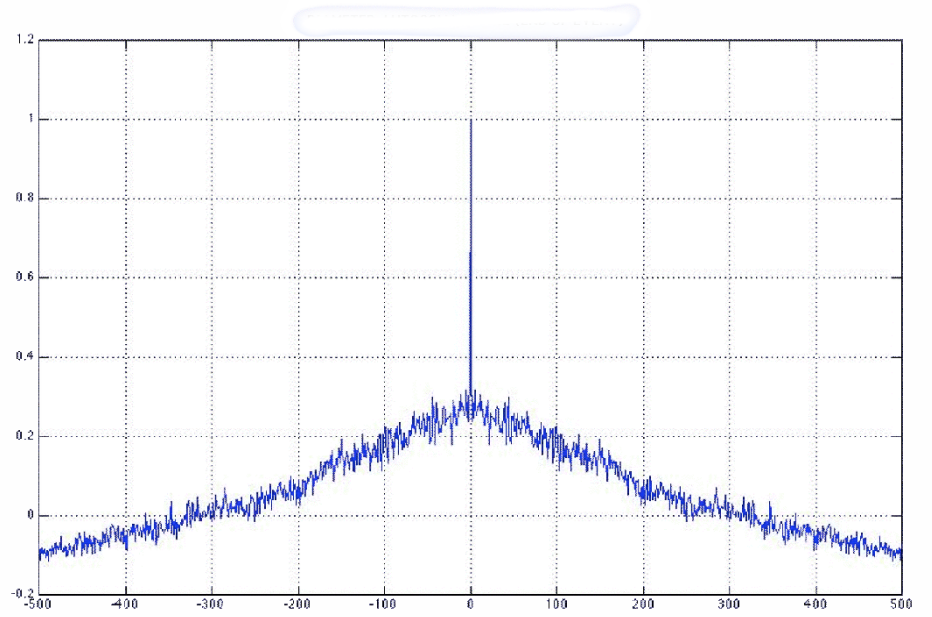


FIGURE 2 – Example of autocorrelation of the same sequence as figure 1

In (5), f_2 is the second order probability density function. If f_2 depends only upon f_1 and the autocovariance R_{xx} of X , therefore (5) can be inverted and once upon a time g has been determined, one gets $R_{xx}(k)$ function of R_{yy} which is the wanted color.

In some way, we have transpose our problem to an other probability distribution modifying accordingly, the color of the sequence. It is therefore advantageous to choose a distribution for which the problem is solvable. Gaussian distribution is clearly a must. Figure 3 sketches the method which is achieved when g and h are known.

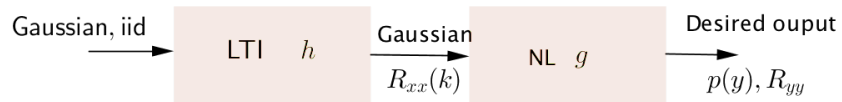


FIGURE 3 – Sketch of the non-linear method

2.2 Determination of $g(x)$

Without any restriction, one can choose :

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$f_2(x', x'', k) = \frac{1}{2\pi\sqrt{1 - \rho_{xx}^2(k)}} \exp \left[-\frac{x'^2 - 2\rho_{xx}(k)x'x'' + x''^2}{2(1 - \rho_{xx}^2(k))} \right] \quad (6)$$

where $\rho_{xx}(k) = R_{xx}(k)/\sigma^2$ denotes the normalized autocorrelation function. If F is the cumulative distribution function of the wanted time sequence, inversion of (4) gives

$$g(x) = F^{-1}\left(\frac{1 + \operatorname{erf}(x/\sqrt{2})}{2}\right) \quad (7)$$

The autocorrelation R_{xx} is now obtained through inversion of

$$R_{yy}(k) = \iint \frac{g(x')g(x'')}{2\pi\sqrt{1 - \rho_{xx}^2(k)}} \exp \left[-\frac{x'^2 - 2\rho_{xx}(k)x'x'' + x''^2}{2(1 - \rho_{xx}^2(k))} \right] dx' dx'' \quad (8)$$

One checks that $-1 \leq \rho_{xx} \leq 1 \Rightarrow m_1^2 \leq R_{yy} \leq m_2$, as expected, where m_n is the n^e order moment of Y . In some cases, the inversion of (8) can be performed analytically. For example, if Y has to be uniformly distributed over $(0, 1)$ one computes [7] :

$$\rho_{xx}(k) = 2 \sin \frac{\pi \rho_{yy}(k)}{6} \quad (9)$$

In most of the cases, as for our red thread, the inversion must be numerically executed.

2.3 Determination of h

The main drawback of the method is that it fails in some cases when $R_{xx}(k)$ is not well behaved. Indeed, $R_{xx}(k)$ has to be definite non-negative [4] and this property is not guaranteed. A simple counter-example is developed in Appendice A.

It is very difficult to find the conditions prevailing this phenomenon. In fact, we have observed its removal when $R_{yy}(k)$ is sufficiently concave.

Fortunately in a lot of questions, the accuracy of the correlation is not mandatory. In such a case, one can take this opportunity for escaping the difficulty by modifying slightly $R_{yy}(k)$.

When the objective $R_{xx}(k)$ is well behaved, it remains to determine the impulse response h of the filter. At this point, it is worthwhile to recall that classical methods such AR, MA or ARMA do not operates in all cases. As previously, a trade off may be to loose accuracy in order to get tractability. Nevertheless, the most appropriate technics seems to be MA because AR leads often to an unstable filter and ARMA requires AR as a first step.

2.4 Example

MA technics consists in finding the zeros of the following polynomial $\Phi(z)$ in order to get a factorized form :

$$\Phi(z) = z^N [R_{xx}(N)z^N + \dots R_{xx}(0) \dots + R_{xx}(N)z^{-N}] = B(z)B(z-1)$$

The zeros inside the unit circle generates $B(z)$ and the zeros outside $B(z-1)$. When zeros are on the unit circle, the factorization is impossible with real values. This constraint is different from the non-negativity one.

In our red thread example, the correlation may extend to 1000 lags. The polynomial $\Phi(z)$ is thus of order 2000 and they are numerical difficulties in finding its zeros. To remove this trouble, we have restrict this polynomial at a less degree (e.g. 20). Once the corresponding $B_r(z)$ is known, it is expanded by interpolation over the desired maximum lag (here 1000). Let $b(i), i = 1, \dots, N$, be the coefficients of $B(z)$. The normal time sequence with normalized autocovariance $R_{xx}(k)$ is obtained by

$$R_{xx}(k) = \sum_{i=0}^N b(i)\nu(k-i)$$

where ν is a white normal noise. As a first step, what is intended is to simulate a pure triangular correlation (without the impulse at the origin). As already quoted, the shifted gamma law (3) does not lead to a well-behaved R_{xx} . Therefore, a new form has been chosen, which is close to a triangle but suitable for the method :

$$R_{yy}(k) = \frac{1 - 0.001|k|}{1 + |k|/750}$$

Figure 4 shows the results of the simulation over 500 000 points. They are very good except the fact we had slightly modify the autocorrelation according to the formula above.

