“Nonlinear” covariance matrix and portfolio theory for non-Gaussian multivariate distributions*

D. Sornette[†], P. Simonetti[‡§] and J.V. Andersen[¶]

University of California, Los Angeles

First Version: March 1998
This Version: February 1999

Abstract

This paper offers a precise analytical characterization of the distribution of returns for a portfolio constituted of assets whose returns are described by an arbitrary joint multivariate distribution. In this goal, we introduce a non-linear transformation that maps the returns onto gaussian variables whose covariance matrix provides a new measure of dependence between the non-normal returns, generalizing the covariance matrix into a non-linear fractional covariance matrix. This nonlinear covariance matrix is chiseled to the specific fat tail structure of the underlying marginal distributions, thus ensuring stability and good-conditionning. The portfolio distribution is obtained as the solution of a mapping to a so-called ϕ^q field theory in particle physics, of which we offer an extensive treatment using Feynman diagrammatic techniques and large deviation theory, that we illustrate in details for multivariate Weibull distributions. The main result of our theory is that minimizing the portfolio variance (i.e. the relatively “small” risks) may often increase the large risks, as measured by higher normalized cumulants. Extensive empirical tests are presented on the foreign exchange market that validate satisfactorily the theory. For “fat tail” distributions, we show that an adequete prediction of the risks of a portfolio relies much more on the correct description of the tail structure rather than on their correlations.

*We are grateful to P.M. Bentler, F. Lacan, P. Santa-Clara, B. Urosevic and S. Xu for helpful discussions.

[†]Institute of Geophysics and Planetary Physics and Department of Earth and Space Science, University of California, Los Angeles, California 90095 and Laboratoire de Physique de la Matière Condensée, CNRS UMR6622 and Université des Sciences, B.P. 70, Parc Valrose, 06108 Nice Cedex 2, France, E-mail: sornette@cyclop.ess.ucla.edu

[‡]now at Servizio Risk Management, Banca Intesa, Via Boito 7, 20121 Milano,Italy

[§]Department of Physics and Astronomy, University of Southern California, Los Angeles, CA 90089-0484

[¶]Nordic Institute for Theoretical Physics, Blegdamsvej 17, DK-2100 Copenhagen, Denmark

Contents

1 Introduction

Many problems in Finance, including risk management, optimal asset allocation and derivative pricing, require an understanding of the volatility and correlations of assets returns. In practice, volatility and correlations are often estimated from historical data and the risk dimension is represented by the variance or volatility for a single asset and by the covariance matrix for a set of assets. In portfolio optimization, the variance of the distribution of returns is minimized by inverting the covariance matrix of returns. However, the covariance matrix is often ill-conditionned and unstable [Litterman and Winkelmann, 1998]. There are several origins to this unstable behavior, in particular the non-normality of the asset returns, i.e. large price variations are more probable than extrapolated from a normal estimation. For Lévy or power law distributions with index less than two, the covariance matrix is not even defined.

Option pricing and hedging relies on representing risk by a single scalar measure, the volatility. In practice, the volatility has the complexity of a time-dependent fluctuating surface defined as a function of the strike price and the time-to-maturity. Again, an important origin of this complexity stem from the fat tail structure of the underlying distributions as well as their non-stationarity.

The need for models that go beyond the Gaussian paradigm is vividly felt by practionners, regulatory agencies and is also advocated in the academic literature. Let us mention Géczy [1998] who proposes to use non-normal multivariate distributions to better assess factor-based asset pricing models : indeed, some factor models are strongly rejected when relying on the presumption that returns or model residuals are independent and identically distributed multivariate normal, while they are no longer rejected when fatter tail elliptic multivariate distributions are used.

The non-Gaussian nature of empirical return distributions has been first addressed by generalizing the normal hypothesis to the stable Lévy law hypothesis [Mandelbrot, 1963,1997; Fama, 1965]. Gaussian and Lévy laws are stable distributions under convolution and enjoy simple additivity properties that allowed the generalization of Markovitz's porfolio theory in a natural way [Samuelson, 1967; Arad, 1975; Bawa et al., 1979]. More recently, a further generalization has been performed [Bouchaud et al., 1998] to situations where the marginal distributions of asset returns may have different power law behaviors for large positive and negative returns, with arbitrary exponents (possibly larger than 2, i.e. not stable in the sense of Lévy laws but only in a sense of large deviation theory).

This stable or quasi-stable property enjoyed by Lévy and power laws, that are instrumental in the generalization of Markovitz's theory, presents however rather stringent restrictions. Indeed, these laws constitute the only solutions obeying the consistencey property [Kano, 1994], according to which any marginal distribution of random vectors whose distribution belongs to a specific family also belongs to the same family. This consistency property is important for the independence of the marginal variances on the order (number of assets constituting a porfolio, for instance) of the multivariate distributions (see the appendix A). Thus, the generalization of portfolio theory from Gaussian to Lévy and then to power laws relies fundamentally on the consistency property [Kano, 1994].

Elliptical distributions [Cambanis et al., 1981] provide a priori maybe the simplest and most natural hope for descrihing fat tails and for generalizing portfolio theory to non-normal multivariate distributions, solely on the basis of the measure of a covariance matrix. Elliptical distributions are defined as arbitrary density functions F of the quadratic form $(X - \langle X \rangle)'\Sigma^{-1}(X - \langle X \rangle)$, where X is the unit column matrix of the N asset returns and Σ is a $N \times N$ dependence matrix between the N assets which is proportional to the covariance matrix, when it exists. It was argued that the results of the CAPM extends to these elliptical distributions [Owen and Rabinovitch, 1983; Ingersoll, 1987] on the basis that, since any linear transformation of an elliptical random vector is also elliptical, the risk measure $W'\Sigma W$ (where W is the column vector of the asset weights) is positively proportional to the portfolio wealth variance. Thus the ranking of portfolios by risk adverse investors appears to retain their ordering. However, Kano [1994] calls attention to the fact that, in general, the density of a marginal distribution has not the same shape as the function F. This leads to an additional mispecification of the risk since it should be in principle captured both by the covariance matrix Σ and the shape and tail structure of the distribution. It is thus not possible to quantify the risk solely by $W'\Sigma W$. In practice, the marginal variance of one of the variable will depend on the number N of assets used in the portfolio. This is clearly an unwanted property for the chosen multivariate distribution. This inconsistency never arises for special cases such as the multivariate t-distribution or for mixtures of normal distributions. In appendix A, we clarify the origin of the problem and show that, for arbitrary elliptic distributions, the portfolio return distribution depends on the asset weights P, not solely via the quadratic form $W'\Sigma W$: the weights W also control the shape as a whole of the portfolio return distribution [Sornette, 1998]. Minimizing the variance $W'\Sigma W$ of the portfolio may actually distord the portfolio return distribution such as to increase the probability weight far in the tail, leading to increased large risks. This result provides a first cautionary note on any attempt to calibrate the portfolio risks by a single volatility measure.

This paper aims precisely at offering a novel approach to address these questions. Our key idea is to use a representation that strives to remain as parsimonious as the Gaussian framework while capturing the non-Gaussian "fat tail" nature of marginal distributions and the existence of nonlinear dependence. For this, a non-linear transformation maps the returns onto Gaussian variables whose covariance matrix provides a new measure of dependence between the non-normal returns, generalizing the covariance matrix into a non-linear fractional covariance matrix. In our approach, the linear matrix calculation of the standard portfolio theory is replaced by Feynman's diagram calculations and large deviation theory resulting from the fact that the portfolio wealth is a nonlinear weighted sum of the Gaussian variables. Notwithstanding this higher sophistication, all results can be derived analytically and thus can be fully controlled. In the present paper, we focus on the risk component of the problem, i.e. we assume symmetric return distributions. In other words, the average returns are sufficiently close to zero that the fluctuations dominates. Even if not true in reality, this approximation provides a precise representation of real data at sufficiently small time scales, corresponding to Sharpe ratios less than one. This situation is reasonably accurate for non-stock market assets, which do not have a long-term trend. To test and validate our theory, we thus use empirical data from the Foreign exchange. We leave for a future work the extension of our theory to the cases where the average return cannot be neglected.

The paper is organized as follows. In section 2, we present our novel method for a full analytical characterization of the distribution of returns for a portfolio constituted of assets with an arbitrary multivariate distribution. In this goal, we introduce a new measure of dependence between non-normal variables, which generalizes the covariance matrix. The proposed nonlinear covariance matrix is chiseled to the specific fat tail structure of the underlying marginal distributions, thus ensuring stability and good-conditionning. In our formulation, the multivariate distribution is normal in terms of a set of reduced variables that are the natural quantities to work with. In section 3, we show how the calculation of the distribution of returns for an arbitrary portfolio can be formulated in terms of a functional integral approach. In the appendices, we provide an extensive synthesis of the tools that allow us to determine analytically the full distribution of returns and work out in details the calculations for multivariate Weibull distributions. In section 4, we use the information captured by the distribution of returns of arbitrary portfolios to compare them with respect to their moderate versus large risks. We develop especially our analysis for uncorrelated assets with "fat tails" for which we compare different portfolios such as the ones that minimize the variance, the excess kurtosis or higher normalized cumulants. We also compare these portfolios with those determined from the maximization of the expected Utility. We show that minimizing the portfolio variance (i.e. the relatively "small" risks) may often increase the large risks, as measured by higher normalized cumulants. In section 5, we offer practical implementations and comparisons with data obtained from the Foreign exchange market. We validate clearly the performance of the proposed nonlinear representation. The comparisons between our theoretical calculations and empirical estimations of the cumulants of the portfolio return distributions are very satisfactory. We find in particular that, for a good representation of the risks of a portfolio, tracking the correlations between assets is much less important than precisely accounting for the "fat tail" structure of the distributions. Section 6 concludes with future extensions.

