Stationnary multiscale graphical models

Philippe St-Jean[*] Jean-Marc Lina[*]
Bernard Goulard[*]

CRM-2909

January 2003

[*]Centre de recherches mathématiques, Université de Montréal Case postale 6128, Succursale centre-ville, Montréal (Québec), H3C 3J7, Canada

Abstract

We present a stationnary model for cascading process defined on dyadic digraphs. This model is intended to describe self-similarity of certain physical processes, such as those encountered in the field of fully-developed turbulence. Those processes respect a differential equation which displays scale symmetry to a certain extent, such as the Navier-Stokes equation. However, the "cascade" model actually used to describe those self-similar processes omit to conform to another fundamental symmetry of the system: invariance under translation. Hence the cascade model generates non-stationnary processes, whereas the non-stationnarity appears in sole consequence of this oversimplifying model.

We propose to modify the standard cascade model by stating a sufficient condition that ensures stationnarity of the process, while keeping the marginal distribution of the coefficients intact such that the scaling behavior of the process remains unchanged.

Keywords. stationnary process; multi-scale

Mathematical Subject Classification. Primary 60G10, Secondary 60K40

1 Introduction

We present a stationnary model for cascading process defined on dyadic digraphs. The motivation for this paper originated from the work of Arnéodo et al. on cascade model in fully developed turbulence, which was introduced to describe multi-fractal behavior (intermittency) in wind tunnel velocity field signals [1, 2, 3]. The model allows to simulate velocity fields with parameter values (such as the Reynold number) that cannot be attained experimentally. For instance there is interest in predicting the behavior when the Reynold number goes to infinity, requiring either the viscosity tending towards zero or the size of the tunnel to be arbitrarily large [5]. The model is defined as a Markov dyadic tree over an orthogonal wavelet decomposition of the velocity signal, and is referred to as a $\mathcal{W}$-cascade model [2].

The $\mathcal{W}$-cascade model is intended to describe self-similarity characteristics of processes. Those processes respect a differential equation which displays scale symmetry to a certain extent, as for the Navier-Stokes equation in the case of turbulence. However, the cascade model omits to conform to another fundamental symmetry of the system: invariance under translation. This invariance is typically verified experimentally [6]. However the $\mathcal{W}$-cascade generates non-stationnary processes, whereas the non-stationnarity appears in sole consequence of an oversimplified model.

Hence $\mathcal{W}$-cascade model on wavelet dyadic trees lacks the stationnarity property, while the stochastic processes to be described are stationnary. A problem arises when it comes to the estimation of the spatial correlation function on simulated data. This problem is usually put aside in a somewhat unelegant fashion by averaging correlation estimators along the spatial dimension, with little or no justification for doing so.

In a different context, efforts were put forward in order to generate "almost stationnary" processes with a multi-scale graphical modelling approach, such as in [7] where a Kalman-filter prediction algorithm on multiscale graphs is studied. Their approach results in apparent stationnarity of the process, meaning that the form of the covariance function "looks" stationnary, i.e. fluctuations of the covariance function when changing the position where it is evaluated are small relative to pointwise variance. This was not a problem in [7] as their algorithm did not require stationnarity. In the case of fully developed turbulence such an approach is invalid; the problem is just hidden deeper, as higher order (n-point) moments are generally not guaranteed to respect this apparent stationarity behavior. Furthermore, one might wonder what kind of behavior on the spatial covariance function (or on higher order moments) should be expected when strict stationnarity of the model is imposed. The main goal of our work is to elaborate a model which is strictly stationnary, while conserving the self-similarity properties of the previous models, those being enlisted by the behavior of pointwise statistics (i.e. on every single node) through scales. For instance, quantities such as the singularity spectrum of the process must be preserved [1, 9].

One important drawback of non-stationnary models on dyadic multiscale graphs lies in the fact that the topology of the dyadic graph becomes visible in the process, since its spatial covariance function depends strongly on the position of the node in the tree. This artefactual structure is of course never observed on the processes one tries to model. This flaw becomes unacceptable when one is interested in modelling the precise form of the spatial correlation function of a wavelet cascade process, or the correlation between dyadic wavelet coefficients.

We propose to modify the standard cascade model by stating a sufficient condition that ensures stationnarity of the process, while keeping the marginal distribution of the coefficients intact. We begin by presenting some non-stationnary models which will be used as a starting point. Then we derive a sufficient condition in this context, assuming that the model is gaussian. We move on to prove that this condition is also sufficient in a much wider context than the gaussian case, i.e. that stationnarity of higher order moments is ensured provided some extra conditions are imposed.

We show that our solution also brings increased versatility to the model, as it allows to control the shape of the spatial correlation function to a certain extent (something that was not possible with the previous models). Then we derive estimators for the extra parameters involved. Finally we show comparison between non-stationnary models and our model on simulated data.

2 Nearest-neighbor models

We begin by describing the two simplest non-stationnary models (which we hereafter refer to as nearest-neighbor models). The first one is defined on a tree structure such as the one depicted on figure (1), a single-rooted binary tree. This tree structure is commonly seen in discrete dyadic wavelet decomposition on orthonormal basis, whereas each node represents a wavelet coefficient. The coefficients are arranged in levels which represent scales. One readily sees that every node is the parent of exactly two child nodes. The cascade process is defined as follows: Let $c_{0,0} = 0$ be the coefficient at the root (top) of the tree, and $c_{j,k}$ the k-th coefficient at level j. Then any $c_{j,k}$ is generated by adding a $W_{j,k}$ random variable to its parent coefficient, i.e.:

$$c_{j,2k} \quad = \quad c_{j-1,k} + W_{j,2k}$$

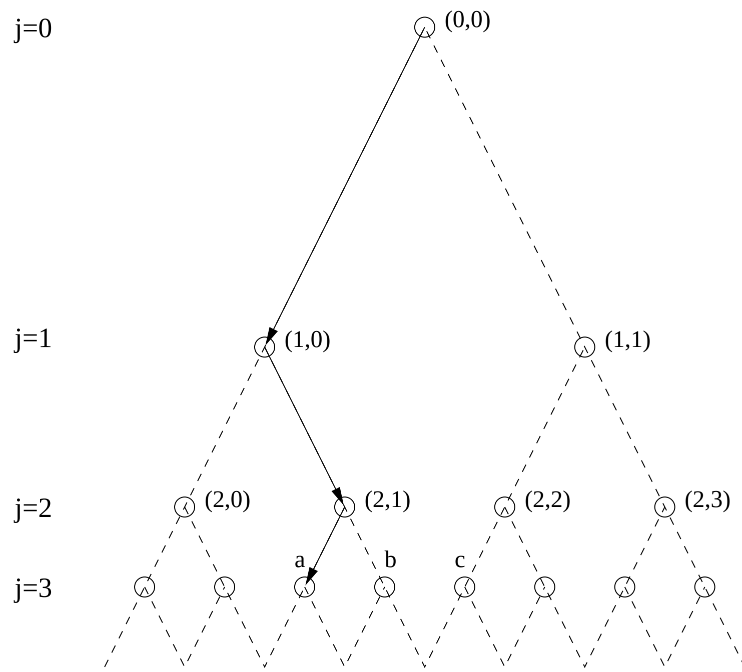

Figure 1: The single parent model. The solid line indicates ancestors of coefficient a.