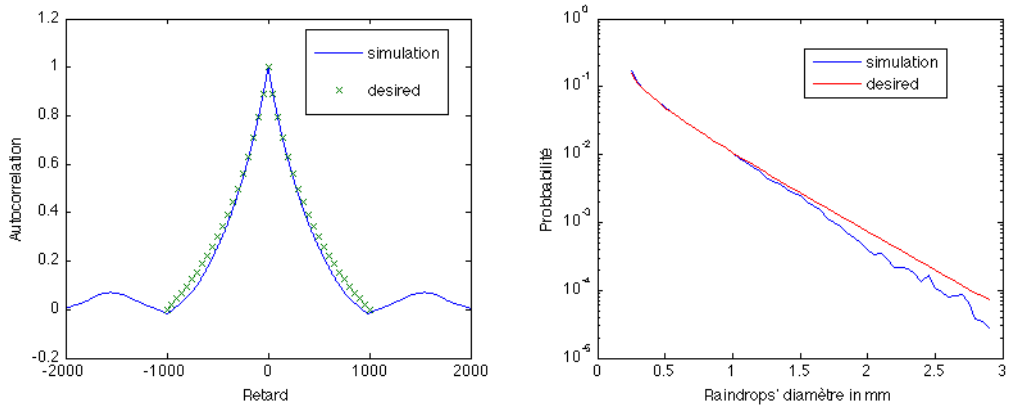


FIGURE 4 – Example of the non-linear method

3 The linear method

3.1 Principles

In essence, the linear method filters a white time sequence, the first order statistical distribution of which is matched with the desired distribution of the output. For the same reasons quoted above, the filtering resorts to MA technics. Thus the output reads :

$$y(k) = \sum_{i=0}^N b(i)x(k-i) \quad (10)$$

The coefficients $b(i)$ are computed as mentioned in the previous section and with the same limitations. X is an iid. sequence whose probability density function has to be specified. From (10) it follows that the pdf of Y is a multiple convolution :

$$g(y) = \prod_{i=0}^N \star \left[\frac{1}{b(i)} f\left(\frac{y}{b(i)}\right) \right] \quad (11)$$

For solving (11), we use the generating functions of the first and second order³. From (11), we deduce (with self-explained notations) :

$$\hat{g}(u) = \prod_{i=0}^N \hat{f}(ub(i)) \quad (12)$$

$$\tilde{g}(u) = \sum_{i=0}^N \tilde{f}(ub(i)) \quad (13)$$

When (12) or (13) are inverted, the problem is solved. Moreover, even if one does not know invert generating functions their knowledge is sometimes sufficient for drawing [12].

3.2 Example : the red thread

In this case, the correlation being triangular the N coefficients $b(i)$ are equal. We can specify $b(i) = 1/\sqrt{N}$. As $g(y)$ is infinitely divisible, $\forall n \in \mathbb{Z}^+$, $\hat{g}(u)^{1/n}$ exists and is a generating function. Therefore, the solution of (12) is simply $\hat{f}(u) = \hat{g}(u\sqrt{N})^{1/N}$ which corresponds to the pdf

$$f(x) = \frac{1}{\Gamma(a/N)(b\sqrt{N})^{a/N}} (x - S/\sqrt{N})^{\frac{a}{N}-1} \exp\left(-\frac{x - S/\sqrt{N}}{b\sqrt{N}}\right) \quad (14)$$

Figure 5 shows the results of the simulation accordingly to this method. Very efficient for the autocorrelation, this method is poorer than the non linear one as pdf is concerned. It is noticeable

3. We use the definitions $\hat{f}(u) = E[e^{ux}]$ and $\tilde{f}(u) = \log[\hat{f}(u)]$

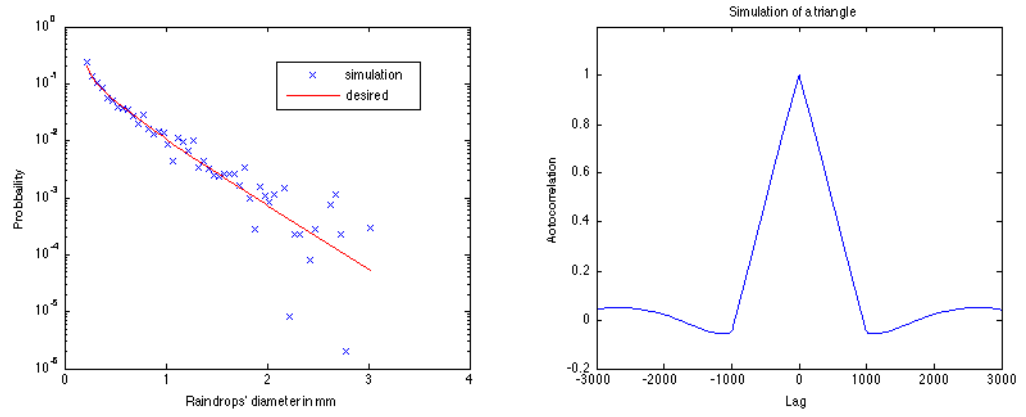


FIGURE 5 – Example of the linear method

and must be underlined that the time series look very different in both method, linear and non-linear.

Figure 6 gives an extract of the simulated sequence obtained in the two methods. This difference is clearly a confirmation of the discussion of the introduction. One does not would be satisfied with getting only the good distribution and correlation. It is mandatory to check more (e.g. third moment) depending upon the simulation uses.

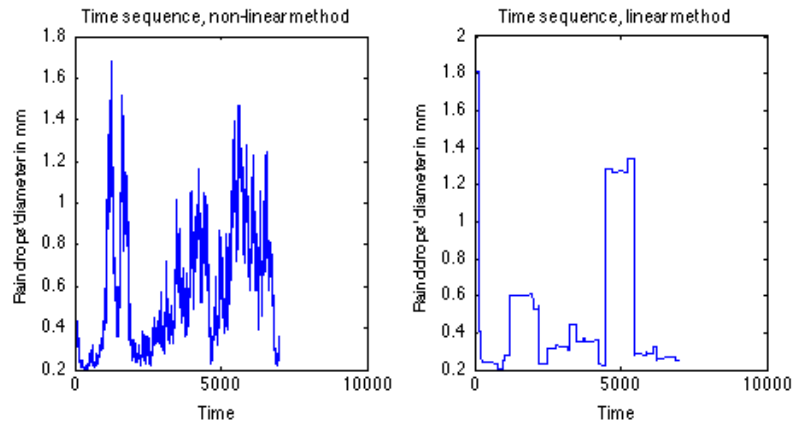


FIGURE 6 – Extract of simulated sequence for both methods

3.3 Further considerations

Our example is so particular than one can wonder if it is the only one for which the linear method is completely achievable! The conjunction of a particular shape of autocorrelation and a infinitely divisible distribution is however not unique. The following deals with two approaches allowing to extend the application's field of the linear method.

3.3.1 The Mellin transform approach

We will restrict ourself to the distributions defined on \mathfrak{R}^+ . Indeed, the Mellin transform of a real function $h(t)$ reads :

$$H(s) = \int_0^\infty h(t)t^{s-1}dt \quad \text{for } \alpha < \text{Real}(s) < \beta$$

The inverse transform is

$$\forall \alpha < c < \beta \quad h(t) = \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} H(s)t^{-s}ds$$

Apart of the linearity, the Mellin transform has the useful property :

$$\forall a > 0 \Rightarrow h(at) \xrightarrow{TM} a^{-s}H(s) \quad (15)$$

If we consider a pdf $f(x)$ defined over \mathfrak{R}^+ , its generating function $\hat{f}(u) = E[e^{ux}]$ is real. Its second order generating function is thus real and non positive. Therefore one gets from (13) :

$$\tilde{g}(u) = \sum_{i=0}^N \tilde{f}(ub(i)) \xrightarrow{TM} G(s) = F(s) \sum_{i=0}^N b(i)^{-s} \quad (16)$$

Formally, the solution of (19) reads

$$F(s) = \frac{G(s)}{\sum_{i=0}^N b(i)^{-s}} \quad (17)$$

In order to achieve the resolution, it just remains to take the inverse Mellin transform of (17), that is often not very easy.

It is worthwhile to recognize $\sum_{i=0}^N b(i)^{-s}$ as a Dirichlet series, some of them are summable in a closed expression. In some cases the inversion of (17) may thus be achieved easier.

For instance, when $b(i) = b = \text{cte}$, (17) becomes $F(s) = G(s)/Nb^{-s}$, thus $\tilde{f}(u) = \tilde{g}(u/b)/N$ and $\hat{f}(u) = \hat{g}(u/b)^{1/N}$. We recover the previous solution for an infinitely divisible distribution.

Therefore, one can build family of autocorrelation functions for which the linear method is available. Among them, one can cite an important class.

