Mathematical Geology, Vol. 34, No. 1, January 2002 (© 2002)

On the Use of Multivariate Lévy-Stable Random Field Models for Geological Heterogeneity[1]

James Gunning[2]

Increasing attention has been paid to the use of non-Gaussian distributions as models of heterogeneity in sedimentary formations in recent years. In particular, the Lévy-stable distribution has been shown to be a useful model of the distribution of the increments of data measured in well logs. Frequently, the width of this distribution follows a power–law type scaling with increment lag, thus suggesting a nonstationary, fractal, multivariate Lévy distribution as a useful random field model. However, in this paper we show that it is very difficult to formulate a multivariate Lévy distribution with any nontrivial spatial correlations that can be sampled from rigorously in large models. Conventional sequential simulation techniques require two properties to hold of a multivariate distribution in order to work: (1) the marginal distributions must be of relatively simple form, and (2) in the uncorrelated limit, the multivariate distribution must factor into a product of independent distributions. At least one of these properties will break down in a multivariate Lévy distribution, depending on how it is formulated. This makes a rigorous derivation of a sequential simulation algorithm impossible. Nonetheless, many of the original observations that spurred the original interest in multivariate Lévy distributions can be reproduced with a conventional normal scoring procedure. Secondly, an approximate formulation of a sequential simulation algorithm can adequately reproduce the Lévy distributions of increments and fractal scaling frequently seen in real data.

KEY WORDS: heterogeneity, multivariate, Lévy, random field, sequential simulation, increments.

INTRODUCTION

In recent years, a series of papers (Gaynor and others, 1998; Gunning and Paterson, 1999; Liu and Molz, 1997; Painter, 1995, 1996a,b, 1998; Painter and Paterson, 1994; Painter, Paterson, and Boult, 1997) devoted to the use of Lévy distributions as models of heterogeneity in the earth sciences has appeared in *Mathematical Geology* and related journals, with a particular emphasis on petroleum applications. These papers range from studies with an emphasis on modelling the distributional characteristics of geological data or their increments (Liu and Molz, 1997; Painter, 1995, 1996a; Painter and Paterson, 1994), to those with an emphasis on simulation

[1]Received 7 April 2000; accepted 10 January 2001.

[2]Australian Petroleum Cooperative Research Centre, CSIRO Division of Petroleum Resources, Glen Waverley, VIC 3150, Australia; e-mail: james.gunning@dpr.csiro.au

and interpolation techniques (Painter, 1996b, 1998; Painter, Paterson, and Boult, 1997), applications (Gaynor and others, 1998; Painter, 1996b; Painter, Paterson, and Boult, 1997), and seismic conditioning (Gunning and Paterson, 1999). The prime motivation for this work is the observation that Lévy distributions model the distributions of *increments* of geophysical measurements rather well, even when these increments are taken over a wide range of length scales. What is usually observed is that the increments form symmetric, heavy-tailed distributions whose general shape is either quite constant over a range of lags, or increasingly "compact" for larger lags. In the extreme tails of the distribution, however, the Lévy model usually has too much weight, and additional bounding of the variable may need to be imposed (Painter, 1998).

These observations are rarely true of the raw (i.e., not differenced) data itself, which usually shows some trend over the scale of the study, thereby making a stationarity hypothesis about the univariate distribution rather suspicious. If the raw data where always susceptible of unambiguous trend-removal techniques and normal scoring of the residuals, the observation of Lévy-distributed increments would be of no great significance for practical modelling work. The problem is that many data sets cannot be directly modelled by stationary processes in a sensible way, and trend removal techniques applied to sparse data are both subjective and can lead to a narrow (nonconservative) estimate of the residual uncertainty.

If we wish to avoid the subjectivity associated with trend removal, suitable multivariate distributions must be constructed, which feature nonstationary behavior, and distributions of increments that have heavy tails and possible long-range correlations. These would provide a very "broad" underlying model for uncertainty studies. In particular, much of the initial enthusiasm for using "wide" distributions was the fact that such distributions give a high probability of large jumps in variables, thereby mimicking the sudden discontinuities that occur at facies boundaries. The prospect of being able to circumvent indicator or Boolean modelling was thus entertained, which would eliminate much of the arduous work in model building. This is appealing in cases where heavy diagenesis of rock properties has occurred, thereby masking the facies boundaries. In such environments, creating an appropriate facies classification can be very subjective. Further, if the control that facies types exert on properties like permeability is substantially weakened by the diagenesis, then there would be little reason to model and simulate facies architecture. Of course, the question of whether the diagenetic processes consistently overwhelm the facies control on permeability is open to debate. Also, in terms of effort and expenditure, it is prudent to consider the fraction of practical cases where such diagenesis masks the facies control, which may not be very high. If so, it is difficult to avoid the conclusion that facies modelling is virtually always necessary.

Attempts to formulate such a "general" model in terms of multivariate Lévy distributions have been the subject of several papers (Painter, 1996b, 1998; Painter

and Paterson, 1994). However, we wish to demonstrate that it is extremely difficult to formulate such a distribution if one desires to preserve a number of features of multi-Gaussian densities, which are essential for any practical random field model. We are not primarily concerned here with the issue of whether the distributions and correlations can be constructed so as to fit the observations, as the more fundamental issue of whether there exists a coherent mathematical formulation of this type of model that is practicable to sample from. It is the purpose of this paper to point out the extent of the difficulties, and suggest some possible remedies.

The paper is organised as follows. In Desirable Features . . . section we describe a set of important properties of practical continuous multivariate densities. Multivariate Lévy Distributions section addresses the question of whether multivariate Lévy random fields can exhibit these properties, and several difficulties are revealed. We devote significant attention to the simpler problem of stationary multivariate Lévy processes, despite the introductory remarks that suggest the nonstationary problem is more relevant. The difficulties are serious enough to warrant close examination of the simplest problem (i.e., stationary processes), before proceeding to the more difficult nonstationary case. Some practical, if approximate, remedies for these problems are outlined in Suggestions section, which may well be sufficient for most practical applications. Our findings are summarised in Conclusions section.

DESIRABLE FEATURES OF A CONTINUOUS MULTIVARIATE DENSITY FOR EARTH SCIENCES APPLICATION

A number of general principles have come to be accepted as essential in the construction of practical multivariate distributions of continuous, spatially distributed properties. Of the points listed below, only the first is truly fundamental. Points 2–4 are so convenient as to be indispensible in practical work.