2 Non-linear mapping to normal distributions

The first step of our approach is to perform a non-linear change of variable, from the return δx of an asset over the unit time scale τ (taken as the daily scale for illustration below) onto the variable $y(\delta x)$ such that the distribution of y is normal. For marginal distributions, this is always possible. We first illustrate this change of variables in simple cases and then proceed to the general multivariate situation.

2.1 The case of a single security

2.1.1 Non-centered case

Let us consider the class of marginal distribution (probability density function or pdf) $p(\delta x)$ which can be parameterized as

$$p(\delta x) d\delta x = C \frac{f'(\delta x)}{\sqrt{|f(\delta x)|}} \, e^{-\frac{1}{2} f(\delta x)} \, d\delta x \,, \tag{1}$$

where C is a normalizing constant and f' denotes the derivative of the function $f(\delta x)$. Note that f must go to $+\infty$ for $|\delta x| \to \infty$) at least logarithmically to ensure normalization. This distribution (1) describes so-called Von Mises variables [Embrechts et al., 1997]. This parameterization (1) covers all cases where the pdf has a single maximum. These are the relevant situations for essentially all securities. Note that for $f(\delta x)/(\delta x)^2 \to 0$ for large $|\delta x|$, the pdf has "fat tails", i.e. the pdf decays slower than a Gaussian for large $|\delta x|$.

Let us then define the variable y by

$$y = \mathrm{sign}(\delta \mathrm{x}) \, \sqrt{|\mathrm{f}(\delta \mathrm{x})|} \,. \tag{2}$$

The sign function is essential in order to keep possible correlations. In the case of fat tails for which $\sqrt{|f(\delta x)|} << |\delta x|$ for large $|\delta x|$, the variable y is the result of a "contraction" of δx (such that its pdf is normal). Indeed, the pdf of y is such that the conservation $p(y)dy = p(x)dx$ of probability holds and we obtain

$$p(y) dy = \frac{1}{\sqrt{2\pi}} \, e^{-\frac{y^2}{2}} \, dy \,. \tag{3}$$

Obviously, the form (1) has been chosen such that a change of variable recovers the normal pdf through a simple transformation. We stress that this parameterization (1) is fully general as any unimodal pdf can be put in this form.

2.1.2 Example : the modified Weibull distribution

Consider the case of the modified Weibull distribution where

$$f(\delta x) = 2\left(\frac{|\delta x|}{\chi}\right)^{c} , \tag{4}$$

and thus

$$p(\delta x) d\delta x = d\delta x \, \frac{1}{2\sqrt{\pi}} \frac{c}{\chi^{\frac{c}{2}}} \, |\delta x|^{\frac{c}{2}-1} \; e^{-(\frac{|\delta x|}{\chi})^{c}} \quad \text{for} \;\; -\infty < \delta x < +\infty \, . \tag{5}$$

The case, where the exponent $c < 1$, corresponds to a "stretched" exponential with a tail fatter than an exponential and thus much fatter than a Gaussian, but still thinner than a power law. Stretched exponential pdf's have been found to provide a parsimonious and accurate fit to the full range of currency price variations at daily intermediate time scales [Laherrère and Sornette, 1998]. This stretched exponential model is also validated theoretically by the recent demonstration that the tail of pdfs of products of a finite number of random variables is generically a stretched exponential [Frisch and Sornette, 1997], in which the exponent c is proportional to the inverse of the number of generations (or products) in a multiplicative process.

In this case (4), the change of variable (2) corresponds to

$$y = \text{sign}(\delta \text{x}) \; |\delta \text{x}|^{\frac{c}{2}} \, . \tag{6}$$

For $c < 2$, $y(\delta x)$ has a negative (resp. positive) curvature for $\delta x > 0$ (resp. for $\delta x < 0$), i.e. is concave (resp. convex). This convexity reflects the contracting nature of the mapping that allows to obtain a normal distribution, with a standard deviation $\xi = \chi^{\frac{c}{2}}$. This change of variable (2) has already been briefly mentionned in the literature. In a footnote of [Jorgensen, 1987], we find mentionned that, for distributions (5) with $c > 2$, the change of variable (6) gives a normal distribution. To our knowledge, the case $c < 2$ has not been investigated.

The pdf (5) is called "modified" Weibull distribution, because it is slightly different by the distinct power in the pre-exponential factor from the standard Weibull pdf defined by

$$p_W(\delta x) \sim |\delta x|^{c-1} \; e^{-(\frac{|\delta x|}{\chi})^{c}} , \tag{7}$$

such that its cumulative distribution is exactly $e^{-(\frac{|\delta x|}{\chi})^{c}}$. This difference accounts for the fact that the change of variable (6) maps exactly the modified Weibull pdf onto a Gaussian pdf. This property does not hold exactly for the standard Weibull distribution. Notice that the modified Weibull distribution with $c = 2$ is nothing but the Gaussian distribution. This provides another incentive in this definition (5) as it retrieves exactly the Normal law as one of its member. From now on, we drop the term "modified" as we will only study (5).

The Weibull distribution (5) is not stable under convolution, i.e. the distribution of weekly returns is not exactly of the same form as the distribution of the daily returns. However, it presents an approximate stability in the sense that it is possible to define an effective exponent c_T and scale χ_T such that the tail of the pdf of the distribution of return over T days is of the form (5) with c replaced by c_T and χ by χ_T. The analytical procecedure to derive this result and numerical tests are given in the appendix B.

2.1.3 Centered case

We generalize (1) further by the following parameterization :

$$p(\delta x)d\delta x = C\frac{f'(\delta x)}{\sqrt{|f(\delta x)|}}\ e^{-\frac{1}{2}(\mathrm{sign}(\delta x)\ \sqrt{|f(\delta x)|}-m)^2}\ d\delta x\ , \tag{8}$$

where $m = \mathrm{E}[(\mathrm{sign}(\delta x_a)\ \sqrt{|f(\delta x)|}]$ and $\mathrm{E}[x]$ is the expectation of x. The parameterization (8) is such that, by the change of variable (2), expression (8) transforms into a standard centered Gaussian distribution

$$p(y)dy = \frac{dy}{\sqrt{2\pi}}\ e^{-\frac{(y-m)^2}{2}}\ . \tag{9}$$

For the Weibull (stretched exponential) (5), this corresponds to

$$p(\delta x)d\delta x = d\delta x\ \frac{1}{2\sqrt{\pi}}\frac{c}{\chi^{\frac{c}{2}}}\ |\delta x|^{\frac{c}{2}-1}\ e^{-\frac{(\mathrm{sign}(\delta x)\ |\delta x|^{\frac{c}{2}}-m)^2}{\chi^c}} \quad \text{for}\ \ -\infty < \delta x < +\infty\ , \tag{10}$$

where $m = \mathrm{E}[(\mathrm{sign}(\delta x)\ |\delta x|^{\frac{c}{2}}]$. The form (10) of the Weibull exponential corresponds to the following cumulative distribution function [1]

$$P_<(\delta x) = \frac{1}{2}\ \mathrm{erfc}\left(\frac{|\delta x|^{\frac{c}{2}}+m}{\chi^{\frac{c}{2}}}\right)\ . \tag{11}$$

2.2 General case of several assets

Let us assume that the returns of N assets over a period of T time intervals are measured : $\delta x_1^{(1)}, \delta x_2^{(1)}, ..., \delta x_T^{(1)}$ for the first asset, $\delta x_1^{(2)}, \delta x_2^{(2)}, ..., \delta x_T^{(2)}$ for the second asset, up to $\delta x_1^{(N)}, \delta x_2^{(N)}, ..., \delta x_T^{(N)}$ for the last asset. From these, the N marginal distributions of the returns for each asset are determined empirically either using a parametric or non-parametric representation. Call $P_j(\delta x)$ the marginal distribution of the jth asset. In general, $P_j(\delta x)$ is non-gaussian and exhibits fat tails, for instance described by a Weibull distribution. But we do not restrict ourselves to this special case and consider a general form for the $P_j(\delta x)$'s.

2.2.1 Non-linear mapping

As for the one security case discussed previously using (2), each marginal distribution $P_j(\delta x)$ is mapped onto a gaussian distribution by performing a change of variable that is specific to each marginal distribution.

[1] The complementary error function is defined as $\mathrm{erfc}(u) \equiv \frac{2}{\sqrt{\pi}}\int_u^\infty dv\ \exp(-v^2)$, and has the asymptotic behavior $\mathrm{erfc}(u) \to \frac{1}{\sqrt{\pi}}\ \frac{1}{u}\ e^{-u^2}$ for large $u > 0$.