$$c_{j,2k+1} = c_{j-1,k} + W_{j,2k+1}.$$

We mention here that if W's are iid gaussian variables, then every coefficient will also be gaussian, and that the whole tree of coefficients defines a multi-gaussian variable. We will assume that this is the case until we reach section 5. At any level j of the tree, it is possible to define a discrete spatial covariance function C as:

$$C^{(j)}(k, \Delta k) = E[c_{j,k} \cdot c_{j,k+\Delta k}],$$

with the additional assumption that the W's are zero-mean. In this model, C is a function of j, Δk and k as well. This means that the covariance of two coefficients on the same level depends not only on their relative distance (Δk), but also on their position in the tree. For instance, coefficients a and b on figure(1) show stronger correlation than coefficients b and c as the former share an immediate common parent, although relative distance between a and b and between b and c are the same. The stationnarity constraint we want to impose is for the covariance C to be a function of j and Δk solely.

We now present another dyadic model, depicted on figure (2). This model is no longer a tree, but still is an acyclic digraph with a single source. In this case, half of the coefficients have only one parent, while the others have two. The expression for those with one parent (the ones with even index k) is the same as in the single-parent model above, while the odd-indexed coefficients are given by the second expression:

$$\begin{aligned} c_{j,2k} &= c_{j-1,k} + W_{j,2k} \\ c_{j,2k+1} &= \alpha \cdot c_{j-1,k} + \beta \cdot c_{j-1,k+1} + W_{j,2k+1} \end{aligned}$$

where α and β can take any real value (typically 1/2 for both). Again, it is obvious that this model is not stationnary in our sense, as the covariance for a pair of even-indexed coefficients will in the general case be different from a pair of odd-indexed ones (as the latter share a parent).

This last model (the "one-two" parent model) will be our starting point in building a stationnary model, although we might as well have chosen the preceeding one. The only difference lies in the computation of the stationnarity constraints which happens to be somewhat simpler to perform using the one-two parent model.

3 Stationnary model (α-model)

In order to obtain stationnarity, it is necessary to add complexity to the model. In the case of nearest-neighbor models, a child coefficient $c_{j,k}$ can be considered as depending solely on the value taken by its direct parents $\gamma(c_{j,k})$

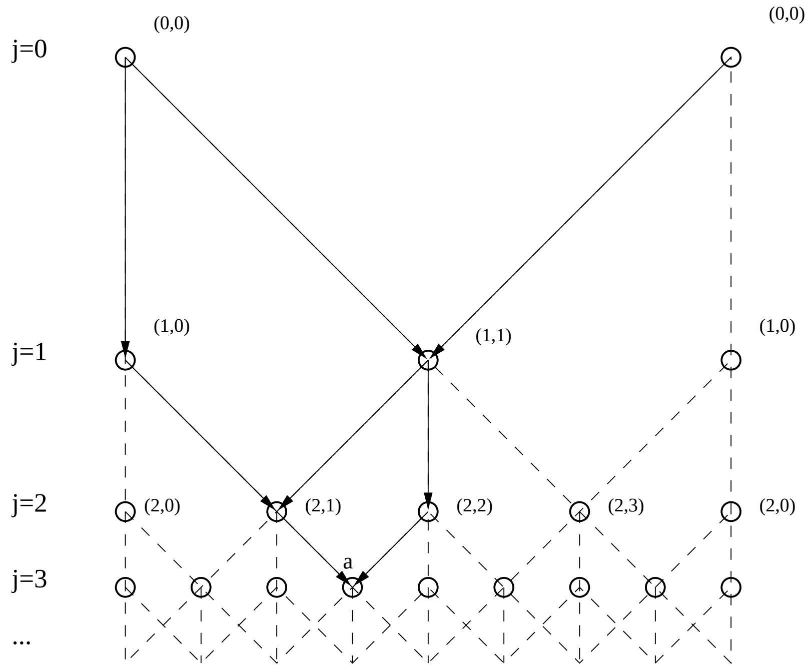

Figure 2: The one-two parent model. Solid lines indicates all ancestors of coefficient a.

(either one or two of them, as above). This is equivalent to a first-order markovian relationship along branches of the tree, of the type (for even-indexed):

$$P(c_{j,k}|\Omega_{j-1}) = P(c_{j,k}|\gamma(c_{j,k}))$$

where Ω_j represents all coefficients on levels $0, 1, ..., j$. However, $c_{j,k}$ will also be correlated to coefficients in Ω_j other than its direct parents. By adding possible direct contribution from other ancestors in the model, one can recover all the properties of nearest-neighbor models (in terms of marginal distributions and parent-child covariances, for instance), while gaining stationnarity. It is actually possible to do so by limiting the influence of direct contribution to the coefficients found at the preceeding level only, i.e. by forcing the digraph to allow edges between nodes on adjacent levels only.

One simple way to achieve this is to impose translation invariance on the covariance of coefficients found on the same level, by exploiting symetries of the tree structure at hand.

In the model we present, the coefficients are written as:

$$c_{j,k} = W_{j,k} + \sum_{k'=0}^{2^{j-1}-1} \alpha^{(j)}_{k,k'} c_{j-1,k'} \tag{1}$$

with $c_{0,0}$=0. The $c_{j,k}$ now depend directly on all coefficients found on level $j-1$. The value of $\alpha^{(j)}_{k,k'}$ for a given triplet (j, k, k') corresponds to the weigth associated to the edge going from $c_{j-1,k'}$ to $c_{j,k}$. We refer to such a model as an α-model. A graphical representation is shown on figure 3.

We still consider the $W_{j,k}$ to be identically distributed, although not necessarily independent from one another (yet independent from the c_{j-1}).

We ask for $\alpha^{(j)}_{k,k'}$ to depend only on some distance function of k and k', i.e.

$$\alpha^{(j)}_{(k,k')} \quad = \quad \alpha^{(j)}_{d(k,k')}.$$

In our case, the obvious choice for $d(k, k')$ is $d(k, k') = |2k' - k|$, which corresponds to the actual distance along the k-axis on fig(3). For simplicity we assume that α is symmetric over its index (i.e. $\alpha_{-i} = \alpha_i$), so that we can drop the absolute value in the distance function. This is done without loss of generality, as the derivation in the following section also holds for the anti-symmetric part of α. Usage of non-symmetric form for α will be discussed further in section 4.1.