AUTOCORRELATION OF EXPONENTIAL TYPE

With $b(i) = \alpha^i$, $0 < \alpha < 1$ and $N = \infty$, one gets successively⁴

$$R_{yy}(k) = \alpha^k \quad \text{and} \quad \sum_{i=0}^{\infty} b(i)^{-s} = \frac{1}{1 - \alpha^{-s}}$$

from which one has

$$F(s) = (1 - \alpha^{-s})G(s) \Rightarrow \tilde{f}(u) = \tilde{g}(u) - \tilde{g}(\alpha u) \Rightarrow \hat{f}(u) = \frac{\hat{g}(u)}{\hat{g}(\alpha u)}$$

4. Notice that this filtering is equivalent to the ARMA recurrence $-\alpha y(n) + y(n-1) = x(n-1)$

Specifying the pdf of Y as an exponential one as $g(y) = \lambda e^{-\lambda y}$ and $\hat{g}(u) = \frac{\lambda}{\lambda - u}$. Therefore it comes

$$\hat{f}(u) = \frac{\lambda - \alpha u}{\lambda - u} \Rightarrow f(x) = (1 - \alpha)\lambda e^{-\lambda x} + \alpha\delta(x)$$

3.3.2 The cumulant approach

When the inversion of (17) is not possible, one can use another method for simulation purposes.

The cumulants of a distribution having a second order moment generating function $\tilde{g}(u)$ are given by the derivations

$$\kappa_j = \left. \frac{d^j \tilde{g}}{du^j} \right|_{u=0} \quad (18)$$

Therefore (13) leads to the relationship

$$\kappa_k = \left[\sum_{j=0}^{\infty} b_j^k \right] \chi_k \quad (19)$$

where κ_k and χ_k are respectively the cumulant of order k of the distributions g and f .

The moment problem [13] deals with the existence and unicity of a distribution having a given sequence of moments (or cumulants). It is beyond the scope of this paper to give an extensive discussion about the existence of a distribution having a sequence of cumulants given by (19).

It has been established some conditions for the existence but there exists some examples for which the unicity is not fulfilled. However, in view of simulation, the unicity is not mandatory due to the fact that the target is not the distribution f but the distribution g .

Nevertheless caution is required and each case must be carefully controlled.

Since the cumulants are known, one is able to generate a lot of distribution functions either using the Edgeworth expansion [14, 15] or orthogonal development [12].

Appendice D gives some details about this technic and a corresponding algorithm. Nevertheless, as explain in the appendice there is no certainty about the convergence of the used expansion.

A COUNTER-EXAMPLE

The sequence $b(i) = (i + 1)^{-m}$, $i = 0, \dots, \infty$, where m is a positive real, generates a family of normalized autocovariance functions which read after straightforward calculations :

$$R_{yy}(k) = \sum_{i=1}^{\infty} \frac{[i(i + k)]^{-m}}{\zeta(2m)}$$

where ζ is the Riemann zeta function.

As $\sum_{i=0}^{\infty} b(i)^k = \sum_{j=1}^{\infty} j^{-mk} = \zeta(mk)$, (19) is easily reducible to :

$$\chi_k = \frac{\kappa_k}{\zeta(mk)}$$

For instance, it is worthwhile to note that, for $0.5 < m < 1$, this family of autocorrelation functions includes long range of correlation ($\sum_k R_{yy}(k) = \infty$), but this method fails trivially when the target distribution has not a null mean. Indeed, $\sum_{j=1}^{\infty} j^{-m}$ diverges when $m < 1$ and thus (19) implies $\kappa_1 = \chi_1 = 0$.

4 Probability mixing and the decalcomania method

4.1 Preliminaries

Let us take two sequences and perform a probability mixing of them. If the first order statistical properties are identical for the two sequences one keeps it after mixing. However it is not the case for the autocorrelation. We take advantage of this behavior for solving our problem.

For fixing the ideas let us take A and B two stationary sequences independent and identically distributed with normalized autocorrelations $\alpha(k)$ and $\beta(k)$. Let U be a new time sequence of zeros and ones, following an « head and tail » law of parameter p with normalized autocorrelation $C(k)$. Let us define the time sequence Z by the relationship :

$$Z(n) = U(n)A(n) + (1 - U(n))B(n) \quad (20)$$

The normalized autocorrelation of Z reads with $q = 1 - p$:

$$\rho_Z(k) = q^2\alpha(k) + p^2\beta(k) + pqC(k)[\alpha(k) + \beta(k)] \quad (21)$$

From (21), it appears that the autocorrelation includes now a Dirac at the origin. That means the probability mixing whiten the sequences by increasing the disorder. At first glance probability mixing is thus not a good tool for building a colored noise. However, generalization of this mixing open new possibilities of transformation of the autocorrelation without changing the first order statistic.

Let us take n times sequences identically distributed and mutually independent. Aggregate them with the vectorial symbol \mathbf{X} .

Let \mathbf{A} be a sequence of random vectors of dimension n whose only one coordinate is equal to 1, the others being 0. The following sequence is created :

$$Z(k) = \mathbf{X} \cdot \mathbf{A} = \sum_{i=1}^n x_i(k)a_i(k) \quad (22)$$

There is a direct equivalence between the multi-dimensionnal sequence \mathbf{A} and a sequence U , the domain of definition of which being the rationals from $1/n$ to 1 by step of $1/n$. Using the different normalization relationships and introducing the normalized autocorrelation of U , $R_{uu}(m)$, straightforward calculations (Appendice B) gives

$$R_{zz}(m) = \left[\frac{n-1}{n} R_{uu}(m) + \frac{1}{n} \right] R_{xx}(m) \quad (23)$$

If $\forall k R_{xx}(k) = 1$, that means the sequence X are constant, then in the limit $n \rightarrow \infty$, one gets :

$$R_{zz}(m) = R_{uu}(m) \quad (24)$$

By construction the series Z has the same distribution than X . But, here, (24) carries out the transfer of the color from U to Z . This result gives mathematical foundations to the ideas introduced by [8].

4.2 Implementation

Outline of the method, which is well named decalcomania, consists in random drawing a sequence of uniform variable U over $(0,1)$ with prescribed color, then taking $F^{-1}(u)$ where F is the wanted distribution function.

Evidently, the construction of U requires another method. The non linear one is particularly attractive taking into account the simplicity of the function $2 \sin[\pi R_{zz}/6]$ - see (9) -.

Unfortunately we are not guaranteed that this function has the well autocorrelation behavior as noted above in section 2.

Nevertheless, whatever our tests are, the results are particularly

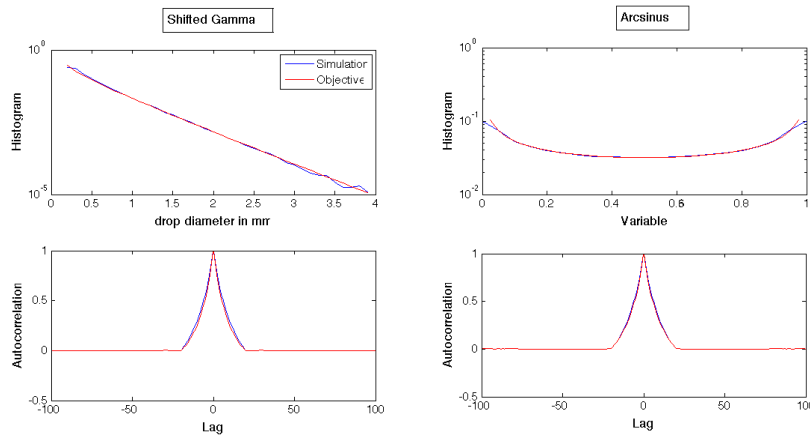


FIGURE 7 – Examples of decalcomania method

good, as figure 7 attests. Moreover it is suitable to emphasize the

simplicity of this method when it works.

Besides, it appears that this technic may be useful when simulating long range memory processes (see below).

5 The swap method

This method is a purely heuristic one. It is often use when the sequence is not too long [10]. Since, generally, one intends to simulate a real situation the idea is to start with a sequence identically distributed which follows the prescribed probability distribution. One postulates that, by rearranging the sequence (among the $N!$ possibilities if N is the length of the sequence) one can find a permutation, the autocorrelation of which is close to the desired one. The question which arises immediately deals with the duration of such a process, because a « bruteforce » is in $N!$ and there is no way *a priori* to find a short path.