We discuss here the symmetric case. For a given asset j, the new variable is such that

$$P_j(\delta x) d\delta x = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy . \tag{12}$$

By construction, the variable y follows a Gaussian distribution of mean 0 and variance 1. Introducing the cumulative distribution $F_j(x)$

$$F_j(\delta x) = \int_{-\infty}^{\delta x} P_j(x')\, dx' , \tag{13}$$

the differential equation (12) for $y(\delta x)$ is integrated into a general implicit equation giving y as a function of δx:

$$F_j(\delta x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{y}{\sqrt{2}}\right)\right] . \tag{14}$$

For a Weibull distribution (5), expression (14) reads

$$e^{-(\frac{|\delta x|}{\chi})^c} = \mathrm{erfc}(\frac{y}{\sqrt{2}}) . \tag{15}$$

We can rewrite (14) as

$$y(\delta x) = \sqrt{2}\ \ \mathrm{erf}^{-1}\left(2F_j(\delta x) - 1\right) \tag{16}$$

where erf^{-1} is the inverse of the error function. If $P_j(\delta x)$ is of the form (1), the solution of (16) retrieves the expression (2).

We perform this change of variable (12,14,16) for each of the N distributions $P_j(\delta x)$. This leads to the transformed measurements $y_1^{(1)}, y_2^{(1)}, ..., y_T^{(1)}$ for the first asset, $y_1^{(2)}, y_2^{(2)}, ..., y_T^{(2)}$ for the second asset, up to $y_1^{(N)}, y_2^{(N)}, ..., y_T^{(N)}$ for the last asset.

2.2.2 Nonlinear covariance matrix

Having thus defined the variables $y^{(j)}$ $(j = 1...N)$ from the N assets, we construct the covariance matrix

$$V = \mathrm{E}[\mathbf{y}\, \mathbf{y}'] , \tag{17}$$

with element V_{ab} defined by

$$V_{ab} = \frac{1}{T} \sum_{i=1}^{T} y_i^{(a)}\, y_i^{(b)} . \tag{18}$$

In (17), we note for short $\mathbf{y}$ as the unicolumn matrix of size N with elements $y^{(j)}$ equal to some realization.

For distributions that can be parameterized as (8), we have seen that the change of variable (16) becomes expression (2). The covariance matrix (17) of the set of y_a's then corresponds to the following dependence matrix for the asset returns δx_a:

$$V_{ab} \equiv \mathrm{E}\left[\left(\mathrm{sign}(\delta x_a)\sqrt{|f(\delta x_a)|} - m_a\right)\ \left(\mathrm{sign}(\delta x_b)\ \sqrt{|f(\delta x_b)|} - m_b\right)\right] , \tag{19}$$

where $m_a = \mathrm{E}[\mathrm{sign}(\delta x_a)\ \sqrt{|f(\delta x_a)|}]$. The distribution of the y_a is multivariate Gaussian with covariance matrix V_{ab}.

The covariance matrix (17,19) measures the covariance between the assets described by the "effective" returns y. This definition (17,19) amounts to a non-standard measure of the covariance in terms of the variables δx's. We use the name "nonlinear covariance" to recall that the change of variable $\delta x \rightarrow y$ for fat tail pdf's $P_j(\delta x)$ lead to a contraction, for instance a concave power law (6) for $c < 2$. In this case, the covariance of y amounts to estimate the covariance of fractional powers of the returns.

2.2.3 Multivariate representation of the joint distribution of asset returns

No approximation has been made until now. We have characterized the marginal distributions of the assets, defined new variables in terms of which the marginal distributions are Gaussian and we have calculated the covariance matrix of these new variables as the new measure of the dependence between the asset returns.

Here comes the simplifying approximation. If the marginal distributions in terms of the new variables $y^{(j)}$ are gaussian, nothing imposes a priori that the multivariate distribution of these variables is also a multivariate Gaussian distribution. From our construction, it is only guaranted that the projections of the multivariate distribution onto each $y^{(j)}$ variable are Gaussian.

However, a standard theorem from Information Theory [Rao, 1973] tells us that, conditioned on the sole knowledge of the covariance matrix, the best representation of a multivariate distribution is the Gaussian. In other words, the multivariate normal distribution contains the least possible a priori assumptions in addition to the covariance matrix (and the vector of the means when they are non-zero), i.e. is the most likely representation of the data.

This implies that, conditionned on the sole knowlege of the covariance matrix (17,18), the best parametric, although not exact, representation of the multivariate distribution $\hat{P}(\mathbf{y})$ is

$$\hat{P}(\mathbf{y}) = (2\pi)^{-N/2}\,|V|^{-1/2} \exp\left(-\tfrac{1}{2}\,\mathbf{y}'\,V^{-1}\,\mathbf{y}\right)\ , \tag{20}$$

where $|V|$ is the determinant of V.

Note that our "nonlinear covariance" approach is very different from the usual approximation which assumes that the asset returns $\delta x_i^{(j)}$ themselves follow a multivariable Gaussian distribution. In particular, the non-linear change of variable ensures that the fat tail structure of the marginal distribution is fully described and, in addition, leads to a novel more stable measure of the covariances. The normal structure of the multivariate distribution in terms of the y variables does *not* lead to a Gaussian multivariate distribution for the returns.

From the expression (20), we obtain the explicit expression of the multivariate distribution of the asset returns by replacing the $y_i^{(j)}$ as a function of $\delta x_i^{(j)}$ as given by (16) for each asset j and use the identity

$$P(\mathbf{x}) = \hat{P}(\mathbf{y})\ \frac{d\mathbf{y}}{d\mathbf{x}}\ , \tag{21}$$

where $\frac{d\mathbf{y}}{d\mathbf{x}}$ is the jacobian of the transformation from $\mathbf{x} \rightarrow \mathbf{y}$. The Jacobian is the determinant of a diagonal matrix whose jth element is simply

$$\frac{dy^{(j)}}{d\delta x^{(j)}} = \sqrt{\pi}\,P_j(\delta x^{(j)})\,e^{(y^{(j)})^2/2} \tag{22}$$

as seen from (16). This finally yields the following approximation for the multivariate distribution $P(\mathbf{x})$ of the column vector $\mathbf{x}$ of N returns of a given realization of the N asset returns :

$$P(\mathbf{x}) = |V|^{-1/2} \exp\left(-\tfrac{1}{2}\,\mathbf{y}'\,(V^{-1} - I)\,\mathbf{y}\right) \prod_{j=1}^{N} P_j(x^{(j)}) \ , \tag{23}$$

where V is again the covariance matrix for $\mathbf{y}$ and I is the identity matrix.

This "nonlinear covariance" approximation is exact for distributions with uncorrelated variables, in which case $V = I$. It is also exact for a Gaussian distribution modified by monotonic one-dimensional variable transformations for any number of variables; or equivalently, multiplication by a non-negative separable function. Note that the multivariate distribution (23) obeys automatically the condition that the corresponding marginal distributions [2] are of the same analytic form. This corresponds to the consistency condition [Kano, 1994] discussed in the introduction. This approach leading to (23) has also been introduced independently for the analysis of particle physics experiments by Karlen [1998].

2.2.4 Multivariate Weibull distributions

Let us now apply this approach to Weibull distributions. The N assets are assumed to have their returns δx_a distributed according to the *unconditional* marginal distributions given by expression (8). With the change of variable (2), we construct the covariance matrix V defined by (19). This defines a signed and fractional dependence matrix for the asset returns. Since $V_{aa} = \mathrm{E}[[(\delta x)^2]^{\frac{c}{2}}]$, we see that $[V_{aa}]^{\frac{2}{c}}$ has the same "dimension" as the usual variance but is fundamentally different in that

$$\mathrm{E}\left[(\delta x)^2\right]^{\frac{c}{2}} \neq \mathrm{E}\left[[(\delta x)^2]^{\frac{c}{2}}\right] \ , \qquad \text{in general} \ . \tag{24}$$

The expectation of a power is in general different from the power of the expectation.

The multivariate Weibull exponential distribution of the N asset returns is then given by expression (23) that we write explicitely

$$P(\delta x_1, \delta x_2, ..., \delta x_N) \prod_{i=1}^{N} d\delta x_i \propto \prod_{i=1}^{N} \left(\frac{c}{2} |\delta x_i|^{\frac{c}{2}-1} \, d\delta x_i \right)$$

$$\exp\left(-\frac{1}{2} \ \sum_{i=1}^{N} \sum_{j=1}^{N} \ V_{ij}^{-1} [\mathrm{sign}(\delta \mathrm{x_i}) \ |\delta \mathrm{x_i}|^{\frac{c}{2}} - \mathrm{m_i}) \ (\mathrm{sign}(\delta \mathrm{x_j}) \ |\delta \mathrm{x_j}|^{\frac{c}{2}} - \mathrm{m_j}) \right) \ . \tag{25}$$

This expression obeys the requirement that the marginal (monovariate) distributions are Weibull exponentials.

[2] i.e. the monovariate unconditional distributions derived from the multivariate distribution by unconditionally integrating over all variables except one.

3 The distribution $P_S(\delta S)$ of portfolio wealth variations

3.1 Theoretical formulation

We now restrict our study to symmetric multivariate distributions. As a consequence, all odd moments (in particular the first moment which is the expected return) and all odd cumulants are vanishing. This means that we are focusing on the aspect of risk embedded in the variance and in all higher even order cumulants. There is no trade-off between risk and return since the expected return is zero. We thus focus exclusively on the risk dimension of the portfolio. In another paper, we will present our results obtained for non-symmetric distributions. The results presented below are already sufficiently rich that we feel compelled to separate the discussion of symmetric and non-symmetric distributions so as not to make the presentation unecessarily cumbersome.