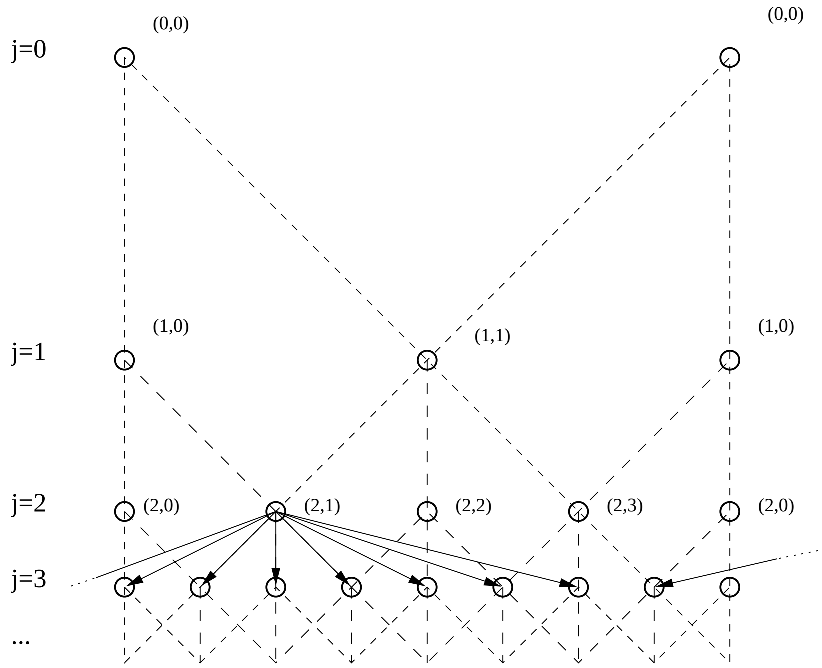

Figure 3: The α model.

4 Stationnarity constraints

From now on we impose that the $W_{j,k}$ are i.i.d. $N(0, \sigma_W)$, without loss of generality. At a given scale j we want the c_j's to share the same marginal distribution $(N(\mu(j), \sigma(j)))$; the form of the pdf is guaranteed to be Gaussian since each coefficient is expressed as a linear combination of the $W_{j,k}$. The zero-mean condition on the W's also implies that $\mu(j) = 0,\ \forall j$. This has no influence on stationnarity, which depends solely on second order moments.

The task is to build a model for which coefficients found at the same level have covariances depending on their mutual distance only; in order to impose this condition, we proceed by induction. We assume that the inner-level covariance for level $j-1$ $\text{Cov}[c_{j-1,k},\ c_{j-1,k+\Delta k}]$ is a function of Δk (and j) only, and demand that the inner-level covariance at level j respects the same condition.

This covariance is written as

$$C^{(j)}(k, \Delta k) = C_1^{(j)}(k, \Delta k) + \text{Cov}_W(\Delta k),$$

where

$$C_1^{(j)}(k, \Delta k) \equiv E\left[\left(\sum_{k'=0}^{2^{j-1}-1} \alpha_{2k'-k}^{(j)}\ c_{j-1,k'}\right)\left(\sum_{k''=0}^{2^{j-1}-1} \alpha_{2k''-(k+\Delta k)}^{(j)}\ \ c_{j-1,k''}\right)\right], \tag{2}$$

and where

$$\text{Cov}_W(\Delta k) = E[W_{j,k} \cdot W_{j,k+\Delta k}]$$

is assumed to depend solely on Δk. Thus the process will be stationnary when $C_1^{(j)}(k, \Delta k)$ depends only on Δk for all j.

We can rewrite (2), using a simple change of variables (from k'' to $k' + k''$):

$$C_1^{(j)}(k, \Delta k) = \sum_{k'=0}^{2^{j-1}-1} \sum_{k''=0}^{2^{j-1}-1} \alpha_{2k'-k}^{(j)}\ \alpha_{2(k''+k')-(k+\Delta k)}^{(j)}\ C^{(j-1)}(0, k''), \tag{3}$$

where we have used the induction hypothesis in the rightmost part of the last expression, as it corresponds exactly to the covariance at level $j-1$.

As we try to make this expression independent of k, we shall first notice that replacing k by $k+2m,\ m \in \mathbb{Z}$ does not change the expression, thanks to the periodicity of α. However, replacing k by $k+1$ (or by $k+2m+1$, which is equivalent) modifies the expression. This means that imposing the expression to be invariant when changing k in $k+1$ is sufficient to obtain stationnarity.

We will first consider the case where Δk takes an even value. The two α's in (3) then share parity on their index, either both even or both odd. Let η and ξ be even and odd sub-samplings of α, respectively:

$$\begin{aligned} \eta_i^{(j)} &= \alpha_{2i}^{(j)} \\ \xi_i^{(j)} &= \alpha_{2i-1}^{(j)} \end{aligned}$$

For C_1 to remain unchanged when either $k=0$ or $k=1$ is to impose

$$E[c_{j,0}\ c_{j,\Delta k}] = E[c_{j,1}\ c_{j,1+\Delta k}].$$

Since $\Delta k = 2m,\ m \in \mathbb{Z}$, we write

$$\begin{aligned} &\sum_{k'=0}^{2^{j-1}-1} \sum_{k''=0}^{2^{j-1}-1} \eta_{k'}^{(j)}\ \eta_{k''+k'-m}^{(j)}\ C^{(j-1)}[k''] = \\ &\quad \sum_{k'=0}^{2^{j-1}-1} \sum_{k''=0}^{2^{j-1}-1} \xi_{k'}^{(j)}\ \xi_{k''+k'-m}^{(j)}\ C^{(j-1)}[k''] \end{aligned}$$

where $C^{(j-1)}[k''] = E[c_{j-1,k'} \cdot c_{j-1,k''+k'}]$ is even around $k''=0$.

The last equation is written in terms of circular convolutions:

$$\left(\eta^{(j)} \circledast \eta_-^{(j)} \circledast C^{(j-1)}\right)[m] = \left(\xi^{(j)} \circledast \xi_-^{(j)} \circledast C^{(j-1)}\right)[m], \tag{4}$$

where $(\eta_-)_i = \eta_{-i}$. Through Fourier transform, this becomes (we omit the j's for simplicity):

$$\hat\eta[l] \cdot \hat\eta[-l] \cdot \hat C[l] = \hat\xi[l] \cdot \hat\xi[-l] \cdot \hat C[l].$$

$\hat C$ and $\hat\eta$ are both real and even, since C and η are (η is the even sub-sampling of α, also even). ξ is also real, but not even. Hence $\hat\xi$ can be written in polar form, $\hat\xi[l] = \hat A_\xi[l] e^{i\theta_\xi[l]}$, with $\hat A_\xi$ and θ_ξ even and odd, respectively.