Oneway to proceed is to use the following algorithm which starts from the iid sequence X_0 of length N . Let us d_i the distance between the autocorrelation R_{X_i} and the desired autocorrelation R ; we will come back on the choice of this distance. ϵ is the upper distance we would reach.

X_i

Swap two values x_i, x_j .

The couple (i, j) is uniformly distributed over $(0, N) \times (0, N)$.

This transformation gives a new sequence X_{i+1} .

If $d_{i+1} \geq d_i$ return to the beginning.

If $d_{i+1} < d_i$ substitute X_{i+1} for X_i and start again but $d_{i+1} < \epsilon$ where one stops.

In order to decrease the computing time, it is advised to compute

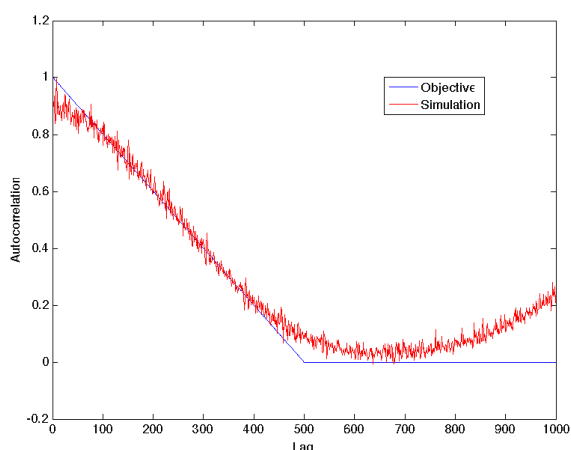


FIGURE 8 – Autocorrelation obtained after 10^7 iterations for the red thread

not d_{i+1} but $d_{i+1} - d_i$. Details are given in Appendice C. The choice

of the distance depends upon what one thinks important in the realization of the autocorrelation.

In the work presented below the distance has been chosen as

$$\sqrt{\sum_m (R_{X_i}(m) - R(m))^2 / N}.$$

The figure 8 shows, for our red thread, the result obtained after 10^7 iterations which last 318 sec whereas one got better approximations with the previous methods (linear and non-linear) in a few seconds with the same computer (see above).

The figure 9 shows how the the computing time is in a power dependence upon the accuracy.

Despite the lack of a thorough theoretical explanation, it is pretty sure that the length of the simulated sequence is very penalizing. In our red thread, this length is of 500 000 and the swap of only two elements modifies slightly the autocorrelation.

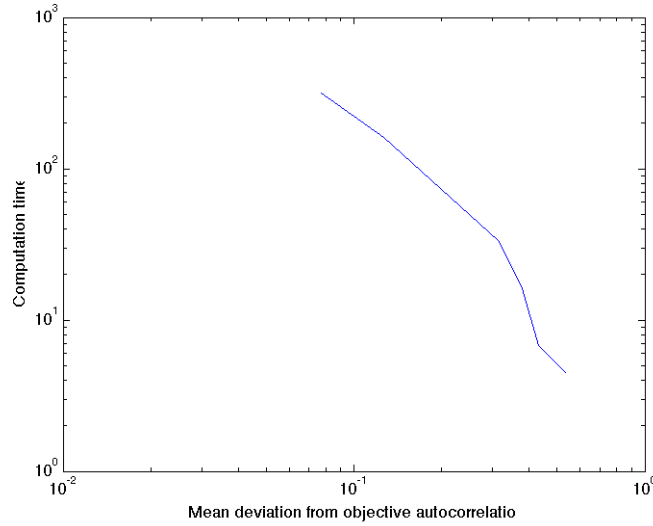


FIGURE 9 – Computing time versus accuracy of the swap method

An improvement of this algorithm may be found in generalizing the method by performing a random permutation over a random sampling of S points. One can use the « Latin Hypercube Sampling » for the choice of these points.

Nevertheless, when one deals with moderate length, the method is very efficient. The following example where the length is 10000 gives a good result (accuracy $\approx 5 \cdot 10^{-2}$) in 2 sec. The figure 10 shows the simulated autocorrelation.

6 Splitting of autocorrelation

In our full red thread the autocorrelation is compound and may be split in to more simple functions. It is quite general and

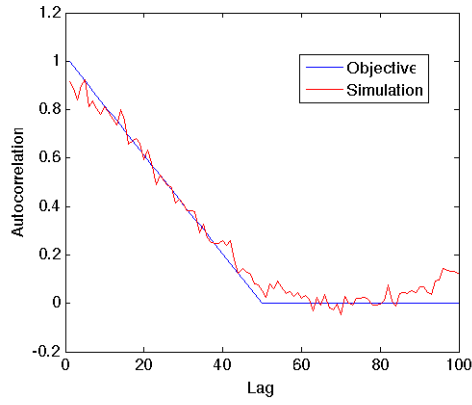


FIGURE 10 – Simulation of the autocorrelation of a sequence of 10000 points following a shifted gamma distribution. Swap method.

consequently useful to examine how one is able to combine autocorrelations. It is not a new method in itself but an easier way to solve practical problems.

Let us take X_1 and X_2 two independent time sequences with variances and normalized autocorrelations, σ_1, σ_2 and $R_{X_1X_1}, R_{X_2X_2}$ respectively. Making the linear combination $Y = X_1 + X_2$ we obtain a new time sequence. Its normalized autocorrelation is easily computed as :

$$R_{YY} = \frac{\sigma_1^2 R_{X_1X_1} + \sigma_2^2 R_{X_2X_2}}{\sigma_1^2 + \sigma_2^2} \quad (25)$$

By this process, it is thus possible to combine autocorrelation so long as their concavity are up.

The generating function of Y is

$$\hat{g}(u) = \hat{g}_1(u)\hat{g}_2(u) \quad (26)$$

If the target Y has an infinitely divisible pdf, equation (26) would be simply used.

APPLICATION

In our red thread

$$\hat{g}(u) = e^{uS}(1 + bu)^{-a}$$

which is easily decomposable into two functions of the same type. Elementary calculation shows that one of the possible choices is :

$$\left. \begin{matrix} (1 + bu)^{-ap} \\ e^{uS}(1 + bu)^{-a(1-p)} \end{matrix} \right\} \Rightarrow R_{YY} = pR_{X_1X_1} + (1 - p)R_{X_2X_2} \quad (27)$$

Figures 11 show the results of simulation using this technics for getting an autocorrelation made of two triangles.

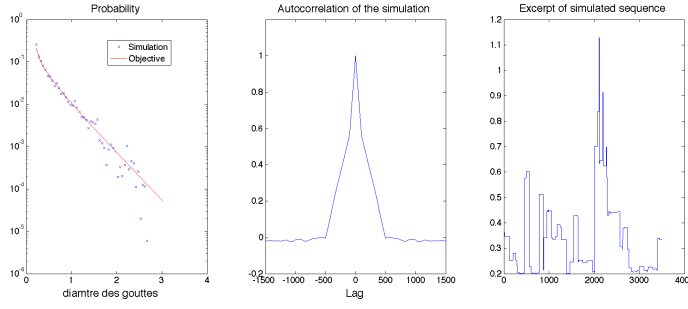


FIGURE 11 – Simulation of the autocorrelation formed by two triangles. The simulation is obtained by the non linear method

7 Self-similar and long range correlation processes

7.1 Background

Processes exhibiting long range correlation (the absolute value of its autocorrelation is not summable) and/or self-similar behavior seem play an more and more important role in many technical and scientific area [16, 17, 18]. It is thus worthwhile to investigate if our approach could be useful in their domain.