1. *Permutation invariance.* There is nothing special about the order in which we choose to label the components in a random field model of some geological property. The full multivariate distribution must therefore be independent of this order.
2. *The uncorrelated limit as a sequence of independent deviates.* Any useful multivariate distribution ought to have a limit where, if the spatial dependence of distinct components in the random field disappears, the multivariate distribution turns into a product of independent univariate distributions. Such a property is vital for the success of methods such as sequential simulation techniques, which are *de rigueur* in large studies. These methods rely on the probabilistic independence of screened components, which enables their probability to factor out when forming the appropriate conditional distribution.

If the factorization-into-independent distributions limit does not hold, this implies that there is something absolute about the number n of variables (grid-blocks) used in the multivariate model. Since n is arbitrary, this would be a troubling conclusion.

3. *The full multivariate distribution can be constructed from a limited number of moments*. This is a special case of the universally agreed principle of *parsimony*, which applies to both continuous and Boolean/marked-point models. One would like to be able to construct the full distribution from a limited number of free parameters that we can infer from the data. In general, a multivariate distribution can only be fully described by an infinite number of moments. In practice, it is difficult to determine anything more than the univariate and bivariate, or surrogates for these, parametrised in a suitably parsimonious way.
4. *All marginal distributions must preferably have the same functional form as the full distribution, and can be constructed directly in the same manner.* This is a very stringent requirement, motivated by the following considerations. All earth science random field models are marginal distributions, in the sense that there are always regions of space that are spatially correlated with the region we choose to model, but do not affect the global phenomena we wish to study (oil production, producible mineral reserves, etc.), and hence are usually neglected. Practitioners are accustomed to omitting such regions from the model, usually on the pretext that they do not directly affect the global phenomena of interest. A more rigorous justification is that multi-Gaussian models permit integration over the parameters in the "uninteresting" region, leaving a marginal distribution insensitive to its inclusion. This very special property of multi-Gaussian distributions need not hold automatically for other multivariate densities of interest.

It may be that these principles have come to seem inviolable through the habitual use of multi-Gaussian fields, particularly the last. Yet it is clear that many practitioners are uneasy about the use of multi-Gaussian fields, particularly in applications where their urbane, maximum entropy character will yield nonconservative estimates of the probability of extreme events. It is thus worthwhile to examine whether other, heavier-tailed, continuous multivariate distributions can be used, and, if so, which of the desirable properties listed above may have to be sacrificed.

MULTIVARIATE LÉVY DISTRIBUTIONS

Of possible candidates for an alternative to the multi-Gaussian density, the multivariate Lévy distributions are the most appealing. They are, in a sense, the

nearest generalization of the Normal distribution, and their *stable* character makes them appealing as models of physical processes that are additive in nature.

Press (1972) has given a general definition of a discrete, symmetric Lévy-stable multivariate random field $\mathbf{Z} = \{z_j\}_{j=1}^{j=n}$ via its characteristic function (the Fourier transform of the probability density):

$$\phi_{\mathbf{Z}}(\mathbf{t}) = \exp\left(i\mathbf{a}^{\mathrm{T}}\mathbf{t} - \frac{1}{2}\sum_{k=1}^{m}(\mathbf{t}^{\mathrm{T}}\boldsymbol{\Omega}_k\mathbf{t})^{\alpha/2} \right). \tag{1}$$

Here $\mathbf{t}$ is the Fourier transform variable, $\mathbf{a}$ is a vector related to the mean of $\mathbf{Z}$, and the $\{\boldsymbol{\Omega}_k\}_{k=1}^{m}$ is a set of matrices, no two of which are proportional. Such a characteristic function defines a *stable distribution of order m*, with "width" parameter α. It reduces to a multivariate normal distribution when $\alpha = 2$, and the practical range of interest is $1 < \alpha \leq 2$. The distribution becomes less "compact" as α decreases away from 2. For simplicity, and without loss of generality, we consider only zero-mean, symmetric distributions, so the factor $i\mathbf{a}^{\mathrm{T}}\mathbf{t}$ will be dropped. Press discusses also a more general class of nonsymmetric Lévy distributions, but we leave these as outside the scope of the present study. A wider generalization is presented by Jurek and Mason (1993), which is also outside our scope.

Order-1 Distributions and Their Properties

Existing work (Kumar and Foufoula-Georgiou, 1993; Painter, 1996a, 1998) on multivariate Lévy-stable random field models has used only the order-1 characteristic function

$$\phi_{\mathbf{Z}}(\mathbf{t}) = \exp\left(-\frac{1}{2}(\mathbf{t}^{\mathrm{T}}C\mathbf{t})^{\alpha/2} \right). \tag{2}$$

We include the factor of $1/2$ here out of deference to the Gaussian distribution, for then C becomes exactly the Gaussian covariance matrix when $\alpha = 2$. The random variables $\mathbf{y}$ thus have the distribution

$$P(\mathbf{y}) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{t}^{\mathrm{T}}C\mathbf{t})^{\alpha/2}}\, e^{i\mathbf{t}^{\mathrm{T}}\mathbf{y}}\, d\mathbf{t}. \tag{3}$$

1. *The generalized covariance.* For Lévy random fields, second moments cannot be found, so a "generalized covariance" defined in terms of alternative measures of the width of the distribution is needed. For any Levy-distributed quantity x, the pseudovariance factor C in the single variable case of Equation (3) is a good substitute for the variance. Fama and Roll

(1968, 1971) write the distribution as

$$P(x) \sim \int_{-\infty}^{\infty} e^{-(C_{\mathrm{FR}}|t|)^{\alpha}}\, e^{itx}\, dt \tag{4}$$

and provide estimators of α and the *scale factor* C_{FR} (see Appendix A). The pseudovariance factor C we use is related to C_{FR} by $C = 2^{2/\alpha} C_{\mathrm{FR}}^2$.