Regardless of the form of $\hat C$, the constraint will be verified when

$$\hat A_\xi^2[l] = \hat\eta^2[l]. \tag{5}$$

This means that $\hat\eta$ and $\hat\xi$ are the same, up to a complex phase. The phase θ_ξ will be such that ξ is even around $1/2$. We lost necessity of the condition in this last step, as possible zeros of $\hat C$ could soften the condition on $\hat\eta$ and $\hat\xi$.

Repeating the same exercise for $\Delta k = 2m+1,\ m \in \mathbb{Z}$, one falls on the following expression:

$$\left(\eta^{(j)} \circledast \xi_-^{(j)} \circledast C^{(j-1)}\right)[m] = \left(\xi^{(j)} \circledast \eta_-^{(j)} \circledast C^{(j-1)}\right)[m+1],$$

which transforms to

$$\hat\eta[l] \cdot \hat\xi[-l] \cdot \hat C[l] = e^{\frac{2\pi i l}{N}} \hat\xi[l] \cdot \hat\eta[-l] \cdot \hat C[l]$$

where $N = 2^{j-1}$. By simplifying both $\hat C$ and $\hat\eta$ in this last expression, all that remains is the same condition on the phase of $\hat\xi$ as mentionned above. Hence no extra condition is to be imposed on η and ξ.

We conclude that the process will be stationnary for any choice of even $\hat\eta$, provided that the modulus of $\hat\xi$ is the same, and that its phase is expressed as

$$\begin{aligned} \hat\xi[l] &= \hat\eta[l] \cdot e^{i\theta_\xi[l]}, \\ \theta_\xi[l] &= \tfrac{-\pi l}{N}, \qquad l = -N/2+1 \ldots N/2. \end{aligned} \tag{6}$$

This means for instance that one can choose values for even coefficients in α, and compute odd ones by performing a frequency shift in Fourier space.

4.1 Symmetry of α

We have assumed so far that α was symmetric over its index. The derivation above can be repeated with the anti-symmetric part of α, with the same result. The only effect of removing the symmetry condition on α will be that (6) now represents the frequency shift between $\hat{\eta}$ and $\hat{\xi}$ rather than the exact phase of the latter, as $\hat{\eta}$ is now allowed a non-null complex phase.

Non-symmetric behavior of α can be used in the modelling of past-future dependencies for time-based cascading processes, as is the case in modelling fluctuations of financial time series. See for instance [4, 9].

4.2 Interpretation of the "frequency shift" condition

We will now show that the "frequency shift" condition (6) tells us that the support of $\hat{\alpha}$ is limited to half of its length, i.e. α is a band-limited discrete function.

We begin by upsampling both η and ξ by a factor of two (by introducing zeros at every odd-indexed value), and then shifting ξ by one, such that the sum of the two leads to an expression for α:

$$\alpha = \eta_{\uparrow 2} + \xi_{\uparrow 2,\text{shft}} \tag{7}$$

The upsampling of η and ξ implies that their Fourier transform is simply duplicated on the domain of $\hat{\alpha}$, which is appropriately twice as large. The shift on ξ introduces a frequency shift in the Fourier domain. Hence we have

$$\begin{aligned} \widehat{(\eta_{\uparrow 2})} &= \hat{\eta}, \\ \widehat{(\xi_{\uparrow 2,\text{shft}})} &= \hat{\eta} \cdot e^{\frac{-2\pi i l}{2^j}} \cdot \begin{cases} e^{\frac{-\pi i l}{2^{j-1}}} & l = -(2^{j-2}) + 1 \ldots 2^{j-2} \\ e^{\pi i - \frac{\pi i l}{2^{j-1}}} & l = 2^{j-2} + 1 \ldots 2^{j-1} \\ e^{-\pi i - \frac{\pi i l}{2^{j-1}}} & l = -(2^{j-1}) + 1 \ldots - (2^{j-2}) \end{cases}, \end{aligned} \tag{8}$$

where it is understood that $\hat{\eta}$ is duplicated over the $[-(2^{j-1}) + 1, 2^{j-1}]$ support of $\hat{\alpha}$. The phase factors in (8) cancel out and give

$$\widehat{(\xi_{\uparrow 2,\text{shft}})} = \hat{\eta} \cdot \begin{cases} 1 & l = -(2^{j-2}) + 1 \ldots 2^{j-2} \\ -1 & l = 2^{j-2} + 1 \ldots 2^{j-1} \\ -1 & l = -(2^{j-1}) + 1 \ldots - (2^{j-2}) \end{cases}$$

Thus the sum in (7) is expressed in the Fourier domain as

$$\hat{\alpha} = 2 \cdot \hat{\eta} \cdot \chi^{(j-1)}, \tag{9}$$

where $\chi^{(j-1)}$ is the characteristic function over $[-(2^{j-2}), 2^{j-2}]$.

Hence the support of $\hat{\alpha}$ is limited to half its domain, which makes α a band-limited function. This implies that the support of α itself is of full length (i.e. 2^j). An important consequence of this condition is that it is not possible to generate a stationnary process with only one or two parents at each node, which in returns explains why the nearest-neighbor models could not be stationnary. For instance the one-two parent model, where α is given as $[0 \ldots 0, 1/2, 1, 1/2, 0 \ldots 0]$, is not band-limited and thus could not lead to a stationnary process.

An interpretation of this condition is found by looking at the set of nodes at scale $j-1$ as a band-limited process; since it only regroups 2^{j-1} nodes, it cannot contain frequencies higher than that number. Consequently it cannot generate by itself higher frequencies unless some additional random variables (such as the W's) are added. Trying to increase artificially the frequency content using solely the nodes at scale $j-1$ always results in creating two different processes, one for odd and one for even nodes at scale j that are then simply interlaced, leading to the artefactual non-stationnarities we are trying to avoid. This makes for an intuitive explanation of why the support of α must be limited to half its length in the frequency domain.

5 Stationnarity of higher order moments

So far we have proven that it is possible to ensure strict-sense stationnarity at least for Gaussian processes, as second-order moments completely describe the statistics in this case. The question arises of whether the same condition implies stationnarity of higher order moments as well (i.e. $p \geq 3$) for non-gaussian distributions, with the benefit of providing stronger stationnarity for a much larger class of marginal distributions on the nodes. It is shown in this section that this is the case to a certain extent.

Let us express this comment formally:

Proposition 1 *For an α-model as defined in (1), with α respecting conditions (5, 6), then:*

$$E[c_{j,k}\ c_{j,k+\Delta k_1} \dots c_{j,k+\Delta k_{p-1}}]$$

does not depend on k, $\forall \Delta k_i \in \mathbb{Z},\ i \in [1 \dots p-1]$, provided that the expected value above exists and that the following condition holds at every scale:

$$\hat{\eta}^{(j)}\left[\sum_i^{p-1} l_i\right] \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] = \tag{10}$$

$$\hat{\eta}^{(j)}\left[\sum_i^{p-1} l_i\right] \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] \cdot \exp\left(i\theta[\sum_i l_i] - i\sum_i \theta[l_i]\right),$$

$\forall l_i \in [0, 2^{j-1}-1]$, $i \in [1 \dots p-1]$.