Let us take a strictly self-similar continuous process of Hurst parameter $H > 0$ (e.g [19]). By definition, its marginal distribution obeys at

$$Y_t \stackrel{d}{=} t^H Y_1 \quad \text{for } t > 0 \quad (28)$$

where $\stackrel{d}{=}$ means equality in the sense of probability distributions. Among them, processes with stationary increments are of particular interest because they are good candidate for representing a lot of practical situations. In such cases, one can write

$$E[(Y_t - Y_s)^2] \begin{cases} = E[(Y_{t-s} - Y_0)^2] \\ = E[Y_t^2 + Y_s^2 - 2Y_t Y_s] \end{cases} \quad (29)$$

If one notes $m = E[Y_1 - Y_0]$ and $\sigma^2 = E[(Y_1 - Y_0 - m)^2]$, formulas (29) allows to get

$$\text{Cov}(Y_t, Y_s) = E[Y_t Y_s] - E[Y_t]E[Y_s] = \frac{\sigma^2}{2} [t^{2H} - (t-s)^{2H} + s^{2H}] \quad (30)$$

Then, a straightforward computation stands that the autocorrelation of the sequence of the increments $X_i = Y_i - Y_{i-1}$ is given by

$$R_{XX}(k) = \frac{1}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}] \quad (31)$$

The asymptotic behavior of R_{XX} is

$$R_{XX}(k) \approx_{k \rightarrow \infty} H(2H-1)k^{2H-2} \quad (32)$$

When $\frac{1}{2} < H < 1$, $R_{XX}(k)$ is slowly decreasing with k , and we can talk about a long range memory ($\sum_k R_{XX}(k) = \infty$) This slowness is the more large since H is close to 1.

Clearly, the discussion of the introduction still stands and the previous results do not mean that the simulation of a self-similar process with stationary increments is equivalent to the simulation of a sequence having prescribed distribution and correlation.

Particularly, one has to take caution because the marginal distribution of Y_1 must have some peculiar properties : absence of atoms, support connected... [20]. An important class [21] is constituted by the strictly stable process of index α^5 , and its Hurst parameter $H = \alpha^{-1}$. However, as for other types of process, it may be suitable to perform a simulation although not exactly self-similar is nevertheless « close » to the experimental data.

7.2 Simulation

7.2.1 Generalities

Unfortunately, even if the filter leading to the Riemann family (cf. page 12) holds an autocorrelation presenting a great similarity (see appendice E) with (31), this way is a dead end as quoted above in the same page (non-convergence of the expansion of the distribution function).

When considering that many algorithms have been proposed for simulating self-similar Gaussian noise, the idea is therefore to use the non-linear method (section 2) for transforming in a self-similar uniform process and then to apply the decalcomania method (section 4).

This approach is very attractive because formula (9) shows that $\rho_{yy}(k) \approx \pi \rho_{xx}(k)/3$ when $\rho_{xx}(k)$ is close to 0, if ρ_{yy} is like (31). It is therefore natural to search if ρ_{xx} is close to a similar function but with a different parameter H .

Figure 12 exhibits, first the autocorrelation after non linear transformation $g(x)$ when the initial one is of type (31) with parameter $H = 0.8$, second the autocorrelation of type (31) with parameter $H = 0.805$. The difference between both curves measured as the maximum of their absolute difference up to a lag of $2 \cdot 10^6$ is only $6.4 \cdot 10^{-3}$. This result remains true for other values of H .

Consequently, the proposed method is composed of three steps.

1. Determine the value of H' corresponding to H
2. Draw a self-similar Gaussian process of Hurst parameter H'
3. Transform the values of the sequence in order that both of corresponding distribution functions - Gaussian and objective - are equal.

5. A process the marginal distribution Y_1 of which is stable is said to be stable with the same index. Stable distributions are subset of infinitely divisible ones.

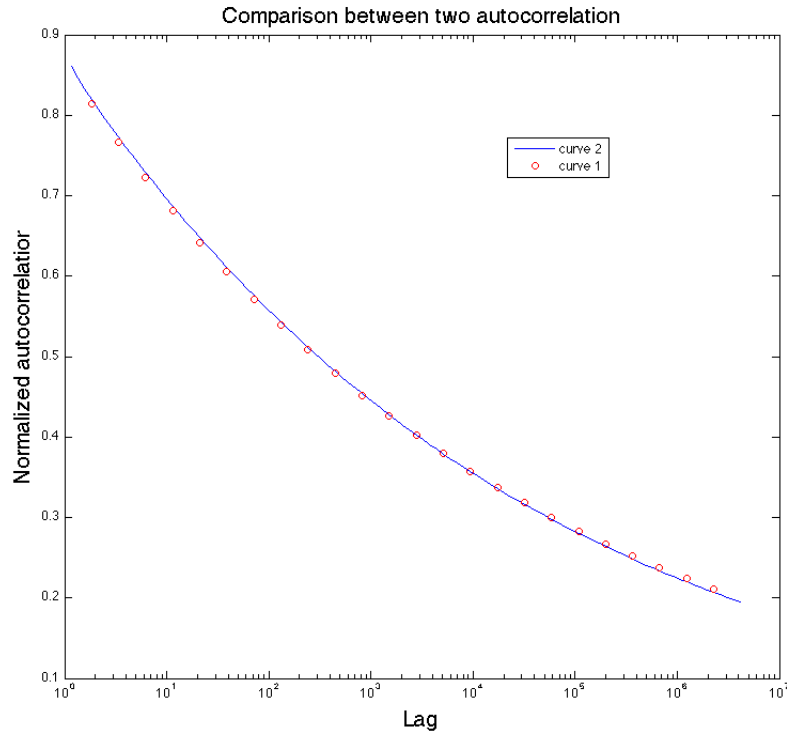


FIGURE 12 – Curve 1 : $2\sin(\pi\rho/6)$ with ρ given by (31) where $H = 0.8$ Curve 2 : ρ given by (31) where $H = 0.805$

7.2.2 Example

One takes as an artificial example the challenge to build a long memory process the marginal distribution of which being an exponential one with parameter 1.

As the « persistence » becomes very important when H is close to 1, a very long sequence is mandatory in order to perform a meaningful statistical work. This requirement leads to some computational difficulties. Therefore we have chosen the value $H = 0.8$ which can be easily managed with a sequence of 2^{17} length.

The three steps of the technic are

1. As previously determined $H' = 0.805$
2. The drawing of a standard fractional Gaussian noise uses the method of circulant embedding matrix [22]
3. The transformation of the normal values of the sequence into an exponential ones via the cumulative distribution function.

Figure 13 shows the results of a simulation following the quoted process. The extract of the sequence is just aimed for illustration. One sees clearly that the initial requirements are fulfilled : the marginal distribution and the autocorrelation are those which has been searched. The method is thus satisfactory.