To a very good approximation, it is harmless and much simpler to replace the returns $\delta x_i(t)$ by $(p_i(t+1)-p_i(t))/p_i(t)$ where $p_i(t)$ is the price of asset i at time t. Over reasonable large time intervals (e.g a year), one can neglect the variation of the denominator in comparison to the variation of the numerator $\delta p_i(t) \equiv p_i(t+1)-p_i(t)$.

The total variation of the value of the portfolio made of N assets between time $t-1$ and t reads

$$\delta S(t) = \sum_{i=1}^{N} w_i \delta p_i(t) \ , \tag{26}$$

where w_i is the weight of the ith asset in the portfolio. By normalization, we have $\sum_{i=1}^{N} w_i = 1$.

In terms of the variables y_i's defined by (2) (resp. (6) for the Weibull exponential case), the expression (26) reads

$$\delta S(t) = \sum_{i=1}^{N} w_i \ \mathrm{sign}(\mathrm{y_i}) \ \mathrm{f}^{-1}(\mathrm{y_i^2}) \ , \tag{27}$$

where f^{-1} is the inverse of f (suitably defined to avoid double valuedness) and respectively

$$\delta S(t) = \sum_{i=1}^{N} w_i \ \mathrm{sign}(\mathrm{y_i}) \ \chi_\mathrm{i} \ |\mathrm{y_i}|^{\frac{2}{\mathrm{c}}} \ , \tag{28}$$

for the Weibull distribution for which $f_i(\delta p_i) = |\frac{\delta p_i}{\chi}|^c$.

All the properties of the portfolio are contained in the probability distribution $P_S(\delta S(t))$ of $\delta S(t)$. We would thus like to characterize it, knowing the multivariate distribution of the δp_i's (or equivalently the multivariate Gaussian distribution of the y_i's) for the different assets. The general formal solution reads

$$P_S(\delta S) = C \prod_{i=1}^{N} \left(\int dy_i \right) \ e^{-\frac{1}{2} \ \mathbf{y}' V^{-1} \mathbf{y}} \ \delta\!\left(\delta S(t) - \sum_{i=1}^{N} w_i \mathrm{sign}(\mathrm{y_i}) \ \mathrm{f}^{-1}(\mathrm{y_i^2}) \right) \ . \tag{29}$$

Taking the Fourier transform $\hat{P}_S(k) \equiv \int_{-\infty}^{+\infty} d\delta S \ P_S(\delta S) \ e^{-ik\delta S}$ of (29) gives

$$\hat{P}_S(k) = \prod_{i=1}^{N} \left(\int dy_i \right) \ e^{-\frac{1}{2} \ \mathbf{y}' V^{-1} \mathbf{y} + ik \ \sum_{i=1}^{N} \ w_i \ \mathrm{sign}(\mathrm{y_i}) \ \mathrm{f}^{-1}(\mathrm{y_i^2})} \ . \tag{30}$$

In the sequel, we present a systematic calculation method of (30), that can be applied in principle to a large variety of functional forms for f. In this paper, we restrict our analysis to the Weibull case, because it is probably one of the simplest non-trivial situation that allows us to make apparent the power of this approach. In addition, as will be shown by extended comparisons with empirical data, the Weibull representation provides a very reasonable description of the "fat tail" structures of empirical distributions. For the Weibull case, the term $f^{-1}(y_i^2)$ is replaced by $\chi |y_i|^{\frac{2}{c}}$.

A further simplification of notation occurs when the exponent c is given by

$$c = 2/q \ , \qquad \text{with } q \text{ integer and odd} \ . \tag{31}$$

Then,

$$sign(u_i)\ |u_i|^{\frac{2}{c}} = sign(u_i)\ |u_i|^q = u_i^q \ , \tag{32}$$

i.e. the sign function disappears. Expression (30) then becomes

$$\hat{P}_S(k) = \frac{1}{(2\pi)^{N/2} \det V^{1/2}} \prod_{i=1}^{N} \left(\int dy_i \right) e^{-\frac{1}{2}\ \mathbf{y}' V^{-1} \mathbf{y} + ik\ \sum_{i=1}^{N}\ w_i\ \chi_i y_i^q} \ . \tag{33}$$

We can absorb the terms χ_i by introducing the new variable $u_i^q = \chi y_i^q$, which shows that the covariance elements (in the formulation where the χ_i do not appear) are proportional to $\chi^{2/q}$.

Expressions (27,28) together with (33) exemplify the origin of the complexity introduced by the non-Gaussian nature of the distributions, namely the portfolio wealth is a nonlinear function of the variables y transformed from the asset returns and its distribution is expressed as a non-Gaussian multivariate integral, which is a priori not obvious to estimate. We notice however that (33) is similar to mathematical objects studied in another context: it is the same as the partition function of a ϕ^q field theory studied in particle physics, with N components and imaginary coupling coefficients ikw_i. $\mathbf{y}$ is like a Gaussian field with interactions described by the second nonlinear term in the exponential. Notice that the case $q = 3$ corresponding to an exponent $c = 2/3$ is particularly interesting, because this value for c seems realistic empirically (see below the empirical section) and because the ϕ^3 theory is the simplest non-trivial case leading to fat tails.

We shall characterize the portfolio distribution $P(\delta S(t))$ by its cumulants c_n defined by

$$\hat{P}_S(k) = \exp\left(\sum_{m=1}^{+\infty} \frac{c_m}{m!} (ik)^m \right) \ . \tag{34}$$

The first cumulant c_1 is the mean, the second cumulant $c_2 \equiv E[(x - \mathrm{E}[x])^2]$ is the variance, the third cumulant c_3 defines the skewness $c_3/c_2^{3/2}$, the fourth cumulant c_4 defines the excess kurtosis $\kappa \equiv c_4/c_2^2$. [3] For Gaussian distributions, all cumulants of order larger than two are zero. The cumulants are thus convenient ways to quantify the departure from normality. Furthermore, provided mild regularity conditions hold, they are equivalent to the knowledge of the full distribution. Since we deal here exclusively with symmetric distributions, all odd-order cumulants are vanishing and it is sufficient to calculate the even-order cumulants.

[3] The kurtosis is defined as the ratio of the fourth centered moment over the square of the variance $\mathrm{E}[(x-\mathrm{E}[x])^4]/\mathrm{E}[(x-\mathrm{E}[x])^2]^2 = 3 + c_4/c_2^2$. In this definition, Gaussian distributions have a kurtosis equal to 3 while the excess kurtosis is zero.

3.2 The Gaussian case $c = 2$

When $c = 2$, the multivariate asset return distribution is Gaussian. The term $\mathrm{sign(y_i)}\ |\mathrm{y_i}|^{\frac{2}{c}}$ becomes simply y_i and the integrals in (30) are standard Gaussian integrals that can be evaluated using the identity

$$\prod_{i=1}^{N}(\int dx_i)\exp\left(-\frac{1}{4}x_iA^{-1}_{ij}x_j + y_ix_i\right) = \sqrt{\frac{\pi^N}{det(A_{ij})}}\exp(y_iA_{ij}y_j)\ . \tag{35}$$

Applied to (30) and after taking the inverse Fourier transform, we get

$$P(\delta S(t)) \propto \exp\left(-\frac{[\delta S(t)]^2}{2W'VW}\right)\ . \tag{36}$$

which recovers the classical result that the distribution of portfolio wealth variations δS is uniquely and fully characterized by the sole value of its variance c_2, expressed in terms of the covariance matrix of the asset price returns and their weight W in the portfolio [Markovitz, 1959, Merton, 1990] :

$$c_2 = W'VW\ . \tag{37}$$

We now turn to the more general case $c \neq 2$ for which the evaluation of the integral (30) is much more involved.

3.3 The case of uncorrelated assets

We first consider the case where the covariance matrix is diagonal, which corresponds to an absence of correlations between assets. Then, the portfolio distribution is only sensitive to the intrinsic risks presented by each asset. We will show in the next subsection how to implement a perturbative diagrammatic expansion in the general correlated case.

For notational convenience, we adopt a change of variable which absorbs the χ's in the exponential term of (33) : the term $\sum_{i=1}^{N}\ w_i\ \chi_i y_i^q$ becomes $\sum_{i=1}^{N}\ w_i\ y_i^q$. As a consequence, the new y_i variables acquire a non-unit variance equal to χ_i^c, with $c = 2/q$.

In the diagonal case, the covariance matrix is given by

$$V = \mathrm{diag}\{d_i\}\ , \qquad \text{where } d_i = \chi_i^{2/q}\ . \tag{38}$$

The relationship between the variances of these y variables and the variances of the asset returns is obtained by noting that $d_i^q = \chi_i^2$ which is proportional to the asset return variance. This illustrates that d_i plays the role of a "nonlinear variance".