The proof can be found in the Appendix.

Let us look at (10) and see what it implies. In the case of $p=2$, the expression is trivially respected and no new condition is imposed, as expected. For $p \geq 3$ we see that (10) will be respected as long as $\theta[l]$ is a linear function of l. Although this is indeed the case for $l \in [-2^{j-2}+1, 2^{j-2}-1]$ where $\theta[l] = -\pi l/2^{j-1}$, the sum in the argument of $\hat{\eta}$ in the LHS of (10) can go beyond this interval. As $\theta[l]$ is reproduced periodically outside the interval, it is easily checked that an extra factor of -1 will appear for certain combinations of values for the l_i. This is the case when $l_1 = l_2 = 2^{j-2}-1$ with $p=3$, for instance. Hence a new condition must be imposed on $\hat{\eta}$ for $p \geq 3$.

A sufficient condition for (10) is derived here. The extra -1 factor appears only when $\sum l_i$ goes out of the $[-2^{j-2}+1, 2^{j-2}-1]$ interval. However we can force $\hat{\eta}[l]$ to be zero on enough points such that the sum of indices on non-zero values never grows out of the interval. For instance with $p=3$, limiting the support of $\hat{\eta}$ to half of its length will achieve this goal, since the sum involves exactly two terms. For any order p, this will hold when limiting the support of $\hat{\eta}$ to a $1/(p-1)$ fraction of its length.

Hence in order to provide high-order moment stationnarity of the model when the marginal distribution on the nodes is not gaussian, it is necessary to "start" the cascading process at a scale where there is enough nodes such that the sufficient condition derived above is not trivial, whereas $\hat{\eta}[l]$ would be forced to be null everywhere except at $l=0$.

This result brings greater generality to the model, as it allows to obtain stationnary of higher order moments for multi-scale process with any non-gaussian multivariate distribution, provided these moments exist.

The proposed α-model can be considered as a method of removing structures induced by the digraph in a multi-scale Markov model rather than being limited to provide wide-sense stationnarity, as the stationnarity constraints do not depend on the specific form of the distribution on the nodes.

6 Form of the covariance function

The stationnarity constraint derived just above leads to some specific behavior of the covariance function, which we describe in this section. Assuming that we have α respecting the stationnarity condition (9), we write the covariance at level j as

$$C^{(j)}[\Delta k] = C_1^{(j)}[\Delta k] + C_W^{(j)}[\Delta k]. \tag{11}$$

For even values of Δk, we can rewrite (4):

$$C_1^{(j)}[2m] = \left(\eta^{(j)} \circledast \eta^{(j)} \circledast C^{(j-1)}\right)[m]. \tag{12}$$

Whereas for odd Δk, (6) gives

$$C_1^{(j)}[2m+1] = \left(\eta^{(j)} \circledast \xi^{(j)} \circledast C^{(j-1)}\right)[m].$$

Considering the last two expressions as null for odd and even values of Δk respectively, we see that adding them together gives the full expression for C_1. Using Fourier transform once again and following the same line as in section 4.2,

$$\hat{C}_1^{(j)}[l] = \hat{C}^{(j-1)}[l] \cdot \hat{\eta}^{(j)}[l] \cdot \left\{\hat{\eta}^{(j)}[l] + e^{\frac{2\pi i l}{2^j}} \hat{\xi}^{(j)}[l]\right\}.$$

The length of the support of $\hat{C}_1^{(j)}$ is 2^j, while that of $\hat{\eta}^{(j)}$, $\hat{\xi}^{(j)}$ and of $\hat{C}^{(j-1)}$ is 2^{j-1}; since those functions are all periodic, the latter are simply replicated twice on the support of $\hat{C}_1^{(j)}$.

By (11), we get a recursive form:

$$\hat{C}^{(j)}[l] = \hat{C}^{(j-1)}[l] \cdot \hat{\eta}^{(j)}[l] \cdot \left\{ \hat{\eta}^{(j)}[l] + e^{\frac{2\pi i l}{2^j}} \hat{\xi}^{(j)}[l] \right\} \; + \; \hat{C}_W^{(j)}[l]$$

Coming back to polar notation for $\hat{\xi}$, and using (5):

$$\begin{aligned} \hat{C}^{(j)}[l] &= \hat{C}^{(j-1)}[l] \cdot (\hat{\eta}^{(j)}[l])^2 \cdot \left\{ 1 + e^{\frac{2\pi i l}{2^j}} \cdot e^{i\theta_\xi[l]} \right\} \; + \; \hat{C}_W^{(j)}[l] \\ &= \hat{C}^{(j-1)}[l] \cdot (\hat{\eta}^{(j)}[l])^2 \cdot 2\chi^{(j-1)} \; + \; \hat{C}_W^{(j)}[l]. \end{aligned} \tag{13}$$

Iterating this last expression, we get a formula for the covariance at any level:

$$\hat{C}^{(j)}[l] \;=\; \left[\sum_{j'=1}^{j-1} 2^{j-j'} \cdot \hat{C}_W^{(j')}[l] \cdot \chi^{(j')}[l] \prod_{j''=j'+1}^{j} (\hat{\eta}^{(j'')}[l])^2 \right] \; + \; \hat{C}_W^{(j)}[l]$$

This last equation is quite general; it describes the form of the covariance function for any stationnary process based on our α-model.

In order to generate a realization of one such process, it is necessary to choose (or estimate) both $\hat{\eta}^{(j)}[l]$ and $\hat{C}_W^{(j)}[l]$. We start by a simple example to show what form of covariance function can be expected. It is readily seen that the following choice for $\hat{\eta}$ is acceptable:

$$\hat{\eta}[l] \;=\; 1 \tag{14}$$

This is just a delta function for η. α, however, will not be a delta function because of the contribution of ξ. α will rather be a "low-pass" delta function. We also consider a delta function for C_W (i.e. W's are i.i.d.), which is admissible as there is no constraint over C_W.

Equation (13) then simplifies to

$$\hat{C}^{(j)}[l] \;=\; \left[\sum_{j'=1}^{j-1} 2^{j-j'} \; \chi^{(j')}[l] \right] \; + \; 1,$$

and we obtain a "staircase" form for the Fourier Transform of the covariance function $\hat{C}$. (fig. (4)). We see that the covariance function behaves like $\log_2(\hat{C}^{(j)}[l]) \sim j - \log_2[l+1]$, i.e. $\hat{C}^{(j)}[l] \sim \frac{j}{l+1}$. Of course, a function with a smoother decrease for $\hat{\eta}$ will result in a smoother $\hat{C}$ too.