From Equation (2), the marginal distribution of any y_j is

$$P(y_j) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}(C_{jj}t^2)^{\alpha/2}}\, e^{ity_j}\, dt, \tag{5}$$

and that of the increment $\Delta y_{ij} \equiv y_i - y_j$ is

$$P(\Delta y_{ij}) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}((C_{ii}+C_{jj}-2C_{ij})t^2)^{\alpha/2}}\, e^{it\Delta y_{ij}}\, dt. \tag{6}$$

For random fields with finite second moments we have the usual result

$$\langle (y_i - y_j)^2 \rangle = \langle y_i^2 \rangle + \langle y_i^2 \rangle - 2\langle y_i y_j \rangle, \tag{7}$$

so it is reasonable to define the generalized covariance W for Lévy fields by the matrix coefficients that appear in the marginal distributions for the increments, that is, we define W_{ij} by the pair of equations

$$P(\Delta y_{ij}) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}(W_{ii}+W_{jj}-2W_{ij})|t|^{\alpha}}\, e^{it\Delta y_{ij}}\, dt \tag{8}$$

$$P(y_j) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}W_{jj}|t|^{\alpha}}\, e^{ity_j}\, dt. \tag{9}$$

In practice, it is the estimated scale factors $C_{\mathrm{FR}}(y_i)$ and $C_{\mathrm{FR}}(\Delta y_{ij})$ of the raw variables y_i and their increments Δy_{ij} (for points separated by a distance Δr_{ij}), which will we modelled, assuming these are adequately described by Lévy distributions with the same α. The generalized covariance W can be derived from these scale factors by the set of relations

$$C_{\mathrm{FR}}(y_i) = \left(\frac{1}{2} W_{ii}\right)^{1/\alpha} \tag{10}$$

$$C_{\mathrm{FR}}(\Delta y_{ij}) = \left(\frac{1}{2} W_{ii} + W_{jj} - 2W_{ij}\right)^{1/\alpha}. \tag{11}$$

However, under the order-1 generalization, the matrix C can be constructed directly from the matrix of Fama–Roll scale factors, without passing through the generalized covariance, viz.,

$$C_{\mathrm{FR}}(y_i)^2 = 2^{-2/\alpha} C_{ii} \tag{12}$$

$$C_{\mathrm{FR}}(\Delta y_{ij})^2 = 2^{-2/\alpha}(C_{ii} + C_{jj} - 2C_{ij}). \tag{13}$$

Typically, we might model the increment scale factors $C_{\mathrm{FR}}(\Delta y_{ij})^2$ by a variogram-like function, for example, $C_{\mathrm{FR}}(\Delta y_{ij})^2 = v(1 - \exp(-3\Delta \mathbf{r}_{ij}/a)$, and fit the parameters v, a. This would give matrix elements $C_{ij} = 2^{2/\alpha-2} v \ \exp(-3|\Delta \mathbf{r}_{ij}|/a)$, which fully defines the model (3). Notice that in the Gaussian case ($\alpha = 2$), the increment scale-factors estimator $C_{\mathrm{FR}}(\Delta y_{ij})^2$ is loosely equivalent to the traditional semivariogram. For general α, we called it a "Fama–Roll variogram."

In summary the full multivariate model (2) can be written down "by inspection," given either W, or, equivalently, the set of scale factors $C_{\mathrm{FR}}(y_i)$, $C_{\mathrm{FR}}(\Delta y_{ij})$. Thus we have a definite and simple algorithm for constructing the full multivariate density from the generalized bivariate moments.

2. *Marginal distributions*. If the vector $\mathbf{y}$ has distribution (3), and is partitioned as $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$, with the corresponding partitioning

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \tag{14}$$

of C, then the marginal distribution of $\mathbf{y}_1$ is easily shown to be

$$p(\mathbf{y}_1) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{t}_1^{\mathrm{T}} C_{11} \mathbf{t}_1)^{\alpha/2}} \ e^{i \mathbf{t}_1^{\mathrm{T}} \cdot \mathbf{y}_1} \ d\mathbf{t}_1, \tag{15}$$

which is obviously independent of the nodes $\mathbf{y}_2$ that contribute to the matrix elements C_{12}, C_{21}, C_{22}.

3. *Invariance under permutations of the* y_i. If the labelling order of the random field vector $\mathbf{y}$ is permuted, as in $\mathbf{y} \to \tilde{\mathbf{y}} = P\mathbf{y}$, then the matrix C will transform as $C \to \tilde{C} = PCP^{\mathrm{T}}$. It is then easy to show that

$$P(\tilde{\mathbf{y}}) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{t}^{\mathrm{T}} \tilde{C} \mathbf{t})^{\alpha/2}} \ e^{i \mathbf{t}^{\mathrm{T}} \cdot \tilde{\mathbf{y}}} \ d\mathbf{t}, \tag{16}$$

that is, the definition is invariant under permutations.

4. *Uncorrelated limit.* In Equation (2), if we take $C = I$, the distribution is

$$P(\mathbf{y}) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{t}^{\mathrm{T}}\cdot\mathbf{t})^{\alpha/2}}\, e^{i\mathbf{t}^{\mathrm{T}}\cdot\mathbf{y}}\, d\mathbf{t} \tag{17}$$

which can only be factored into n independent Lévy distributions if $\alpha = 2$ (the Gaussian case). Firstly, this result is problematic for parameter estimation, since a typical data set will be treated as a set of *i.i.d.* Lévy deviates, with the multivariate density

$$P(\mathbf{y}) \sim \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{j=1}^{n} e^{-\frac{1}{2}|t_j|^{\alpha}} \right) e^{i\mathbf{t}^{\mathrm{T}}\mathbf{y}}\, d\mathbf{t}, \tag{18}$$

for which, for example, certain estimators of α have been constructed by Fama and Roll (1971). With this definition of Lévy white noise, it is clear that spatial and ensemble averages can be interchanged freely, but this is not so with Equation (17). It is not clear that Fama and Roll's estimators are appropriate for the distribution (17), or, indeed, whether any classical estimation technique that invokes the interchange of spatial and ensemble averages has any meaning for this distribution.