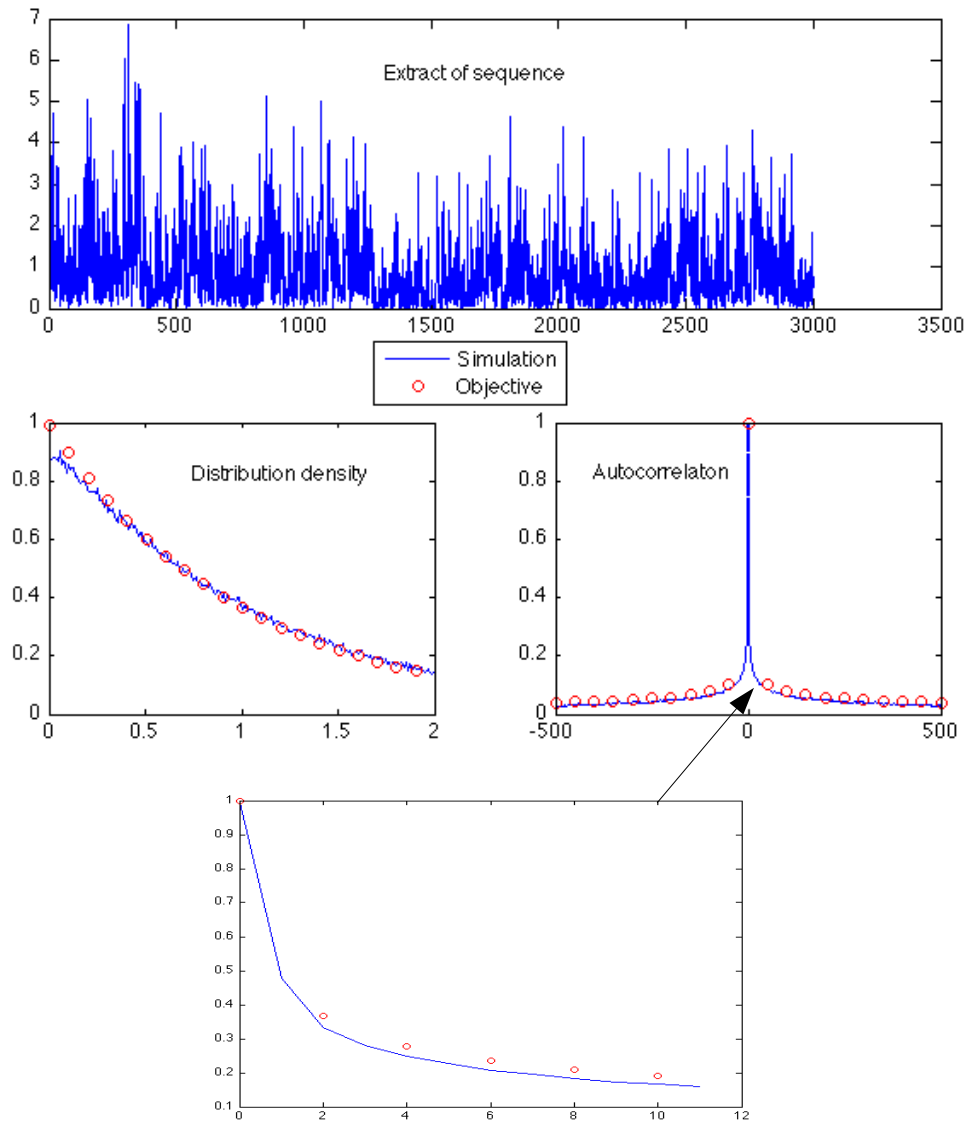


FIGURE 13 – The three parts are respectively an extract of the simulated sequence, the marginal density distribution compared with the target one and the autocorrelation of the sequence together with the desired one. The last figure is just a zoom near the origin.

8 Conclusion

Clearly this paper deals with the possibility of simulate experimental processes and not modeling them. Not any of the four methods which have been studied is suitable in any situation. The table 1 sums up the main possibilities inherent to each of them. Nevertheless, each method has its own limitations described in the corresponding section. For instance, the decalcomania technic assumes a simulated uniform distribution : it can be only approximated in the case of long memory. Decalcomania method is thus also heuristic sometimes...and the table 1 just proposes coarse features.

	Non-linear	Linear	Decalcomania	Swap
Exact	✓	✓	✓	
Heuristic				✓
Short sequence	✓	✓	✓	✓
Long sequence	✓		✓	
Short memory	✓	✓	✓	✓
Long memory			✓	

TABLE 1 – Behavior of the methods along different point of view

The red marks emphasize the most efficient way to proceed. It has to be recall (1.1) that, in any case, marginal distribution and autocorrelation are insufficient to characterize a random process. Therefore even if these both requirements are fulfilled it must be care that the resulting sequence look like the experimental one.

Appendices

Appendice A: Counter-example for the non-linear method

Restrict ourselves to an autocorrelation function R , the z transform of which being

$$\Phi(z) = R(2)z^2 + R(1)z + R(0) + R(1)z^{-1} + R(2)z^{-2}$$

R , real, is non-negative if and only if

$$\forall u_i, u_j \in \mathfrak{R} \quad \sum_{i,j} u_i u_j R(|j-i|) \geq 0$$

This condition is equivalent, in the present case, to

$$\begin{vmatrix} R(0) & R(1) & R(2) \\ R(1) & R(0) & R(1) \\ R(2) & R(1) & R(0) \end{vmatrix} \geq 0$$

which reads

$$[R(0) - R(2)][R(0)(R(0) + R(2)) - 2R(1)^2] \geq 0$$

With $R(0) = 1 \geq R(2)$ the condition becomes simply

$$1 + R(2) - 2R(1)^2 \geq 0 \quad (33)$$

Therefore, using formula (9), one gets R_{yy} satisfies the condi-

$R_{yy}(0) = 1$	$R_{xx}(0) = 1$
$R_{yy}(1) = 0.781$	$R_{xx}(1) = 0.82$
$R_{yy}(2) = 0.247$	$R_{xx}(2) = 0.3$
$R_{yy}(k > 2) = 0$	$R_{xx}(k > 2) = 0$

TABLE 2 – Example of relation between R_{xx} and R_{yy}

tion (33) and therefore is definite non-negative, while $1 + R_{xx}(2) - 2R_{xx}^2(1) = -0.045 < 0$, that implies R_{xx} is not an autocorrelation function. Consequently the non-linear method does not work in this example.

Appendice B: Autocorrelation of mixing sequences

Taking the notations of section 4, we have due to the independence of different sequences

$$E(Z) = E(\mathbf{X}) \cdot E(\mathbf{A}) = m \sum_{i=1}^n p_i = m \quad (34)$$

$$E(Z^2) = E(\mathbf{X} \cdot \mathbf{X}^t) E(\mathbf{A} \cdot \mathbf{A}^t) \quad (35)$$

where $m = E(X_i)$, $\forall i$ and $p_i = \text{Prob}[a_i = 1]$.

By construction, the matrix $\mathbf{A} \cdot \mathbf{A}^t$ is the identity matrix and therefore $E(Z^2) = E(X_i^2) \sum_{i=1}^n p_i = E(X_i^2)$ whatever i .

Similarly

$$E[Z(k)Z(k+m)] = E[\mathbf{X}(k) \cdot \mathbf{X}^t(k+m)]E[\mathbf{A}(k) \cdot \mathbf{A}^t(k+m)]$$

The matrix $\mathbf{A}(k) \cdot \mathbf{A}^t(k+m)$ has only one element different of 0 and equal to 1. Its probability will be noted

$$\pi_{ij}^m(k) = \text{Prob}[a_i(k) = 1 \text{ and } a_j(k+m) = 1] \quad (36)$$

According to our hypothesis, this probability does not depend upon k and we have

$$\sum_{i=1}^n \pi_{ij}^m = p_j \quad \sum_{j=1}^n \pi_{ij}^m = p_i \quad \sum_{i,j=1}^n \pi_{ij}^m = 1 \quad (37)$$

As

$$\begin{aligned} i \neq j &\Rightarrow E(X_i X_j) = m^2 \\ i = j &\Rightarrow E(X_i(k) X_i(k+m)) = \text{Corr}(X_i) + m^2 \end{aligned}$$

it results

$$E[Z(k)Z(k+m)] = \sum_{i,j=1}^n \pi_{ij}^m m^2 + \sum_{i=1}^n \pi_{ii}^m \text{Corr}(X)(m) \quad (38)$$

Therefore the autocorrelation of Z reads

$$R_{zz} = \sum_{i=1}^n \pi_{ii}^m R_{xx} \quad (39)$$

Assuming that $U(k)$ is an exchangeable sequence, it comes

$$p_i = 1/n \quad \pi_{ii}^m = \alpha(m) \quad \pi_{i \neq j}^m = \beta(m)$$

with the normalization $\sum_{i,j} \pi_{ij}^m = n\alpha + n(n-1)\beta = 1$

Then, one computes

$$\begin{aligned} E(U) &= \frac{n+1}{2n} \quad \text{Var}(U) = \frac{n^2-1}{12n^2} \\ \text{Var}(U) R_{uu} &= E(U(k)U(k+m)) - E(U)^2 \end{aligned}$$