7 Estimators for $\hat{\eta}$ and $\hat{C}_W$

We shall now describe how to estimate parameters of the model given a realization of the process. Estimators for $\hat{\eta}$ and $\hat{C}_W$ are sufficient to completely characterize the α-model under the gaussian assumption; this is understood from the fact that an expression for pairwise covariances of all wavelet coefficients can be derived from $\hat{\eta}$ and $\hat{C}_W$, regardless of whether the coefficients belong to the same level or not.

Indeed, in order to obtain a simple estimator for $\hat{\eta}$ we begin by computing the covariance function between coefficients found on two adjacent levels:

$$C^{(j-1,j)}[\Delta k] \equiv \mathrm{Cov}\left(c_{j-1,k'}, c_{j,k}\right),$$

where $\Delta k = 2k' - k$ still. We will only need this covariance for even values of Δk, i.e. the even subsampling $C_e^{(j-1,j)}[\Delta k]$. Then we have:

$$\begin{aligned} C_e^{(j-1,j)}[\Delta k] &= \sum_{k'=0}^{2^{j-1}-1} \alpha_{2k'-\Delta k}^{(j)} C^{(j-1)}[k'] \\ &= \left(\eta^{(j)} \circledast C^{(j-1)} \right) \left[\frac{\Delta k}{2} \right] \end{aligned}$$

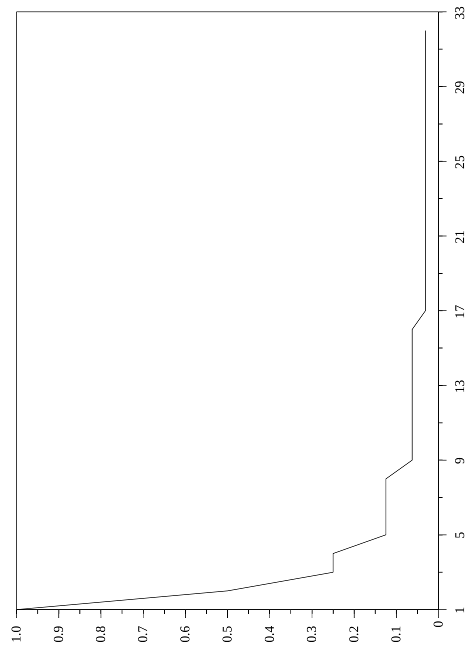

Figure 4: $\hat{C}[l]$ for $\hat{\eta}$ defined in (14).

Considering that $C_e^{(j-1,j)}$ is null for odd values of Δk, we have $\hat{C}_e^{(j-1,j)}[l] = \hat{\eta}^{(j)}[l] \cdot \hat{C}^{(j-1)}[l]$. Hence we obtain an estimator for $\hat{\eta}^{(j)}[l]$:

$$\hat{\eta}^{(j)}[l] = \frac{\hat{C}_e^{(j-1,j)}[l]}{\hat{C}^{(j-1)}[l]} \tag{15}$$

For $\hat{C}_W$, we use equation (12) to write:

$$\hat{C}_W^{(j)}[l] \quad = \quad \hat{C}^{(j)}[l] - \hat{C}^{(j-1)}[l] \cdot (\hat{\eta}^{(j)}[l])^2 \cdot 2\chi_{[-(2^{j-2}),2^{j-2}]}$$

It is possible to get a simpler expression for the even sub-sampling of $\hat{C}_W$:

$$\begin{aligned} \hat{C}_{W,e}[l] &= \hat{C}_e^{(j)}[l] - \hat{C}_e^{(j-1)}[l] \cdot (\hat{\eta}^{(j)}[l])^2 \\ &= \hat{C}_e^{(j)}[l] - \hat{C}_e^{(j-1,j)}[l] \cdot \hat{\eta}^{(j)}[l] \end{aligned} \tag{16}$$

where we used (15) in order to obtain (16).

8 Simulations and analysis

In this section, we compare the behavior of non-stationnary models for cascading process with our proposed stationnary model. To that end, we have generated multiple realizations of the process for each model, estimated correlations between nodes (i.e. pairs of nodes at distance Δk from each other, both nodes being found at the same scale) at the finest scale $j = J$ and plotted their values as a function of the position k of one of the node in the graph.

We have generated three series of process realizations, one with the single-parent model, one with the one-two parent model and the other using the stationnary α-model. For the latter we chose the simple case of $\hat{\alpha}$ being a truncated gaussian window and C_W a delta function. This was done over 8 scales ($j = 1...8$), and the correlations are estimated on nodes found at the last scale $j = 8$. Figure (5) shows estimation of the correlation of the $c_{j,k}$'s as a function of position k and inter-nodal distance Δk for the single-parent model. One very well sees the dyadic structure of the tree. This process is highly non-stationnary; for instance certain nodes are almost completely uncorrelated ($\rho \simeq 0$) while most are strongly correlated ($\rho > 0.6$) when $\Delta k = 1$. All figures are subsampled by a factor of 4 for clarity.

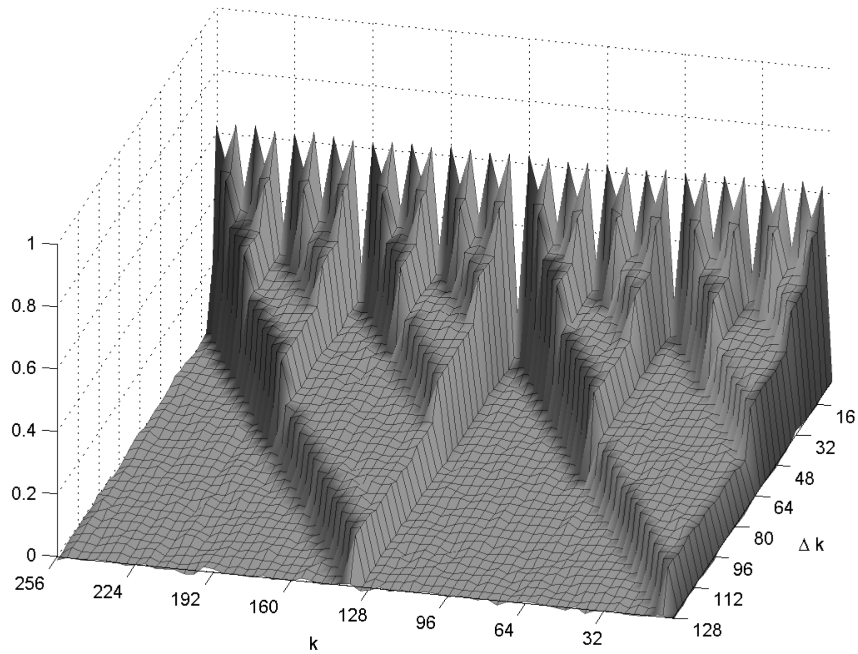

Figure 5: Correlations between nodes as a function of absolute position of the node (k) and inter-nodal distance (Δk) for the single-parent non-stationnary model (j=8, $\sigma_W = 1$). Blocky non-stationnary artifacts are obvious.