The second problem is that the absence of a sensible white noise limit precludes the possibility of a theoretically sound sequential simulation algorithm. Techniques for sampling the correlated distribution (2), conditioned to some fixed y_i, invariably make use of linear transforms, which turn the problem into one of sampling from the multivariate white noise form. If the white noise form is not of "product" form, then it cannot be sampled from sensibly using a sequential algorithm. Recall that any sequential simulation algorithm relies on the factorization of the total density into the Bayesian product

$$\begin{aligned} p(y_n \ldots y_1) = {} & p(y_n \mid y_{n-1} \ldots y_1) p(y_{n-1} \mid y_{n-2} \ldots y_1) \ldots \\ & \times p(y_j \mid y_{j-1} \ldots y_1) \ldots p(y_1). \end{aligned} \tag{19}$$

First, at any stage j, the density $p(y_j \mid y_{j-1} \ldots y_1)$ will be constructed from the marginal density $p(y_j \ldots y_1)$ obtained by integrating over $y_{j+1} \ldots y_n$, so this marginal density must have a tractable form. Second, we will approximate $p(y_j \mid y_{j-1} \ldots y_1)$ as $p(y_j \mid y_{k_1} \ldots y_{k_m})$, where we use some neighborhood $y_{k_1} \ldots y_{k_m}$, $m \leq j-1$, which screens the remaining fixed nodes $y_{k_{m+1}} \ldots y_{k_{j-1}}$, such that the distribution $p(y_j \mid y_{j-1} \ldots y_1)$ is virtually independent of $y_{k_{m+1}} \ldots y_{k_{j-1}}$. Then, by the very definition of independence, the terms containing $y_{k_{m+1}} \ldots y_{k_{j-1}}$ should *factor out* as a constant.

But the screened nodes $y_{k_{m+1}} \ldots y_{k_{j-1}}$ will not contribute a factorizable term in a form like (17), they have the effect only of increasing n. The consequence of this is that all other fixed nodes in the system must be included at any stage of the sequential decomposition. There is no legitimate notion of an approximate conditional distribution formed from only a local neighborhood.

A practical consequence of this is that, although the marginal distribution (15) indicates that all nodes in the system have a univariate Lévy distribution with parameter α, a sequential simulation using the local-neighborhood algorithm described in Painter (1998) will yield a field whose estimated α is usually significantly higher than that used in the simulation (except in the Gaussian case) (see Appendix B). Experiments with this algorithm have shown that it has great difficulty generating processes whose increments have Lévy width factors $\alpha < 1.5$, irrespective of the spatial correlations. Figure 1 illustrates the difference between the theoretical α and the observed α when the sequential algorithm is used to produce white noise using this formulation. Note that these scale factors have been estimated using a single large realizations of the process, unlike the ensemble averages used by Painter (1998). The property of ergodicity (interchangeability of spatial and ensemble averages) is crucial to the

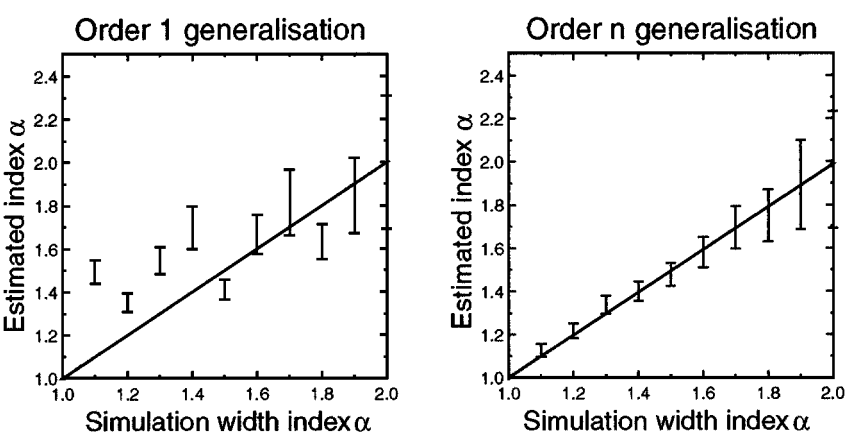

Figure 1. Left: estimated α parameters versus "theoretical" α, from a long (10^5 units) 1D "Lévy white noise" process generated using a sequential algorithm with a maximum of 20 neighbors, and Equation (17) for the local marginal distribution, rendered conditional to the presimulated points by the method presented in Painter (1998) (summarized in Appendix B). Right: the same graph using the order-n multivariate generalization, sampled using a sequential algorithm. The increasing uncertainty in the estimators of α is due to certain properties of the quantile-based estimators of Fama and Roll (1971).

practice of geostatistics, so we use a spatial average to emphasise this point (in practice, parameters are only ever estimated from one "realization").

Alternative, Higher Order Generalizations

Since the order-1 multivariate generalization fails most conspicuously in the uncorrelated limit, we construct an alternative generalization starting directly from this problem. Notice first that the uncorrelated Lévy density (18) is already an order-n multivariate Lévy distribution in the sense of Press' definition, with the $k = 1 \ldots n$ matrices

$$(\mathbf{\Omega}_k)_{ij} = \begin{cases} 1 & i = j = k \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

And since, in the Gaussian case, a correlated multivariate distribution could be written as

$$P(\mathbf{y}) \sim \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{j=1}^{n} e^{-\frac{1}{2}|t_j|^2} \right) e^{i\mathbf{t}^{\mathrm{T}} L^{-1} \mathbf{y}} \, d\mathbf{t}, \tag{21}$$

where L is any convenient "square root" of C ($LL^{\mathrm{T}} = C$), a natural extension of (18) to the correlated case is

$$P(\mathbf{y}) \sim \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{j=1}^{n} e^{-\frac{1}{2}|t_j|^{\alpha}} \right) e^{i\mathbf{t}^{\mathrm{T}} L^{-1} \mathbf{y}} \, d\mathbf{t}, \tag{22}$$

for some suitable L. This is still an order-n multivariate Lévy distribution, as the transformation $\mathbf{t} \to L^{\mathrm{T}}\mathbf{t}'$ and a few lines of algebra will show.

1. *Generalized covariance*. The marginal distributions of any y_i or increment Δy_{ij} are still Lévy distributions under this generalization, as given by Equations (8) and (9), but the generalized covariance takes on the less tractable form

$$W_{ij} = \frac{1}{2} \left(\sum_{k=1}^{n} |L_{ik}|^{\alpha} + \sum_{k=1}^{n} |L_{jk}|^{\alpha} - \sum_{k=1}^{n} |L_{ik} - L_{jk}|^{\alpha} \right). \tag{23}$$

In the Gaussian limit $\alpha = 2$, this equation degenerates to $W = LL^{\mathrm{T}}$ (as expected), but also provides $n(n+1)/2$ equations for the n^2 elements of L, since W is symmetric. The resulting incompleteness of definition is

of no consequence, since only W is observable. *Any* matrix that satisfies $LL^{\mathrm{T}} = W$ is satisfactory; the triangular Cholesky decomposition, the eigenvalue decompositions $L = U\Omega^{1/2}$ or $L = U\Omega^{1/2}U^{\mathrm{T}}$ (where $W = U\Omega U^{\mathrm{T}}$) being three simple examples.