3.3.1 Cumulants

In the diagonal case, the multiple integral over the multiple asset contributions in (33) becomes the product of single integrals of the kind

$$I_i = \int_{-\infty}^{+\infty} du\ e^{-\frac{u^2}{2d_i} + ikw_iu^{q_i}}\ . \tag{39}$$

We consider the general case where we allow the exponent $c_i = 2/q_i$ to vary from asset to asset. From this integral, we derive in appendix C the exact and explicit expression of the general $2r$-th cumulant of $P(\delta S)$:

$$c_{2r} = \sum_{i=1}^{N} C(r, q_i) (w_i^2 d_i^{q_i})^r \; , \tag{40}$$

where

$$C(r,q) = (2r)! \; 2^{qr} \left\{ \sum_{n=0}^{r-2} (-1)^n \frac{\Gamma\left((r-n)q + \frac{1}{2}\right)}{(2r-2n)! \pi^{1/2}} \left[\frac{\Gamma\left(q+\frac{1}{2}\right)}{2! \pi^{1/2}} \right]^n - \frac{(-1)^r}{r} \left[\frac{\Gamma\left(q+\frac{1}{2}\right)}{2! \pi^{1/2}} \right]^r \right\} , \tag{41}$$

and Γ denotes the Gamma function.

The expression (40) is valid even when the q_i's are real and the interaction term in (33) is proportional to $\text{sign}(u_i)|u_i|^{q_i}$ and thus applies to arbitrary Weibull distributions.

For $r = 1$ (variance) and $r = 2$ (fourth order cumulant), the expression (41) gives

$$C(1,q) = (2^q/\sqrt{\pi}) \Gamma(q+1/2) \tag{42}$$

$$C(2,q) = (2^{2q}/\sqrt{\pi}) \; \Gamma(2q+1/2) - (3 \; 2^{2q}/\pi) [\Gamma(q+1/2)]^2 \; . \tag{43}$$

Note that the dependence of the cumulants in the d_i's as given in (40) enters through their qth power d_i^q since, as we noticed before, V_{ab} is proportional to the moment $\langle [(\delta x)^2]^q \rangle$. As already mentionned, $[V_{ab}]^q$ (with $q = \frac{2}{c}$) has the same "dimension" as the usual variance. The dependence of the cumulants $c_{2r}(q)$ on the weights w_i's enters only through the terms $w_i^2 d_i^q$ which are the same for all cumulant orders.

3.3.2 Tails of the portfolio distribution for $c < 1$

We now characterized the extreme tails of the distribution of wealth variations of the total portfolio constituted of N assets. We present results for the case where all assets have the same exponent. It is easy to reintroduce the dependence of $c_i = 2/q_i$ in the formulas, as it will become necessary when comparing to the empirical data in section 5.

For this, we use (40) in (34) to get

$$\hat{P}_S(k) = \exp\left(\sum_{m=1}^{+\infty} \frac{(-k^2)^m}{(2m)!} C(m,q) \sum_{i=1}^{N} (w_i^2 d_i^q)^m \right) \; . \tag{44}$$

From the expression (41) of $C(m,q)$ (see also Appendix C), we see that

$$\ln C(m,q) \rightarrow_{m \rightarrow +\infty} \quad (\frac{q}{2} - 1) \; 2m \ln 2m + O(2m) \; . \tag{45}$$

Two cases must be distinguished.

1. $q > 2$: $C(m,q)$ increases faster than exponentially for large cumulant orders $2m$. As a consequence, the tail of the distribution $P_S(\delta S)$ is controlled by the large cumulants.

2. $q \leq 2:$ $C(m,q)$ decreases with m and we cannot derive the structure of the tail from the large cumulants.

For $q > 2$, i.e. $c < 1$ (stretched Weibull distributions), the behavior of large order cumulants embodies the structure of the tail of $P_S(\delta S)$. From (40), we see that

$$c_{2m}(q) \rightarrow_{m \rightarrow +\infty} \quad C(m,q) \left[\text{Max}_i \;\; w_i^2 d_i^q \right]^m , \tag{46}$$

i.e. the high-order cumulants of the portfolio distribution are controlled by the single asset that maximizes the product $w_i^2 d_i^q$ of the square of the weight with the variance of the return. We call this maximum $w_{\max}^2 d_{\max}^q$. Keeping only this term (46) for the expression of the cumulants $c_{2m}(q)$ of large order, we see that $\hat{P}_S(k)$ takes the form of the cumulant expansion for a single Weibull distribution with d_i^q replaced by $w_{\max}^2 d_{\max}^q$. We then deduce the expression for the extreme tail

$$P_S(\delta S) \rightarrow_{|\delta S| \rightarrow \infty} \exp\left[-\left(\frac{|\delta S|}{w_{\max} d_{\max}^{q/2}} \right)^c \right] = \exp\left[-\left(\frac{|\delta S|}{w_{\max} \chi_{\max}} \right)^c \right] \quad \text{where } \chi_{\max} \equiv d_{\max}^{q/2} . \tag{47}$$

For $c < 1$, we have thus shown that the extreme tail of the portfolio distribution is of the same functional form as the distribution of the individual assets with a characteristic decay rate $w_{\max}\chi_{\max}$ controlled by a single asset. The selection of this asset is a function of the parameters of the asset distributions and of the portfolio weights. Note that, for w_{max} of order $1/N$,

$$P_S(\delta S) \sim \exp\left[-N^c \left(\frac{|\delta S|}{\chi} \right)^c \right] , \tag{48}$$

where χ is a coefficient independent of N. The extreme tail of $P_S(\delta S)$ thus decays as the exponential of a power of N smaller than one.

When the exponents $c_i = 2/q_i$ are distinct, the tail of the portfolio distribution is controlled by the asset with the smallest c, i.e. largest q.

3.3.3 Tails of the portfolio distribution for $c > 1$

For $q < 2$, i.e. $c > 1$, the above derivation does not hold. However, we can use the extreme deviation theorem [Frisch and Sornette, 1997] to determine the shape of the extreme tail of the portfolio wealth distribution. This theorem applies only for the exponent $c > 1$ and the stretched Weibull case previously analyzed is excluded from this analysis. This is due to the existence of a log-convexity condition $f_i''(\delta x_i) > 0$, where f'' is the second derivative of f defined in (1), as is shown in the Appendix D and in [Frisch and Sornette,1997].

The Appendix D generalizes the extreme deviation theorem [Frisch and Sornette, 1997] to the case of non-identically distributed random variables. The result is

$$P_N(\delta S) \rightarrow_{\text{large } \delta S} \quad \frac{\pi^{\frac{N-1}{2}}}{X \; \prod_{j=1}^N w_j \chi_j} \left[\frac{2}{c(c-1)} (\frac{\delta S}{\chi})^{2-c} \right]^{\frac{N-1}{2}} \exp\left(-\frac{N}{2(c-1)} (\frac{\delta S}{\hat{\chi}})^c \right) , \tag{49}$$

where

$$\hat{\chi}^c = \frac{(\sum_{j=1}^N w_j \chi_j)^c}{c \frac{(\sum_{j=1}^N w_j \chi_j)^2}{N \sum_{j=1}^N w_j^2 \chi_j^2} + c - 2} . \tag{50}$$

This result is valid for $c > 1$. Note that the extreme tail of $P_N(\delta S)$ decays as the exponential of the number N of assets. This result gives the correct cross-over to the result (48) obtained for $c < 1$.

3.4 The general case of correlated assets

3.4.1 Cumulants of the portfolio distribution with N assets

We assume that all assets are characterized by Weibull distribution with the same exponent $c = 2/q$. Our goal is to compute the characteristic function

$$\hat{P}_S^q(k) = \frac{1}{(2\pi)^{N/2} \det V^{1/2}} \int \left(\prod_i^N du_i \right) e^{-\frac{1}{2} U V^{-1} U + ik \sum_i w_i u_i^q} . \tag{51}$$

For this, we introduce a perturbation analysis and the functional generator [Brezin et al., 1976; Sornette, 1998] defined by

$$\hat{P}_S^q(k, J_i) = \frac{1}{(2\pi)^{N/2} \det V^{1/2}} \int \left(\prod_i^N du_i \right) e^{-\frac{1}{2} u V^{-1} u + ik \sum_i w_i u_i^q + \sum_i J_i u_i} . \tag{52}$$

When the integral is a Gaussian ($k = 0$), we get

$$\begin{aligned} \hat{P}_S^q(0,0) &= 1 \\ \hat{P}_S^q(0,J_i) &= e^{\frac{1}{2} J V J} . \end{aligned} \tag{53}$$

With the relationship

$$f\left(\frac{\delta}{\delta J_i}\right) \int \left(\prod_i^N du_i \right) e^{-\frac{1}{2} U V^{-1} U + \sum_i J_i u_i} = \int \left(\prod_i^N du_i \right) e^{-\frac{1}{2} U V^{-1} U + \sum_i J_i u_i} f(u_i) , \tag{54}$$

we can formally express the characteristic function as

$$\hat{P}_S^q(k) = e^{ik \sum_i w_i \frac{\delta^q}{\delta J_i^q}} \; e^{\frac{1}{2} J V J} \Big|_{J_i = 0} . \tag{55}$$

This formulation is the most natural for a perturbative analysis of (51).