Figure (6) shows the same plot for the one-two parent model this time. As this model is closer to the stationnary model in terms of number of parent nodes involved in the graph, non-stationnarities do not appear as strickingly as for the single-parent model. Still the dyadic structure is obvious and the process is non-stationnary.

Figure (7) shows the same plot for the α-model. Elimination of the artefactual dyadic structure is obvious when compared to figure (5).

All simulations involved 20000 realizations of graphs.

9 Discussion and conclusions

A method has been proposed to ensure stationnarity of multi-scale graphical models processes. The method relies on the necessary independence of second-order statistics over spatial translation. In the case of a multivariate gaussian joint distribution on the set of nodes, the condition is also sufficient and thus leads to strict-sense stationnarity of the process rather than approximate stationnarity as was the case in [7].

The construction of the model leaded us to an expression for the spatial covariance on the nodes of the graph given the parameter values of the model. Hence we have constructed a whole family of processes for which a set of parameters (the α's) can be estimated.

We have also shown that it is possible to ensure stationnarity of higher order moments with additional conditions imposed on the model. The question whether a modification of the model could lead to strict sense stationnarity for any marginal distribution on the nodes is the subject of ongoing work. For instance it can be shown that a dyadic model starting with a number of nodes (on the first level) which is not a power of two can be made strictly stationnary regardless of the marginal distribution on the nodes; however the stationnarity conditions in this case involve a strong delocalization of the spatial covariance function.

Regarding the problem of fully developed turbulence cascades, the proposed method allows not only to model scaling behavior on the marginal distribution on the nodes, but also to model the form of the spatial covariance function. For instance a $\log^2(\Delta k)$ behavior for the shape of the spatial covariance has been observed with great universality over different Reynolds number and experimental setup [5]. The method proposed here can be used to explore cascading processes that would explain the $\log^2(\Delta k)$ behavior.

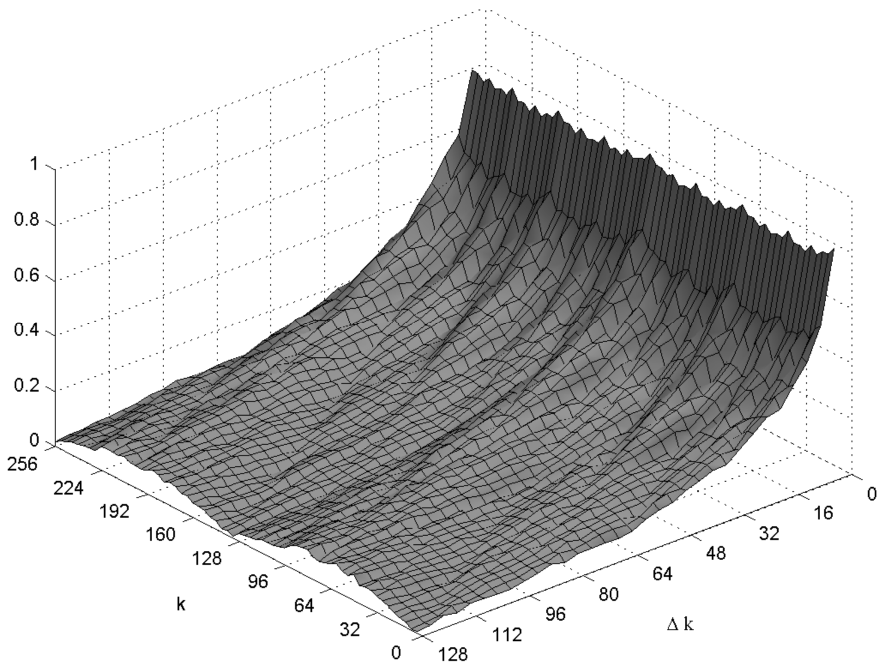

Figure 6: Correlations between nodes as a function of absolute position of the node (k) and inter-nodal distance (Δk) for the two-parent non-stationnary model (j=8, $\sigma_W = 1$). Artifacts are subtler than the ones on figure (5), yet still visible at any inter-nodal distance.

A Appendix

Proof of Proposition 1. Again we will proceed by induction. The p-order moments of the joint distribution of level j nodes can be expressed as:

$$\begin{gathered} C_1(c_{j,k},\ c_{j,k+\Delta k_1}, ...c_{j,k+\Delta k_{p-1}}) = \\ \sum_{\{K\}} \alpha^{(j)}_{2k_1-k}\ \alpha^{(j)}_{2(k_2+k_1)-(k+\Delta k_1)} \cdots \alpha^{(j)}_{2(k_p+k_1)-(k+\Delta k_{p-1})} \cdot \\ E[c_{j-1,k_1} \cdot c_{j-1,k_2+k_1} \cdot\ \dots\ \cdot c_{j-1,k_p+k_1}] \end{gathered} \tag{17}$$

where C_1 is defined as in (2), and $K = \{k_i | i = 1...p,\ k_i = 0...2^{j-1} - 1\}$. The induction hypothesis allows us to write the expectation in the rightmost part of (17) as $C^{(j-1,p)}(k_2,\ \dots, k_p)$. We want this expression to be equal for $k = 0$ and $k = 1$ regardless of the parity of the Δk_i. First consider the case where all Δk_i are even. We write $m_i = \Delta k_i/2$. For $k = 0$, we have

$$C_1 = \sum_{\{K\}} \eta^{(j)}_{k_1}\ \eta^{(j)}_{k_2+k_1-m_1} ...\eta^{(j)}_{k_p+k_1-m_{p-1}}\ C^{(j-1,p)}(k_2,\ \dots, k_p)$$

Let us compute the Fourier transform along the m_i:

$$\hat{C}_1 = \sum_{\{M\}} e^{-2\pi i \mathrm{l}\cdot\mathrm{m}/2^j} \cdot \sum_{\{K\}} \eta^{(j)}_{k_1}\ \eta^{(j)}_{k_2+k_1-m_1} ...\eta^{(j)}_{k_p+k_1-m_{p-1}}\ C^{(j-1,p)}(k_2,\ \dots, k_p)$$

Changing all m_i to $m_i + (k_i + k_1)$:

$$\begin{aligned} \hat{C}_1 = \sum_{\{M\}} \sum_{\{K\}} & e^{-2\pi i \mathrm{l}\cdot\mathrm{m}/2^j} \cdot e^{-2\pi i l_1 (k_1+k_2)/2^j} \cdot\ \dots\ e^{-2\pi i l_{p-1}(k_1+k_p)/2^j} \cdot \\ & \eta^{(j)}_{k_1}\ \eta^{(j)}_{-m_1} ...\eta^{(j)}_{-m_{p-1}}\ C^{(j-1,p)}(k_2,\ \dots, k_p) \end{aligned}$$

leading to

$$\hat{C}_1 = \hat{\eta}^{(j)}[\sum_{i=1}^{p-1} l_i]\ \cdot\ \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i]\ \cdot\ \hat{C}^{(j-1,p)}[l_1,\ \dots, l_{p-1}]$$