In the Lévy case, however, the higher order moments are explicitly functions of L, not simply LL^{T}, so the particular choice of L matters. Its full definition is then completed by the requirement of permutation invariance.

2. *Permutation invariance*. The simplest way to guarantee suitable permutation behavior is to insist that L be symmetric, so $L = L^{\mathrm{T}}$. This yields a decomposition analogous to the eigenvalue decomposition $L = U\Omega^{1/2}U^{\mathrm{T}}$. To demonstrate that this works, we can rewrite (23) as

$$\begin{aligned} W_{ii} + W_{jj} - 2W_{ij} = &\left| \left(W_{ii} - \sum_{k\neq i} |L_{ik}|^{\alpha} \right)^{1/\alpha} - L_{ji} \right|^{\alpha} \\ &+ \left| \left(W_{jj} - \sum_{k\neq j} |L_{jk}|^{\alpha} \right)^{1/\alpha} - L_{ij} \right|^{\alpha} \\ &+ \sum_{k\neq i,j} |L_{ik} - L_{jk}|^{\alpha} \end{aligned} \tag{24}$$

This set of $n(n-1)$ equations for the off-diagonal elements whose solution L_{ij} will be invariant under the interchange of any two labels i, j if L is symmetric. Similarly, the diagonal elements

$$L_{ii} = \left(W_{ii} - \sum_{k\neq i} |L_{ik}|^{\alpha} \right)^{1/\alpha} \tag{25}$$

will transform correctly. At this stage, no elementary sequential algorithm has been found to solve this system, but it can be solved numerically by standard nonlinear root-finding algorithms for small n. For arbitrary n, in systems that are near-uncorrelated ($|W_{ij}| \ll W_{ii}$), the approximate solution

$$L_{ii} = W_{ii}^{1/\alpha} \tag{26}$$

$$L_{ij} = \frac{2W_{ij}}{\alpha\left(W_{ii}^{1-1/\alpha} + W_{jj}^{1-1/\alpha}\right)} \tag{27}$$

can be derived, which exhibits the required permutation symmetry nicely.

The two serious problems here are then the nonlinearity of (24), and, more disturbingly, the fact that a positive-definite W no longer guarantees a solution for L.

3. *Marginal distributions.* The marginal distributions are messy. Integrating over the n_2 degrees of freedom in $\mathbf{y}_2$ to form the marginal of $\mathbf{y}_1$ gives a long expression which can be shown to be Lévy-stable only if $\alpha = 2$. This cannot reduce to the distribution $P(\mathbf{y}_1)$ one would construct directly using only the generalized covariance submatrix W_{11}. It can only be made to work if we choose L to be lower triangular (which breaks permutation invariance), or in the trivial case that the nodes $\mathbf{y}_2$ and $\mathbf{y}_1$ are independent.
4. *White noise limit.* By construction, the distribution (22) factors into n independent Lévy distributions if L is diagonal, and then $W_{ij} = |L_{ii}|^{\alpha}$.

SUGGESTIONS

It is apparent from the preceding discussion that either possible generalization to multivariate Lévy distributions has serious difficulties. If we choose the order-1 generalization, we get attractive marginal distributions, permutation invariance, and a simple set of equations for the generalized covariance, but a complete breakdown of the white noise limit, and hence the possibility of a well-behaved sequential simulation algorithm. If we choose the order-n generalization, we get a sensible white noise limit and recover the generalized covariance and permutation invariance with some pain, but then completely lose the simplicity of the marginal distributions, which make sequential algorithms possible.

In view of this, it is worth asking, firstly, whether any of the empirical observations that motivated the earlier enthusiasm for such generalizations cannot be adequately explained by more conventional approaches, and secondly, whether some series of approximations can be made to produce an algorithm that is adequate for practical use. In particular, we would desire an approximate algorithm to produce the correct width parameters α for the increments, and produce the correct spatial correlations over the entire range of white noise to power–law, as desired. Apropos the first question, for the case of stationary processes, it can be shown that a normal scoring technique is capable of producing processes whose increments are Lévy distributed with a width parameter α that is remarkably constant over different length scales. This possibility is discussed in the **Stationary Processes** . . . subsection following. As for the second possibility, for the case of nonstationary fractal processes, we demonstrate that an ad hoc technique using the order n generalization (22) and a crude Cholesky decomposition in place of a true solution to (23) is remarkably successful. Details are given in the **Fractal Processes** . . . subsection.

Stationary Processes: Heavy-Tailed Distributions of Increments From an Inverse Normal-Scoring Method

Firstly, consider the traditional idea of creating a process whose univariate distribution is Lévy-stable by inverse normal-scoring a multi-Gaussian process to the desired Lévy distribution. Figure 2 shows the estimated α factors for the increments of a 10^4 point, 1D exponentially correlated Gaussian process subjected to this transformation. The left figure shows the estimated α factors for a Lévy white noise process generated by transforming Gaussian white noise to a Lévy distribution with parameter $\alpha = 1.5$. The increments have width parameter $\alpha = 1.5$ within statistical sample error. The right figure shows the same estimated α's versus lag when the pre-transformed Gaussian process is now exponentially correlated, with correlation length of 50 units. The increments at shorter lags in this case are perhaps slightly more compact ($\alpha \approx 1.6$), but quickly settle back to 50 units.