In appendix E, we use the Feynman diagram method to calculate

$$\hat{P}_S^q(k) = \left[1 + i \frac{g_q}{q!} \sum_i w_i \frac{\delta^q}{\delta J_i^q} + \frac{1}{2} \left(\frac{i g_q}{q!} \right)^2 \sum_{i,j} w_i w_j \frac{\delta^q}{\delta J_j^q} \frac{\delta^q}{\delta J_i^q + \ldots} \right] e^{\frac{1}{2} J V J} \Big|_{J=0} , \tag{56}$$

where we have defined the auxiliary coupling constant

$$g_q \equiv q! k . \tag{57}$$

Up to second order in k (which will thus provide the variance of $P_S(\delta S)$), we get

$$\hat{P}_S^q(k) = \left[1 - g_q^2 \sum_{l=0}^{(q-1)/2} \frac{1}{(q-2l)!} \frac{1}{(2!)^{2l+1}} \frac{1}{(l!)^2} \sum_{i,j} w_i \left(V_{ii}\right)^l \left(V_{ij}\right)^{q-2l} \left(V_{jj}\right)^l w_j \right] . \tag{58}$$

The diagrammatic expansion used in appendix E becomes very useful for higher order terms in g_q to obtain a systematic classification of their proliferation. A well-known result of diagrammatic perturbation theory [Veltman, 1995] tells us that neglecting the disconnected diagrams, which appear beyond second order, corresponds to compute the logarithm of the characteristic function. Therefore, the set of connected diagrams at m-th order give us directly the m-th cumulant coefficient c_m defined by expression (34).

We give now the expressions of the second $c_2(3)$ and fourth cumulant $c_4(3)$ corresponding to $q=3$, i.e. $c=2/3$. The second cumulant is read from expression (159) in the appendix

$$c_2(3) = \sum_{ij}\left[6w_i\,(V_{ij})^3\,w_j + 9w_iV_{ii}V_{ij}V_{jj}w_j\right] , \tag{59}$$

where V_{ij} is the i,j element of the covariance matrix V between the variable y_i of asset i and the variable y_j of asset j. The sum is over all pairs of assets. The fourth order cumulant is

$$\begin{aligned} c_4(3) \quad = \quad & 4!(3!)^4 \sum_{i_1,i_2,i_3,i_4} w_{i_1}w_{i_2}w_{i_3}w_{i_4}\left\{\frac{1}{2^4}V_{i_1i_2}^2V_{i_1i_3}V_{i_2i_4}V_{i_3i_3}V_{i_4i_4}+\right. \\ & \frac{1}{2^3}V_{i_1i_2}^2V_{i_1i_3}V_{i_2i_3}V_{i_3i_4}V_{i_4i_4}+\frac{1}{2^4}V_{i_1i_2}^2V_{i_1i_3}V_{i_2i_4}V_{i_3i_4}^2+ \\ & \left.\frac{1}{3!2^3}V_{i_1i_2}V_{i_1i_3}V_{i_1i_4}V_{i_2i_2}V_{i_3i_3}V_{i_4i_4}+\frac{1}{4!}V_{i_1i_2}V_{i_1i_3}V_{i_1i_4}V_{i_2i_3}V_{i_2i_4}V_{i_3i_4}\right\} . \end{aligned} \tag{60}$$

The sum is carried over all possible quadruplets of assets.

This result generalizes for higher order m-th cumulants and for arbitrary q as

$$C_m(q) = m!(q!)^m \sum_{i_1,\dots,i_m} w_{i_1}\cdots w_{i_m} \sum_{G_m(q)} \frac{1}{S(\{l_r\},\{n_{rs}\})}\prod_{r=1}^{m}(V_{i_ri_r})^{l_r}\prod_{r<s=1}^{m}(V_{i_ri_s})^{n_{rs}} \tag{61}$$

where the sum is carried over the set $G_m(q)$ of all the topologically inequivalent connected diagrams with m vertices of q legs as defined in the Appendix E and the symmetry factor S is of the form (see Appendix E)

$$S(\{l_r\},\{n_{rs}\}) = (2!)^{\sum_{r=1}^{m} l_r}\prod_{r=1}^{m} l_r!\prod_{r<s=1}^{m} n_{rs}!S_v(\{l_r\},\{n_{rs}\}) . \tag{62}$$

3.4.2 Portfolio with two assets

As an illustration, for the case of two assets ($N=2$), Eqs.(59,60) reduce to

$$c_2(3) = 15V_{11}^3w_1^2+15V_{22}^3w_2^2+(V_{11}V_{22})^{\frac{3}{2}}\,[6\rho_{12}^3+9\rho_{12}]\,w_1w_2 , \tag{63}$$

where the correlation coefficient is defined by

$$\rho_{12} \equiv \frac{V_{12}}{\sqrt{V_{11}V_{22}}} . \tag{64}$$

The fourth cumulant is

$$\frac{c_4(3)}{4!(3!)^4} = aV_{11}^6w_1^4+aV_{22}^6w_2^4+\left[(a-\frac{1}{2^4}-\frac{1}{4!})V_{11}V_{12}V_{22}+(\frac{1}{2^4}+\frac{1}{4!})V_{12}^3\right](V_{11}^3w_1^3w_2+V_{22}^3w_1w_2^3)$$

$$+2\left((a-\frac{1}{4!})V_{11}^2V_{12}^2V_{22}^2+\frac{1}{4!}V_{11}V_{12}^4V_{22}\right)w_1^2w_2^2\,, \tag{65}$$

where

$$a\equiv\frac{1}{2^4}+\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{3!2^3}+\frac{1}{4!}\,. \tag{66}$$

If the two assets are not correlated, $V_{12}=\rho_{12}=0$, expressions (63) and (65) recover the non-correlated expression (40) as they should, since $4!(3!)^4\,a=C(2,3)=9720$.

For $V_{11}=V_{22}=V$, the expression (65) of the fourth cumulant simplifies into

$$\frac{c_4(3)}{4!(3!)^4V^6}=a[w_1^4+w_2^4]+\rho_{12}[a-(\frac{1}{2^4}+\frac{1}{4!})(1-\rho_{12}^2)](w_1^3w_2+w_1w_2^3)+2\rho_{12}^2[a-\frac{1}{4!}(1-\rho_{12}^2)]w_1^2w_2^2\,. \tag{67}$$

4 Portfolio optimization for uncorrelated assets

For non-Gaussian distributions, all the cumulants and not only the variance of the distribution of the portfolio wealth variations must be considered as relevant measures of risk. The variance as well as all higher order cumulants depend on the weights w_i of the assets constituting the portfolio with different functional forms. It is thus important to determine the relative variation of the cumulants when the weights w_i are modified. In particular, we show now that the portfolio which minimize the variance, i.e. the relatively “small” risks, often increases larger risks as measured by higher normalized cumulants.

4.1 Minimization of the variance

For simplicity of notation, we restrict to the case where all assets have the same exponent $c=2/q$. It is straightforward to reintroduce the asset dependence of the q's in the formulas. In order to take into account the normalization constraint $\sum_i w_i=1$, we solve for the N-th weight $w_N=1-\sum_{i=1}^{N-1}w_i$ and we substitute it in the explicit formula for the variance

$$c_2(q)=(2q-1)!!\left[\sum_{i=1}^{N-1}w_i^2d_i^q+\left(1-\sum_{i=1}^{N-1}w_i\right)^2d_N^q\right]\,. \tag{68}$$

The variance being quadratic in the weigths w_i, the solution to the minimization problem is the solution to the linear algebraic system of $N-1$ equations in $N-1$ variables

$$w_id_i^q=\left(1-\sum_{i=1}^{N-1}w_i\right)d_N^q, \qquad i=1,\ldots,N. \tag{69}$$

The solution can be written in a compact form by using the symmetric polynomials $\sigma_n^{(N)}(x_1,\ldots,x_N)$ generated by $\prod_{i=1}^N(x+x_i)=\sum_{n=0}^N x^{N-n}\sigma_n^{(N)}(x_1,\ldots,x_N)$ such that $\sigma_0^{(N)}=1$, $\sigma_1^{(N)}=x_1+x_2+\ldots+x_N$, ... and $\sigma_N^{(N)}=x_1x_2\cdots x_N$. The solution takes the form

$$w_1d_1^q=w_2d_2^q=\ldots=w_Nd_N^q=\frac{\sigma_N^{(N)}(d_1^q,\ldots,d_N^q)}{\sigma_{N-1}^{(N)}(d_1^q,\ldots,d_N^q)}=\frac{1}{\sum_i\frac{1}{d_i^q}}\,. \tag{70}$$

The weights w_i that minimize the portfolio variance are inversely proportional to the corresponding asset variance d_i^q.

Noting

$$x_i = \frac{1}{d_i^q} \ , \tag{71}$$

the cumulants corresponding to the weights that minimize the variance are

$$c_{2r}^V(q) = C(r,q) \ \frac{\sum_i x_i^r}{\left(\sum_i x_i\right)^{2r}} \ . \tag{72}$$

4.2 Minimizing the higher-order cumulants

The weights w_i that minimize the cumulants $c_{2r}(q)$ given by (40) are

$$w_1 d_1^{\frac{qr}{2r-1}} = w_2 d_2^{\frac{qr}{2r-1}} = \ldots = w_N d_N^{\frac{qr}{2r-1}} = \frac{1}{\sum_i \frac{1}{d_i^{\frac{qr}{2r-1}}}} \ . \tag{73}$$

Note that these weights vary with the order r of the cumulant that is minimized. It is thus not possible to minimize simultaneously all the cumulants of order larger than two. This is in contrast to the result for the normalized cumulants discussed below. The weights that minimize the very large $r \to +\infty$ order cumulants approach asymptotically the values that minimize all the normalized cumulants of order larger than two, as given by (84) below. Thus, conclusions obtained for the normalized cumulants carry out for the cumulants in the limit $r \to \infty$.