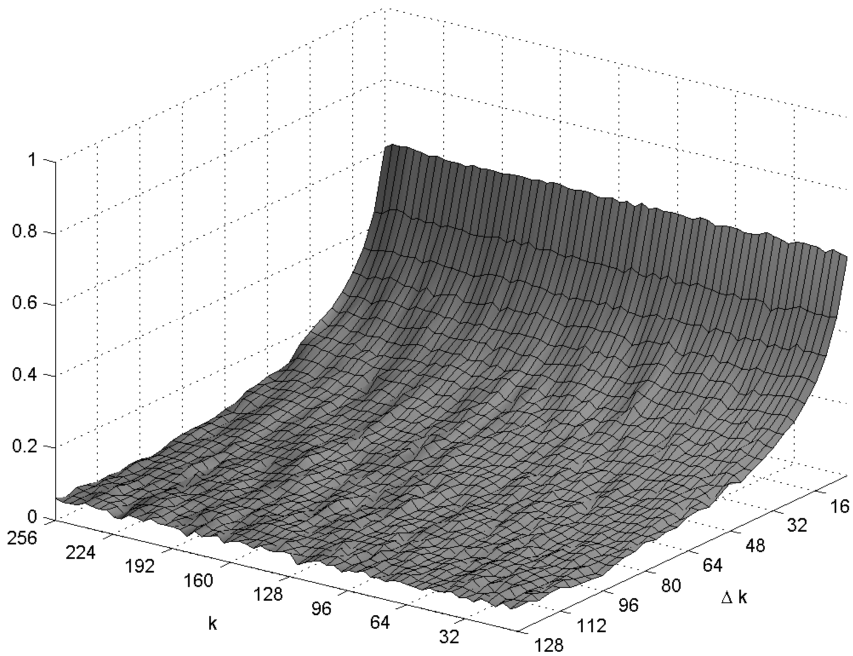

Figure 7: Correlations for the stationnary α-model as a function of absolute position of the node (k) and inter-nodal distance (Δk). (j=8, $\sigma_W = 1$)

For $k = 1$, the exact same expression as above holds where all η are replaced with ξ. The second-order condition imposed $\hat{\xi}[l] = \hat{\eta}[l] \cdot e^{i\theta[l]}$. Hence for $k = 1$, $\hat{C}_1$ becomes

$$\hat{C}_1 = \hat{\eta}^{(j)} \left[\sum_i^{p-1} l_i \right] \cdot e^{i\theta[\sum_i l_i]} \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] \cdot e^{-i\theta[l_i]} \cdot \hat{C}^{(j-1,p)}[l_1, \ \dots \ , l_{p-1}] \tag{18}$$

Even though $\theta[l] = -\pi l/2^j$ is a linear function of l over the interval $[-2^{j-2}+1, 2^{j-2}-1]$, the sum over i of the l_i can go out of this interval for $p \geq 3$. Thus (18) imposes a new condition on $\hat{\eta}$ and $\hat{\xi}$. This condition is expressed as

$$\hat{\eta}^{(j)} \left[\sum_i^{p-1} l_i \right] \cdot \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] = \hat{\eta}^{(j)} \left[\sum_i^{p-1} l_i \right] \cdot e^{i\theta[\sum_i l_i]} \prod_{i=1}^{p-1} \hat{\eta}^{(j)}[-l_i] \cdot e^{-i\theta[l_i]}, \tag{19}$$

$\forall l_i \in [0, 2^{j-1}-1]$, $i = 1 \ \dots \ p-1$.

Let us now consider the case where one of the Δk_i is odd while the others remain even. Assume without loss of generality that Δk_1 is the odd one. For $k = 0$, the effect is simply to change the $\hat{\eta}[-l_1]$ into $\hat{\xi}[-l_1]$, such that in this case

$$\hat{C}_1 = \hat{C}_{1,\text{all even}} \cdot e^{i\theta[l_1]} \tag{20}$$

Conversely for $k = 1$, we have $\hat{\xi}_-[l_1]$ changing into $\hat{\eta}_-[l_1] \cdot e^{-2\pi i l_1/2^j}$; the "missing" $\hat{\xi}_-[l_1]$ implies the absence of the $e^{i\theta[l_1]}$ term in (18). $\hat{C}_1$ becomes

$$\hat{C}_1 = \hat{C}_{1,\text{all even}} \cdot e^{-2\pi i l_1/2^j} \cdot e^{-i\theta[l_1]} \tag{21}$$

Replacing $\theta[l] = -\pi l/2^j$ in this last expression proves that both (20) and (21) are the same, provided the condition (19) for the case where all Δk_i were even is respected. Hence we find no new condition on α when considering one Δk_i to be odd.

The same exercise can be repeated for all combinations of even and odd Δk_i. Hence condition (19), along with (5, 6) is sufficient to provide stationnarity of order p moments of the joint distribution of the nodes.

Acknowledgements

The authors would like to thank Alain Arnéodo, J.F. Muzy, Jean Delour and Noel Cressie for insightful discussions. This work is supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada, by the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche and by the Network for Computing and Mathematical Modelling (RCM_2).

References

[1] A. Arneodo, E. Bacry and J.-F. Muzy (1995). The thermodynamics of fractals revisited with wavelets. *Physica A* **213,** 232–275.

[2] A. Arnéodo, E. Bacry and J.F. Muzy (1998). Random cascades on wavelet dyadic trees. *J. Math. Phys.* **39,** 4142–4164.

[3] A. Arnéodo, S. Manneville and J.F. Muzy (1998). Towards log-normal statistics in high Reynolds number turbulence. *Eur. Phys. J. B***1,** 129–140.

[4] A. Arnéodo, J.-F. Muzy and D. Sornette (1998). "Direct" causal cascade in the stock market. *Eur. Phys. J. B* **2,** 277–282.

[5] J. Delour, J.-F. Muzy and A. Arnéodo (2001). Intermittency of 1D velocity spatial profiles in turbulence: a magnitude cumulant analysis. *Eur. Phys. J. B* **23,** 243–248.

[6] U. Frish (1995). *Turbulence*. Cambridge University Press.

[7] H.-C. Huang and N. Cressie (2001). Multiscale graphical modeling in space: Applications to command and control. In *Spatial Statistics: Methodological Aspects and Applications*, ed. M. Moore. Springer Lecture Notes in Statistics **159,** Springer-Verlag, New York, 83–113.

[8] J.-F. Muzy (1995). Analyse de distributions fractales à partir de leur transformée en ondelettes. *Annales de physique* **20,** 63–231.

[9] J.F. Muzy, J. Delour and E. Bacry (2000). Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *Eur. Phys. J.* manuscript.