A challenging example for the method is furnished by the covariance function (Wackernagel, 1995)

$$C(r) = C_0 \frac{2}{\Gamma(H)} \left(\frac{r}{2a}\right)^H K_H \left(\frac{r}{a}\right), \tag{28}$$

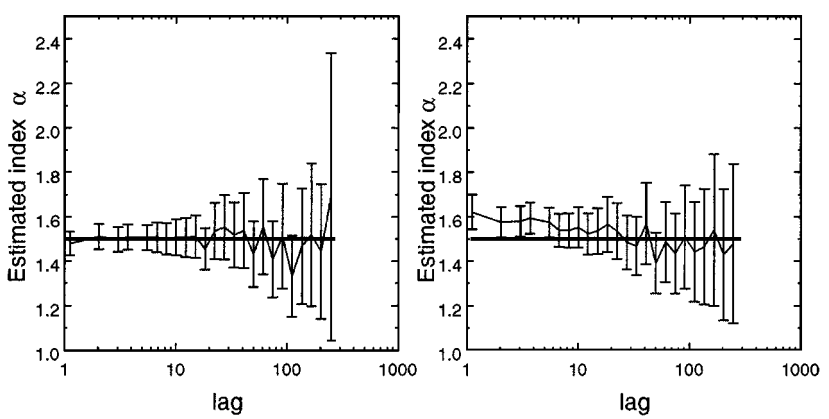

Figure 2. Estimated α parameters versus increment lag, from a long (10^4 units) 1D multi-Gaussian process back-transformed to a univariate Lévy distribution with $\alpha = 1.5$. The left inset shows the case when the original multi-Gaussian process is uncorrelated, the right when the correlation is exponential with range $a = 50$ units ($C(r) \equiv \exp(-3r/a)$). The error bars are $\approx$1 SD, so the increments have $\alpha = 1.5$ for all lags, in both cases, within experimental error.

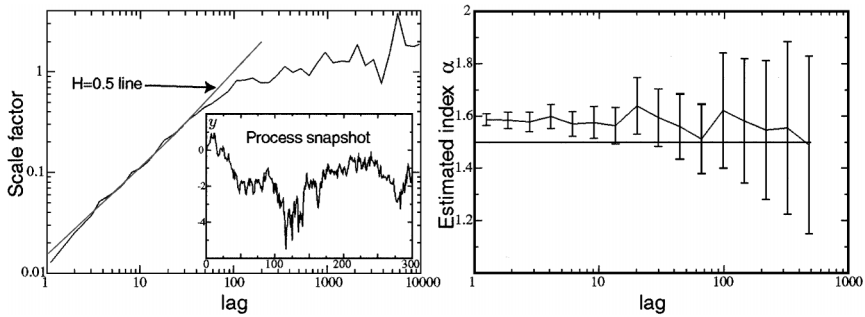

Figure 3. Pseudo-fLm for $\alpha = 1.5$, $H = 0.5$ generated using quasi-stationary covariance, Equation (28), with range parameter $a = 100$ and sill $C_0 = 1$. The inset on the left graph shows a snapshot of the process y. Since this is a stationary process with univariate distribution a Lévy distribution with $\alpha = 1.5$, scale factor $C = 1$, the stretching effect due to the back transformation gets noticeable for $|y| > 3$ (a kind of heteroscedasticity is induced). For a 1D realization of 10^5 points, the sample Fama–Roll scale factors are shown on the left graph, and the α (width) parameters of the increments is shown in the right graph. The fractal behavior is nicely evident on the left graph, which tapers off to virtual independence once the lags exceed the scale parameter $a = 100$. On the right, the Lévy width parameter is quite stable over a range of lags, with a slight upward bias as explained in the text.

where K_H is a modified Bessel function. This is a very interesting process, having the variogram asymptotic properties ($\gamma(r) = C(0) - C(r)$)

$$\gamma(r) \sim C_0 \frac{\Gamma(1-H)}{\Gamma(1+H)} \left(\frac{r}{a}\right)^{2H} \qquad r \ll a. \tag{29}$$

It is thus capable of simulating a fractional Brownian motion process by setting a much larger than the system size, and adjusting the ratio C_0/a^{2H} suitably. Figure 3 shows an example of what kind of behavior this process gives under the normal scoring method. A speculation as to why the increments at small lags have slightly too high an α value is that these lags are more likely to be in the region near the peak of the univariate Lévy distribution, so the inverse transformation from Gaussian to Lévy distribution does not "stretch" them very much.

This method is very good at reproducing α values that are at the volatile end of the practical range ($1 < \alpha < 1.5$). The method works by forcing the increments between uncorrelated locations to have the correct distribution (via the back transform), and this transform is sufficiently pervasive to induce a comparable distribution in the more correlated increments.

Fractal Processes

Nonstationary, fractal processes are simulated by Painter (1998) by a sequential method where a set of n known (fixed or presimulated) neighbors are

located close to the current unknown point of interest. One of the neighbors is chosen arbitrarily as a local origin, and the multivariate distribution of the increments with respect to this point is constructed using (2), with the fBm-like covariance

$$C_{ij} = 2^{2/\alpha-1}C_0^2(|r_i|^{2H} + |r_j|^{2H} - |r_i - r_j|^{2H}). \tag{30}$$

Because this method yields increments that tend to have too compact a distribution, as shown in Appendix B, we prefer to use the order-N generalization (22), somewhat sportily ignoring the fact that we cannot rigorously approximate the marginal for the local neighborhood of size $n \ll N$, as noted in the **Alternative, Higher Order Generalizations** subsection. Instead, at each step, we *assert* the form (22) directly, where n is the size of the local neighborhood. We construct L directly from the Cholesky decomposition of C, as a crude approximation to Equation (23), which we know happens to be strictly true only in the Gaussian case. Given the heavy approximations made, we chose the Cholesky decomposition in preference to the eigenvalue decomposition for reasons of speed. The permutation invariance is already broken by the approximation of Equation (23) by $LL^{\mathrm{T}} = W$. The Cholesky form allows the local conditional density for an unknown y_n in terms of fixed neighbors $y_1, \ldots y_{n-1}$ to be derived simply from (22) as

$$P(y_n) \sim \int_{-\infty}^{\infty} e^{-\frac{1}{2}|t_n|^{\alpha}}\, e^{it_n(\sum_{j=1}^{n-1} L_{nj}^{-1} y_j + L_{nn}^{-1} y_n)}, \tag{31}$$

which is simply a standard Lévy distribution for the "offset + scaled" variable $\sum_{j=1}^{n-1} L_{nj}^{-1} y_j + L_{nn}^{-1} y_n$. This makes the sampling much easier.