The cumulants corresponding to (73) are

$$c_{2r}^{min}(q) = \frac{C(r,q)}{\left(\sum_i x_i^{\frac{r}{2r-1}}\right)^{2r-1}} \ , \tag{74}$$

where the x_i's are given by (71). By definition, the two expressions (74) and (72) coincide for $r = 1$, i.e. give the same minimum variance.

It is interesting to compare the values taken by the cumulants for the weights that minimize the variance with the values taken by the cumulants for the weights that minimize a cumulant of given order $2r$. For this, we study the ratio of the cumulants which is also the ratio of the normalize cumulants since the two expressions (74) and (72) coincide for $r = 1$:

$$\frac{c_{2r}^V(q)}{c_{2r}^{min}(q)} = \frac{\lambda_{2r}^V(q)}{\lambda_{2r}^{min}(q)} = \left(\sum_i X_i^r\right) \left(\sum_i X_i^{\frac{r}{2r-1}}\right)^{2r-1} \ , \tag{75}$$

where

$$X_i = \frac{x_i}{\sum_j x_j} \qquad \text{are normalized} \ \sum_{j=1}^N X_j = 1 \ , \tag{76}$$

and the x_i's are defined by (71).

Differentiating $c_{2r}^{V}(q)/c_{2r}^{min}(q)$ given by (75) with respect to one of the X_i's, we find that the derivative vanishes when all X_j's are equal to $1/N$. This corresponds to an extremum of all $c_{2r}^{V}(q)/c_{2r}^{min}(q)$ equal to 1. This extremum is in fact a minimum. To see this, we parameterize

$$X_i = \frac{1}{N}(1+\varepsilon_i) \ , \tag{77}$$

where the ε_i's are small and sums up to zero ($\sum_{i=1}^{N} \varepsilon_i = 0$). We then expand (75) in powers of ε_i's and find

$$\frac{c_{2r}^{V}(q)}{c_{2r}^{min}(q)} = 1 + \frac{1}{N} \ \frac{r(r-1)^2}{2r-1} \ \sum_{i=1}^{N} \varepsilon_i^2 \ . \tag{78}$$

Thus, for unequal X_i's, the ratios $c_{2r}^{V}(q)/c_{2r}^{min}(q)$ are generically larger than one. In words, the values of the cumulants for the weights that minimize the variance are higher than those for the weights that minimize a cumulant of given order $2r > 2$: minimizing "small" risk increases large risks.

4.3 Minimization of the excess kurtosis and higher normalized cumulants

Normalized cumulants provide a better measure of large risks than the (non-normalized) cumulants. The normalized cumulants are defined by

$$\lambda_{2m} = \frac{c_{2m}}{[c_2]^m} \ . \tag{79}$$

The fourth normalized cumulant λ_4 is often called the excess kurtosis κ. Recall that it is identically zero for Gaussian distributions.

The quantitative deviation of a distribution from normality is given by normalized cumulants $\frac{c_{2m}}{[c_2]^m}$. Indeed, the difference $P_>(z) - g(z)$ between the cumulative distribution function of the sum of N random variables and its Gaussian asymptotic value is given by [Gnedenko and Kolmogorov, 1954]

$$P_>(z) - g(z) \simeq \frac{\exp(-z^2/2)}{\sqrt{2\pi}} \left(\frac{Q_1(z)}{N^{1/2}} + \frac{Q_2(z)}{N} + \ldots + \frac{Q_k(z)}{N^{k/2}} + \ldots \right) \tag{80}$$

where $Q_k(x)$ are polynomials parameterized by the normalized cumulants λ_n of the distribution of the N random variables. For instance, the two first polynomials are

$$Q_1(x) = \frac{1}{6}\lambda_3(1-x^2), \tag{81}$$

and

$$Q_2(x) = \frac{1}{72}\lambda_3^2 x^5 + \frac{1}{8}(\frac{1}{3}\lambda_4 - \frac{10}{9}\lambda_3^2)x^3 + (\frac{5}{24}\lambda_3^2 - \frac{1}{8}\lambda_4)x. \tag{82}$$

A more straightforward way to recognize the role of the normalized cumulants is to work with the characteristic function defined in (34). Since k is the variable conjugate to δx and the natural scale for δx is the standard deviation $\sqrt{c_2}$, this means that the natural "dimensionless" variable is $\delta x/\sqrt{c_2}$, i.e. in the conjugate variable, it is $k\sqrt{c_2}$ that we define as $\tilde{k}$. In terms of $\tilde{k}$, $\hat{P}$ become

$$\hat{P}(\tilde{k}) = \exp\left(-\frac{\tilde{k}^2}{2} + \sum_{n=2}^{\infty} (-1)^n \frac{\lambda_{2n}}{(2n)!} \tilde{k}^{2n}\right) . \tag{83}$$

Using (40) to construct the normalized cumulants λ_{2m} defined by (79), we find that the asset weights w_i that minimize λ_{2m}, *irrespective* of the order $2m$, are given by

$$w_1 d_1^{q/2} = w_2 d_2^{q/2} = \ldots = w_N d_N^{q/2} = \frac{\sigma_N^{(N)}(d_1^{q/2}, \ldots, d_N^{q/2})}{\sigma_{N-1}^{(N)}(d_1^{q/2}, \ldots, d_N^{q/2})} = \frac{1}{\sum_i \frac{1}{d_i^{\frac{q}{2}}}} . \tag{84}$$

We recover (84) from the analysis of the tail of the portfolio distribution for $c < 1$ given in section (3.3.2). The natural criterion is to find the weights w_i that minimize the characteristic decay rate $\hat{\chi}$ of the tail of the portfolio distribution given by (47) :

$$\frac{\partial \hat{\chi}}{\partial w_j} = 0 , \qquad \text{for all } w_j' s . \tag{85}$$

After some calculation, we find that the weights that minimize $\hat{\chi}$ are exactly those given by (84), that minimize the high order normalized cumulants. This confirms that the normalized cumulants are the relevant measures of the tail of the portfolio distribution.

The expressions of the cumulants corresponding to the weights (84) are

$$c_{2m}^{(K)} = NC(m,q) \frac{1}{\left(\sum_i \frac{1}{d_i^{\frac{q}{2}}}\right)^{2m}} . \tag{86}$$

If all assets have the same nonlinear variance $d_i = d$, then $c_{2m}^{(K)} = c_{2m}^{(V)}$ for all orders $2m$. Thus, minimizing the variance minimizes all normalized cumulants at the same time. This special result is not true in general as we now show.

Using (76) and the definition (79) of the normalized cumulants, we get

$$\frac{\lambda_{2m}^{(K)}}{\lambda_{2m}^{(V)}} = \frac{1}{N^{m-1}} \frac{\left(\sum_i X_i^2\right)^m}{\left(\sum_j X_j^{2m}\right)} . \tag{87}$$

Using the Lagrange multiplier method to impose the normalization condition on the X_i's, we find that the values X_i's that *maximize* the $\frac{\lambda_{2m}^{(K)}}{\lambda_{2m}^{(V)}}$ for $m > 1$ are all equal to $1/N$, for which $\frac{\lambda_{2m}^{(K)}}{\lambda_{2m}^{(V)}} = 1$. For any other set of "nonlinear variances", the ratio is less than one. Thus, the weights that minimize the variance *increase* the higher normalized cumulants such as the excess kurtosis.

We now compare the cumulants of the portfolio with minimum variance to the cumulants of the portfolio with minimum normalized excess kurtosis, for forming the ratio

$$\frac{c_{2m}^{(K)}}{c_{2m}^{(V)}} = N \frac{\left(\sum_i \frac{1}{d_i^q}\right)^{2m}}{\left(\sum_j \frac{1}{d_j^{qm}}\right)\left(\sum_k \frac{1}{d_k^{\frac{q}{2}}}\right)^{2m}} \ . \tag{88}$$

With the definition (76,71), we get

$$\frac{c_{2m}^{(K)}}{c_{2m}^{(V)}} = N \frac{\left(\sum_i X_i^2\right)^{2m}}{\left(\sum_j X_j^{2m}\right)} \ . \tag{89}$$

In particular,

$$\frac{c_2^{(K)}}{c_2^{(V)}} = N \sum_i X_i^2 \ , \tag{90}$$

and

$$\frac{c_4^{(K)}}{c_4^{(V)}} = N \frac{\left(\sum_i X_i^2\right)^4}{\left(\sum_j X_j^4\right)} \ . \tag{91}$$

We find that the ratio $\frac{c_2^{(K)}}{c_2^{(V)}}$ is a minimum equal to 1 when all X_i's are equal to $1/N$. For any other set of nonlinear covariance matrix, the variance of the portfolio which minimize the excess kurtosis is always larger or equal to the minimum variance. This is expected by construction.

The ratio $\frac{c_4^{(K)}}{c_4^{(V)}}$ is also found to be a minimum equal to 1 when X_i's being equal to $1/N$. This result generalizes to higher orders. For any other set of nonlinear covariance matrix, we find that the higher order cumulants of the portfolio which minimize the excess kurtosis are always larger or equal to the cumulants of the portfolio that minimizes the variance.