As

$$E(U(k)U(k+m)) = \alpha \sum_{i=1}^n \frac{1}{i^2} + 2 \frac{1-n\alpha}{n(n-1)} \sum_{i>j} \frac{1}{ij}$$

one deduces

$$\alpha = \frac{n-1}{n^2} R_{uu} + \frac{1}{n^2} \quad (40)$$

Therefore (39) reads

$$R_{zz} = \left[\frac{n-1}{n} R_{uu} + \frac{1}{n} \right] R_{xx} \quad (41)$$

Appendice C: Computation of the autocorrelation in the swap method

The autocorrelation of a sequence $X_i, i = 1, \dots, n$ is given for a second order stationary sequence by

$$R_{XX}(k) = E[(X_i - E(X))(X_{i+k} - E(X))]/\text{Var}(X) \quad (42)$$

Its estimation for a finite sequence of N length is

$$d(k) = \frac{\sum_{j=1}^{N-k+1} (X_j - m)(X_{j+k} - m)}{\sum_{j=1}^N (X_j - m)^2 / (N-1)} \quad \text{with } m = \frac{\sum_{j=1}^N X_j}{N} \quad (43)$$

If we swap X_l and X_s the new estimation of autocorrelation reads now

$$d(k) \rightsquigarrow d(k) - \frac{(X_l - X_s)(X_{k+l} - X_{k+s})}{\sum_{j=1}^N (X_j - m)^2 / (N-1)} \quad (44)$$

Formula (44) is valid under the condition

$$(l+k) \text{ and } (s+k) < N \quad (45)$$

In the most common situations, the autocorrelation is specified over a range $M \ll N$. It is thus convenient to discard the swaps which do not satisfy this condition over the range M .

Appendice D: Generation of positive random variables given their cumulants

We do not present here the method based upon Edgeworth expansion due to its more involved implementation.

Following [12], one assumes that the distribution has a continuous density $f(x)$ defined over (α, β) , $\alpha, \beta \in (-\infty, \infty)$ with finite moments at all orders.

Over this interval, one presumes to develop $f(x)$ on a complete closed orthogonal base of polynomials, $P_i(x), i = 0, \dots, \infty$ with weight function $w(x)$ such that

$$\forall n, x \quad \left| f(x) - \sum_{i=0}^n a_i w(x) P_i(x) \right| \leq R_n(x) \quad (46)$$

a_i are the coefficients of the expansion and $\lim_{n \rightarrow \infty} R_n(x) = 0$ almost everywhere.

Due to the classical identity :

$$\int_{\alpha}^{\beta} w(x) P_n(x) P_m(x) dx = \delta_n^m$$

The coefficients a_i read

$$a_i = \int_{\alpha}^{\beta} f(x) P_i(x) dx \quad (47)$$

This expansion of $f(x)$ is an asymptotic one when it converges because the coefficients are independent of the truncation order. From (47) it appears that a_i depends upon the first i moments of $f(x)$, hence the first i cumulants.

Let us take μ_k the moment of order k of the distribution f . Moments μ_n and cumulants κ_n are related to the moment generating functions as

$$\hat{f}(u) = \sum_{k=0}^{\infty} \mu_k \frac{u^k}{k!} = \exp[\tilde{f}(u)] = \exp\left[\sum_{k=1}^{\infty} \kappa_k \frac{u^k}{k!}\right]$$

The moments et cumulants are thus linked by the Bell polynomials [23].

$$\mu_n = B_n(\kappa_1, \dots, \kappa_n) \quad (48)$$

$$B_n(x_1, \dots, x_n) = \sum_{\substack{k_1 + \dots + k_n = n \\ 1k_1 + \dots + nk_n = n}} \frac{n! x_1^{k_1} \dots x_n^{k_n}}{k_1! \dots k_n! (1!)^{k_1} \dots (n!)^{k_n}} \quad (49)$$

A direct computation is very involved. It is thus more convenient to use the recurrence⁶

$$B_n(x_1, \dots, x_n) = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k(x_1, \dots, x_k) x_{n-k} \text{ with } B_0 = 1 \quad (50)$$

After computing the value of the coefficients a_i , we have to check the convergence of the expansion $\sum_{i=0}^n a_i$ which implies the one of $\sum_{i=0}^n a_i w(x) P_i(x)$. At this stage it is noticeable that some computation difficulties arise for large i . In fact we are generally obliged to stop the expansion at a moderate order over which the numerical errors induce too important errors.

If it is true, we construct a dominating function $g(x)$ of $f(x)$ along the circumstances. Then we apply the following acceptance-rejection algorithm

```

repeat
generate  $X$  with density proportional to  $g$ 
generate a uniform  $[0, 1]$  random variate  $U$ 
set  $T \leftarrow Ug(X)$ ,  $n \leftarrow 0$ ,  $S \leftarrow 0$ 
  repeat  $S \leftarrow S + a_n w(X) P_n(X)$ ,  $n \leftarrow n + 1$ 
  until  $|S - T| > R_{n-1}(X)$ 
until  $T < S$ 
return  $X$ 

```

FIRST EXAMPLE : $(\alpha, \beta) = (0, \infty)$

One can chose the Laguerre polynomials associated with the weight

6. When considering in B_n the monomial including at least one x_j , the remaining is based only on x_1 to x_{n-j} . There are $\binom{n-1}{j}$ ways to chose such a disposition among n factors. Therefore one gets easily the recurrence $B_n = \sum_{j=0}^n \binom{n-1}{j} B_{n-j} x_j$ which is the same as noted in the main corpus

function e^{-x}

$$L_n(x) = \frac{e^x}{n!} \frac{d^n}{dx^n} [e^{-x} x^n] = \sum_{k=0}^n (-1)^k \binom{n}{n-k} \frac{x^k}{k!}$$

which obey to the recurrence relation

$$nL_n(x) = (2n-1-x)L_{n-1}(x) - (n-1)L_{n-2}(x) \quad n \geq 2$$

with $L_0 = 1$ and $L_1 = 1 - x$.

Let us write $c_k = \mu_k/k!$, simple calculation from (50) leads to

$$c_k = \sum_{j=0}^{k-1} \frac{c_j \kappa_{k-j}}{k(k-1-j)!} \quad (51)$$

and now from (47), a_i reads

$$a_i = \sum_{k=0}^i (-1)^k \binom{i}{i-k} c_k \quad (52)$$

SECOND EXAMPLE : $(\alpha, \beta) = (-1, 1)$

Many bases of polynomials are convenient. The most simple are the Legendre polynomials associated with the weight function $w(x) = 1$.

$$\begin{aligned} P_i(x) &= 2^{-i} \sum_{k=0}^{[i/2]} (-1)^k \frac{(2i-2k)!}{k!(i-k)!(i-2k)!} x^{i-2k} \\ &= 2^{-i} \sum_{k=0}^{[i/2]} (-1)^k \binom{i}{k} \binom{2i-2k}{i} x^{i-2k} \end{aligned}$$

where $[x]$ is the integer part of x .

These polynomials follow the recurrence

$$(i+1)P_{i+1}(x) = (2i+1)xP_i(x) - iP_{i-1}(x)$$

and the orthogonality relation

$$\int_{-1}^1 P_i(x) P_j(x) dx = \frac{2\delta_i^j}{2i+1}$$

Therefore, a_i reads now

$$a_i = 2^{-(i+1)} (2i+1) \sum_{k=0}^{[i/2]} (-1)^k \frac{(2i-2k)!}{k!(i-k)!} c_{i-2k} \quad (53)$$

where c_k has the same meaning than in (51).

Appendix E: Estimation of expression (32)

One starts from the identity

$$\begin{aligned} \sum_{i=1}^{\infty} [i(i+x)]^{-m} + \underbrace{\sum_{i=1}^{\infty} (-1)^i [i(i+x)]^{-m}}_A &= 2 \sum_{i=1}^{\infty} [2i(2i+x)]^{-m} \\ &= 2^{1-2m} \sum_{i=1}^{\infty} [i(i+x/2)]^{-m} \end{aligned} \quad (54)$$

The term A is an alternating series, thus simple to compute with a given accuracy : the value of the last retained term is less than the desired accuracy. Therefore, from x , every $\sum_{i=1}^{\infty} [i(i+2^n x)]^{-m}$ can be computed by iteration and thus, it just remains to evaluate the value for $x < 1$.