If the sequential simulation consisted of finding a local neighborhood with only one known neighbor at each step ($n = 1$), this neighbor would become the local origin, and the matrix C would be 1×1, with entry (computed from (30))

$$C_{11} = 2^{2/\alpha} C_0^2 |h|^{2H}, \tag{32}$$

where h is the distance from the unknown point to the local origin. The Cholesky decomposition would then give $L_{11} = 2^{1/\alpha} C_0 h^H$, and thus the true generalized covariance, computed from Equation (23), would be $W_{11} = 2C_0 h^{\alpha H}$. The scale factor is then computed from Equation (10) to be

$$C_{\mathrm{FR}}(h) = C_0 h^H. \tag{33}$$

As in Painter's work, modelling consists in fitting the scale factors to this power–law form to find the constants C_0 and the Hurst parameter H, which are then used

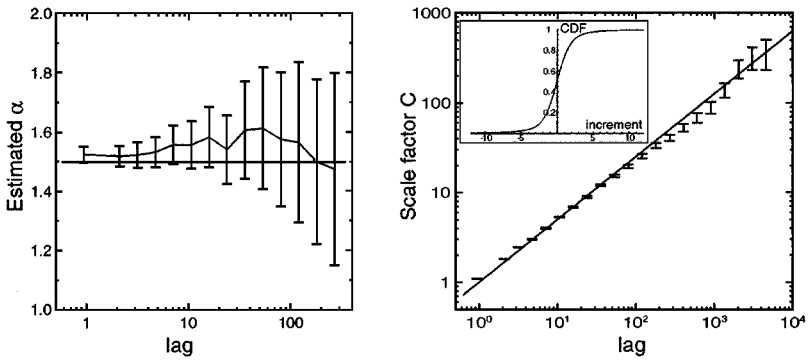

Figure 4. Lévy α factors (left) and scale factors (right) estimated from a single 10^5-long fLm process generated with $\alpha = 1.5$, $H = 0.7$, as described in Fractal Processes subsection. The theoretical α and C curves are shown solid. Inset into the right figure is the experimental CDF of the lag 1 increments plotted with the theoretical CDF. The curves are virtually indistinguishable.

at each step of the process to form the local pseudocovariance matrix C at each step of the sequential simulation via (30).

Of course, the preceding argument is not a proper derivation of a suitable estimation framework for the parameters C_0 and H, but the following points are salient.

1. We use the "preferentially farthest" random node selection method described by Painter. Each new node is chosen, from among a small number of purely random candidates, to be as far away as possible from known nodes. This multigrid-like scheme can be shown to largely guarantee the generation of the power–law behavior.
2. In practice the coefficient $\hat{C}_0$ estimated by Fama–Roll methods tend to be a little higher than that used to form C, due to the approximations made. We have found that using a heuristic factor of $C_0 = 2^{1-2/\alpha}\hat{C}_{\mathrm{FR}}$, gives good results in the parameter ranges of most interest to us ($0.3 < H < 0.7$, $1.3 < \alpha < 2$). This coefficient could of course be fine-tuned in any particular application.

Figure 4 illustrates the estimated scale factors and α values from a typical 1D, 10^5-step fLm process generated using this technique. The correct fractal scaling and stable α estimates are evident.

CONCLUSIONS

In summary, it can be seen that there is no rigorous way to formulate a theory of multivariate Lévy distributions, which is practical for modern geostatistical applications. For smaller systems, with perhaps a few thousand nodes, the more

satisfactory distribution (22) can be sampled exactly, but this is of limited use in large models.

For the case of stationary processes, a satisfactory approximate way to simulate such processes has been presented, hinging essentially on the inverse normal-score transform. The method reproduces the desired α parameter over the range $1 < \alpha < 2$ — estimated from spatial averages of a single large realization—with a small upward bias that will typically be small compared to the uncertainty due to the limited data set. By contrast, the upward biases evident in spatial averages of realizations generated with the algorithm of Painter (1998), are usually statistically significant. These observations carry across to the nonstationary case.

Traditionalists will doubtless urge that the stationary fLn method presented here is no different than the classical methods of normal scoring, simulation, and then inverse scoring, and this is exactly so. The persistent, stable non-Gaussianity displayed by the algorithm presented in the **Alternative, Higher Order Generalizations** subsection merely reveals a little-known aspect of normal-scoring methods: if the data distribution is symmetric and has thick Lévy-like tails, then the distributions of the increments will be similar. Of course, if the univariate distribution is nonsymmetric, the distribution of the increments will be symmetric, but this alone will not be a sufficient basis on which to construct a model. One can conclude then that, in the case of stationary processes, there is no advantage in using multivariate Lévy formulations over traditional normal-scoring methods. Interestingly though, it is possible to exploit the normal-scoring procedure such that a stationary model with a Bessel function covariance can be used to generate a pseudo-fLm process, valid for lags up to a certain cutoff.

In the nonstationary case, an approximate algorithm has been described which adequately reproduces the desired properties of power–law correlations and Lévy distributed increments. The method produces the desired α parameter, within statistical error, over the full range of interest $1 < \alpha < 2$. It reproduces the correct generalized covariance and increment distributions only approximately, with a small error that averages out over many different spatial configurations of data and random simulation paths. This error will be typically small compared to the modelling error induced by the availability of only sparse, clustered data in typical applications. The tendency of the longer-range increments to approach to a more Gaussian distribution can be captured by introducing hard or soft bounds into the simulation algorithm with corresponding rejection criteria, as has been implemented by Painter within the context of the order-1 generalization (Painter, 1998).

For practical applications, we recommend that data whose trends are unambiguous be treated with more classical methods if possible (e.g., sequential Gaussian simulation). The approximate nonstationary Lévy model may be appropriate for data that has unknown trends, if the property variations appear homoscedastic, and the Fama–Roll variogram fits a power–law model. In addition, the increment distribution α factors must be reasonably constant or gently rising as the lag

increases, the latter effect being imposed by the choice of sensible bounds on the variable. The simulated univariate distribution may then only approximately represent that of the data, but this is consistent with the decision of nonstationarity.

REFERENCES

Fama, E. F., and Roll, R., 1968, Some properties of symmetric stable distributions: J. Am. Stat. Assoc., v. 63, p. 817.

Fama, E. F., and Roll, R., 1971, Parameter estimates for symmetric stable distributions: J. Am. Stat. Assoc., v. 66, p. 331–338.

Gaynor, G. C., Chang, E. Y., Painter, S. L., and Paterson L., 1998, Application of Lévy random fractal simulation techniques in modelling reservoir mechanisms in the Kuparuk River Field, North Slope, Alaska: SPE asia pacific conference on integrated modelling for asset management, Kuala Lumpur, March 23–24, SPE 39739.

Gunning, J., and Paterson, L., 1999, Conditioning of Lévy-stable fractal reservoir models to seismic data: SPE annual technical conference and exhibition, Houston, TX.

Jurek, Z. J., and Mason, J. D., 1993, Operator-limit distributions in probability theory: Wiley, New York.

Kumar, P., and Foufoula-Georgiou, E., 1993, A multicomponent decomposition of spatial rainfall fields 2. Self-similarity in fluctuations: Water Resour. Res., v. 29, p. 2533–2544.

Liu, H. H., and Molz, F. J., 1997, Comment on "evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations" by Scott Painter: Water Resour. Res., v. 33, p. 907–908.

Painter, S., 1995, Random fractal models of heterogeneity: The Lévy-stable approach: Math. Geol., v. 27, p. 813–830.

Painter, S., 1996a, Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations: Water Resour. Res., v. 32, p. 1183–1195.

Painter, S., 1996b, Stochastic interpolation of aquifer properties using fractional Lévy motion: Water Resour. Res., v. 32, p. 1323–1332.

Painter, S., 1998, Numerical method for conditional simulation of Lévy random fields: Math. Geol., v. 30, p. 163–179.

Painter, S., and Paterson, L., 1994, Fractional Lévy motion as a model for spatial variability in sedimentary rock: Geophys. Res. Lett., v. 21, p. 2857–2860.

Painter, S., Paterson, L., and Boult, P., 1997, Improved technique for stochastic interpolation of reservoir properties: SPE J., v. 2, no. 1, p. 48–57.

Press, S. J., 1972, Multivariate stable distributions: J. Multivar. Anal., v. 2, p. 444–462.

Wackernagel, H., 1995, Multivariate geostatistics: Springer, Berlin.

APPENDIX A: SUMMARY OF FAMA AND ROLL ESTIMATORS OF LÉVY DISTRIBUTION PARAMETERS

Fama and Roll (1971) have proposed robust quantile-based estimators for the scale factor C_{FR} and width parameter α of the Lévy distributions (4), which are discussed further by Painter (1996a). The parameter α is estimated by computing

$$\eta = \frac{Q_{0.95} - Q_{0.05}}{Q_{0.72} - Q_{0.28}} \tag{34}$$

where the Q_p are sample pth quantiles of a set of N observations of a deviate we assume to be Levy-distributed. We then use the relationship $\alpha(\eta)$ implicit in Table A1 of Painter (1996a) to compute an estimate of α from η. A good approximation of this function, in the range of interest $1 < \alpha \leq 2$, is

$$\alpha(\eta) = \frac{\eta^{1.45}}{-2.3 + 0.94\,\eta + 0.24\,\eta^2}. \tag{35}$$

The scale factor C_{FR} is estimated using

$$\hat{C}_{\mathrm{FR}} = \frac{Q_{0.72} - Q_{0.28}}{1.654}. \tag{36}$$

Simple equations for computing the uncertainties associated with these estimators—which naturally scale like $N^{-1/2}$—are presented in the appendix of Painter (1996a).

APPENDIX B: SEQUENTIAL SIMULATION USING THE ORDER-1 LÉVY GENERALIZATION

The sequential simulation algorithm devised by Painter (1998) proceeds in the usual way to find a set of fixed nodes close to the current unknown node. The conditional density for the unknown node is constructed, sampled, and the algorithm proceeds. A maximum of $n_{\max} = 10$–30 nearest neighbors is imposed, leaving, effectively, the problem of sampling from (17) with n truncated at $n_{\max}$, and $y_{n-1} \ldots y_1$ fixed. According to Painter (1998), the problem of sampling the marginal distribution $P(y_n \mid y_{n-1} \ldots y_1)$ from (3) can be reduced to sampling the density

$$W_{n-1}(z) = \int_0^\infty \cos(z\tau) \int_0^\infty \exp\!\left(-(r^2 + \tau^2)^{\alpha/2}\right) r^{n-2}\, dr\, d\tau \tag{37}$$

for z. Here $z = \sqrt{y_1'^2 + \cdots y_n'^2}$, with $\mathbf{y}' = A\mathbf{y}$, with A a lower triangular matrix that satisfies $A^{\mathrm{T}}A = 2^{2/\alpha}C^{-1}$. Hence the $y_1' \ldots y_{n-1}'$ are fixed. This is the same as sampling from the white noise form (17), with $y_1' \ldots y_{n-1}'$ fixed. For large z, the asymptotic result $W_{n-1}(z) \sim z^{-(n+\alpha)}$ will be of use in the subsequent argument.

It is clear that the basic ideas behind this simulation algorithm are not dependent on the particular functional form of the correlation in C. Painter uses long-range, power–law correlation structures, but the algorithm does not depend on this. All that is required is the operation of some screening behavior, and

the factorization of the multivariate density, which does not hold for the order-1 generalization.

In order to verify that the algorithm can actually produce increments that have a Lévy distribution with parameter α, it is better to work with a short range correlation function, and estimate the α from a large realization using the methods of Fama and Roll. If these tests are carried out on fields with short range correlation structures (or better still, pure nugget effects), it is observed that the estimated α is always significantly greater than that used in the algorithm.

A simple thought experiment will show why. Suppose we wish to generate Lévy white noise using only one nearest neighbor in the sequential algorithm. The sequential algorithm then forms a simple Markov chain for the probability of the nth sample from the probability of the $(n-1)$th sample, using (37) and Bayes' rule:

$$P(y_n) = \int W_1\left(\sqrt{y_{n-1}^2 + y_n^2}\right) P(y_{n-1})\, dy_{n-1} \tag{38}$$

This system must reach stationarity when $P(y_i) = P(y)$ (with the additional constraint $\int P(y)\, dy = 1$ of course). Then Equation (38) constitutes an integral equation for the univariate distribution, $P(y)$, which we desire to be Lévy-stable with parameter α. This distribution has a power–law tail $P(y) \sim y^{-\alpha-1}$. It is then straightforward to show that a first-kind integral equation like (38) can only have a solution with such a power–law tail if the kernel $W_1(z) \sim z^{-1}$ for large z. From the known asymptotic properties of $W_1(z)$, this will clearly not be the case—the kernel decays as $z^{-2-\alpha}$, much faster. What will then happen is that the compact kernel $W_1(z)$ will force a weaker tail in $P(y)$, making the solution more compact than a Lévy distribution, and hence the estimated α values will be closer to the Gaussian case.