In this case, the classical expansion of $(1+x/i)^{-m}$ leads to

$$\sum_{i=1}^{\infty} [i(i+x)]^{-m} = \sum_{r=0}^{\infty} \binom{-m}{r} x^r \zeta(2m+r) \quad (55)$$

where the binomial coefficient is

$$\binom{-m}{r} = (-1)^r \frac{\Gamma(r+m)}{r! \Gamma(m)}$$

The expansion (55) is also an alternating series, and it is thus easy to control the accuracy of an extracted limited series.

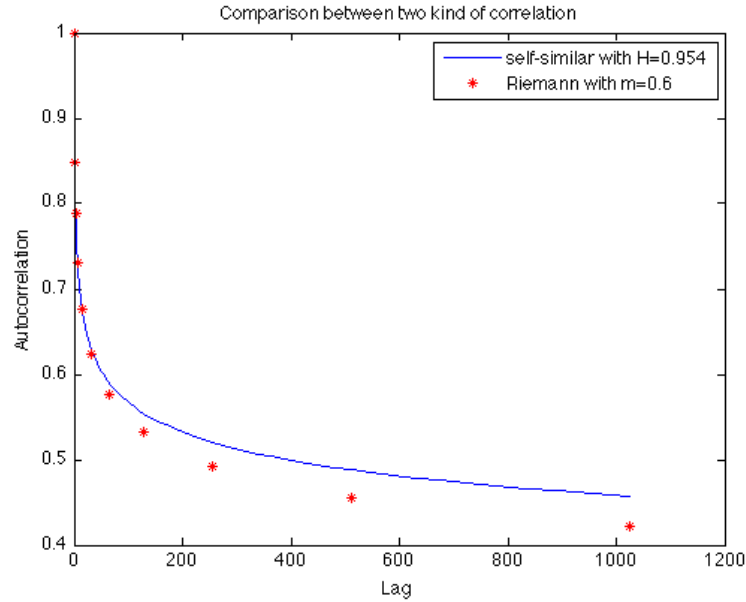


FIGURE 14 – Comparaision between autocorrelations of self-similar and Riemann family

Figure 14 shows that for moderate lags (<1000) the Riemann family presents autocorrelations close to the self-similar ones. In

this example the value of the Hurst parameter has been evaluated for minimizing the distance between the two autocorrelations. The chosen distance is the maximum of absolute difference between the two functions up to a lag of 1000.

The gotten distance is 0.036 which is a small value. Moreover, making the same work on $\approx 4 \cdot 10^6$ lags, it is noticeable that the Hurst parameter and the distance are nearly identical as those obtained for 1000 lags. In fact H is 0.954 in place of 0.953 and the distance increases up to 0.0367. Figure 15 shows the two functions.

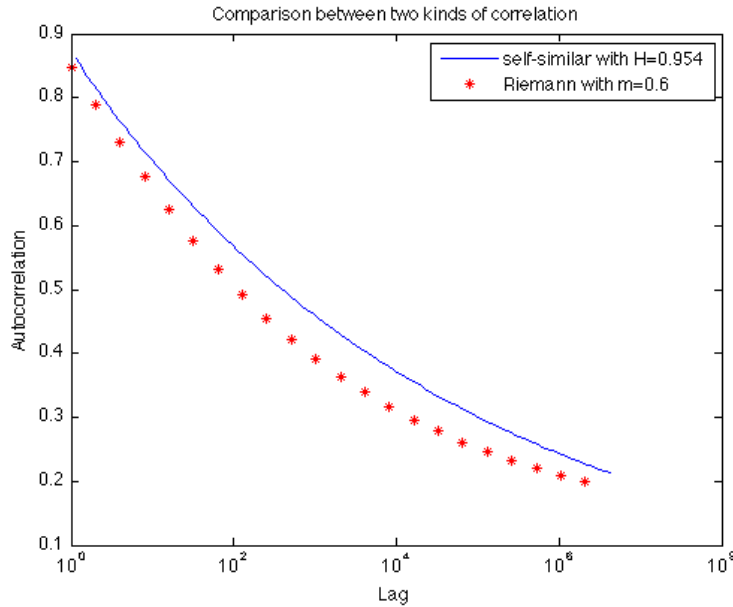


FIGURE 15 – Comparison between autocorrelations of self-similar and Riemann family upon large lags

Bibliographie

- [1] I. Popescu and I. Vaduva. A survey on computer generation of some classes of stochastic processes. *Math. Comp. Simul.*, 33 :223–241, 1991.
- [2] S. Chiaw-Hock. Simulation of weibull and gamma autoregressive stationary process. *Com. Stat. Simul*, 15(4) :1141–1146, 1986.
- [3] M. Bladel. Génération de nombres aléatoires corrélés. *math. Comp. Simul.*, 21 :42–49, 1979.
- [4] A. Blanc-Lapierre and B. Picinbono. *Fonctions Aléatoires*. Masson, 1981.
- [5] L. Devroye. *Non Uniform Random Variate Generation*. Birkhäuser, 1987.

- [6] V.G. Gugar and R.J. Kavanagh. Generation of random signals with specified probability density functions and power density spectra. *IEEE-AC*, 13 :716–719, 1968.
- [7] S.T. Li and J.L. Hammond. Generation of pseudo-random numbers with specified univariate distributions and correlation coefficients. *IEEE-SMC*, 5 :557–561, 1975.
- [8] R.J. Polge, E.M. Holliday, and B.K. Bhagavan. Generation of pseudo-random set with desired correlation and probability distribution. *Simulation*, 20(153-158), 1973.
- [9] V.C. Lakhan. Generating autocorrelation pseudo-random numbers with specific distributions. *J. Statist. Comp. Simul.*, 12 :303–309, 1981.
- [10] I.W. Hunter and R.E. Kearney. Generation of random sequences with jointly specified probability density and autocorrelation functions. *Biol. Cybern.*, 47 :141–146, 1983.
- [11] J. Lavergnat and P. Golé. A stochastic raindrop time distribution model. *J. Appl. Meteor.*, 37(805-818), 1998.
- [12] L. Devroye. On random variate generation when only moments or fourier coefficients are known. *Math. Comp. Simul.*, 31 :71–89, 1989.
- [13] W. Feller. *An Introduction to Probability Theory and Its Applications*. John Wiley and Sons, 2nd edition, 1971.
- [14] M. Kendall and A. Stuart. *The advanced Theory of Statistics*, volume I. Macmillan, 1977.
- [15] S. Blinnikov and R. Moessner. Expansions for nearly gaussian distributions. *Astron. Astroph. Suppl. Ser.*, 130 :193–205, 1998.
- [16] O. Cappé, E. Moulines, JC. Pesquet, A. Petropulu, and X. Yang. Long-range dependence and heavy-tail modeling for teletraffic data. *IEEE Signal Processing Magazine*, 19(3) :14–27, 2002.
- [17] F. Lombardo, E. Volpi, and D. Koutsoyiannis. Rainfall downscaling in time : Theoretical and empirical comparison between multifractal and hurst-kolmogorov discrete random cascades. *Hydrological Sciences Journal*, 57(6) :1052–1066, 2012.
- [18] D. Emmanoulopoulos, I. M. McHardy, and I. E. Papadakis. Generating artificial light curves : Revisited and updated. *MNRAS*, 443 :907–927, 2013.
- [19] W. Vervaat. Properties of general self-similar processes. *Bull. Int. Statist. Inst*, 52(4) :199–216, 1987.
- [20] G.L. O’Brien and W. Vervaat. Marginal distributions of self-similar processes with stationary increments. *Z. Wahrscheinlichkeitstheorie verw.*, 64 :129–138, 1983.

- [21] J.W. Lamperti. Semi-stable stochastic processes. *Trans. Amer. Math. Soc.*, 104(1) :62–78, 1962.
- [22] A.T.A. Wood and G. Chan. Simulation of stationary gaussian process in $[0, 1]^d$. *Journal of Computational and Graphical Statistics*, 3 :409–432, 1994.
- [23] L. Comtet. *Advanced Combinatorics*. Dordrecht, 1974.