4.4 Comparison between the excess kurtosis of the minimum-variance portfolio and the benchmark $w_i = 1/N$

The benchmark with $w_i = 1/N$ has the following cumulants

$$c_{2m}^{(1/N)} = \frac{C(m,q)}{N^{2m}} \sum_i d_i^{qm} \ . \tag{92}$$

We construct the ratio of the excess kurtosis of the portfolio with minimum variance to the excess kurtosis of the benchmark :

$$R \equiv \frac{\lambda_4^{(V)}}{\lambda_4^{(1/N)}} = \frac{\left(\sum_i x_i^2\right)\left(\sum_j \frac{1}{x_j}\right)^2}{\left(\sum_k x_k\right)^2 \left(\sum_l \frac{1}{x_l^2}\right)} , \tag{93}$$

where the x_i's are defined by (71).

For all x_i's equal, $R = 1$ as it should. Changing all x_i's into their inverse change the ratio R into its inverse $1/R$. This implies that the two situations are equally probable : either the excess kurtosis of the benchmark is smaller/larger than the excess kurtosis of the portfolio with minimum variance. Indeed, for each set of assets with x_i's for which $R < 1$, then the set of assets with inverse $1/x_i$'s gives $R > 1$. Finding the portfolio with minimum variance may thus either increase or decrease its excess kurtosis compared to that of the benchmark.

Notice that for $N = 2$, R is identically equal to unity for any possible choice of x_1 and x_2. A portfolio constituted of two uncorrelated assets is thus such that the excess kurtosis of the benchmark and of the optimized variance portfolio are the same. This results is not true anymore for $N > 2$ for which R is usually different from one.

It is also interesting to investigate the ratio of the cumulants

$$\frac{c_{2m}^{(V)}}{c_{2m}^{(1/N)}} = N^{2m} \frac{\sum_i X_i^m}{\sum_j \frac{1}{X_j^m}} . \tag{94}$$

Expanding the X_i's around the equal nonlinear variance case $X_i = \frac{1}{N}(1+\varepsilon_i)$ where the ε_i's are small and sums up to zero $\sum_{i=1}^N \varepsilon_i = 0$, we get

$$\frac{c_{2m}^{(V)}}{c_{2m}^{(1/N)}} = 1 - \frac{m}{N} \sum_{i=1}^N \varepsilon_i^2 , \tag{95}$$

up to second-order in ε_i's. The higher-order cumulants of the portfolio with minimum variance are smaller than those of the benchmark.

4.5 Synthesis

We have obtained the following results.

1. *Minimizing the variance versus minimizing the normalized higher-order cumulants* λ_{2m} : the portfolio with minimum variance may have smaller cumulants than the portfolio with minimum normalized cumulants but may have larger normalized cumulants (which characterize the large risks). Thus, minimizing small risks (the variance) may increase the large risks.

2. *Minimizing the variance versus minimizing the non-normalized higher-order cumulants* c_{2m} : the portfolio with minimum variance may have larger non-normalized and normalized cumulants of order larger than two than the portfolio with minimum non-normalized cumulants. Thus, minimizing small risks (the variance) may increase the large risks.

3. *Minimizing the variance versus the benchmark portfolio* : the portfolio with minimum variance has smaller non-normalized higher-order cumulants than the benchmarck portfolio, even if it may have smaller or larger higher-order normalized higher-order cumulants. This situation is similar to the first case, i.e. it is preferable to minimize the variance than taking the benchmark, even if the resulting portfolio distribution is less Gaussian.

4.6 The expected utility approach

The investor, starting the period with initial capital $W_0 > 0$, is assumed to have preferences that are rational in the von-Neumann-Morgenstern [1944] sense with respect to the end-of-period distribution of wealth $W_0 + \delta S$. His preferences are therefore representable by a utility function $u(W_0 + \delta S)$ determined by the wealth variation δS at the end-of-period. The expected utility theorem states that the investor's problem is to maximize $\mathrm{E}[u(W_0 + \delta S)]$, where $\mathrm{E}[x]$ denotes the expectation operator :

$$\mathrm{E}[u(W_0 + \delta S)] = \int_{-W_0}^{+\infty} d\delta S \; u(W_0 + \delta S) \; P_S(\delta S) \; . \qquad (96)$$

$u(W)$ has a positive first derivative (wealth is prefered) and a negative second derivative (risk aversion).

Consider first the case of a constant absolute measure of risk aversion $-U''/U' = a$, for which $U(W) = -\exp(-aW)$. In the limit of an initial wealth W_0 that is sufficiently large such that the probability of ruin occuring in a one-period (i.e. $\delta S < -W_0$) is negligible, expression (96) transforms into

$$\mathrm{E}[u(W_0 + \delta S)] = -e^{-aW_0} \; \hat{P}(k = ia) = -e^{-aW_0} \exp\left(\sum_{n=1}^{+\infty} \frac{c_{2n}}{(2n)!} \; a^{2n} \right) \; . \qquad (97)$$

Due to the negative sign in front of the expression, maximizing $\mathrm{E}[u(W_0 + \delta S)]$ is equivalent to minimizing $\sum_{n=1}^{+\infty} \frac{c_{2n}}{(2n)!} \; a^{2n}$, i.e. a weighted sum over the cumulants, each of them function of the asset weights w_i given by (40). For a small risk aversion $a << 1$, the sum in the exponential is essentially given by the first term $c_2 a^2/2$. In this limit of small risk aversion, maximizing the expected utility retrieved the standard procedure of minimizing the portfolio variance. However, for larger risk aversions, the higher-order cumulants bring in non-negligible contributions. It is straightforward to solve for the optimization with explicit analytical formulas.

Consider as another illustration the case where the utility is a member of the power (isoelastic) functions

$$u(W) = W^{\gamma} \qquad \text{with } \; 0 < \gamma < 1 \; . \qquad (98)$$

Then, assuming again that $|\delta S| < W_0$ for a single period (which is not at all restrictive an assumption in practice), we expand

$$u(W_0 + \delta S) = (W_0 + \delta S)^{\gamma} = W_0^{\gamma} (1 + \delta S/W_0)^{\gamma} = W_0^{\gamma} \sum_{j=0}^{+\infty} \frac{\gamma!}{(\gamma - j)! \; j!} \; (\delta S/W_0)^j \; . \qquad (99)$$

Putting (99) in (96) gives

$$\mathrm{E}[u(W_0 + \delta S)] = W_0^{\gamma} \sum_{r=0}^{+\infty} \frac{\gamma!}{(\gamma - 2r)! \; (2r)!} \; \mathrm{E}[(\delta S/W_0)^{2r}]$$

$$= W_0^\gamma \left(1 - \frac{\gamma(1-\gamma)}{2} \, \frac{c_2}{W_0^2} - \frac{\gamma(1-\gamma)(2-\gamma)(3-\gamma)}{4!} \frac{c_4 + 3c_2^2}{W_0^4} - ... \right) \tag{100}$$

where we have replaced the lower bound $-W_0$ by $-\infty$ in the integral and $\mathrm{E}[(\delta S/W_0)^2]$ by c_2/W_0^2 and $\mathrm{E}[(\delta S/W_0)^4]$ by $(c_4 + 3c_2^2)/W_0^4$. Recall that the cumulants c_{2n} are given by (40) with (41) and contain the dependence on the asset weights w_i. The parameters controlling the behavior of $\mathrm{E}[u(W_0 + \delta S)]$ are the exponent γ and the N ratios $d_i^{q_i/2}/W_0$ of the asset price standard deviations normalized by the initial wealth.

For large initial wealth W_0 (compared to the one-period price standard deviations), only the first few cumulants need to be considered in the sum as the others are weighted by a negligible factor. In this limit, the best portfolio retrieves the optimal variance portfolio, since maximizing $\mathrm{E}[u(W_0 + \delta S)]$ corresponds to minimizing the first term c_2/W_0^2 in the expansion (100).

In the other limit where the initial wealth W_0 is not large compared to the one-period price standard deviations, $\mathrm{E}[u(W_0 + \delta S)]$ receives non-negligible contributions from higher-order cumulants. The best portfolio is a weighted compromise between the different dimensions of risks provided by the different cumulants.

An important insight obtained by this analysis is that the optimal portfolio depends on the initial wealth W_0, everything else being equal. This results from the existence of the many different dimensions of risks provided by the different cumulants which are each weighted by the appropriate power of the initial wealth. In other words, the initial wealth quantifies the relative importance of the high-order cumulants.

Notice the parallel between the analysis of the two cases with constant absolute measure a of risk aversion and with a power (isoelastic) utility function : the relevant relative measure of the impact of the various cumulants (i.e. the importance of the tail of the distributions) is a for the first case and $1/W_0$ in the second case. It is interesting to retrieve the standard variance minimization approach in the limit of small risk aversion or large initial wealth. In other words, when you use the standard variance minimization method, you implicitely express (perhaps unwillingly) a small risk aversion. This is logical since you drop all information on the large-order cumulants that control the large risks. In contrast, the full treatment incorporating the higher cumulants as relevant measures of risks allows us to respond to a much larger spectrum of trader's sensitivity and risk aversions.

5 Empirical tests

5.1 $r \to y$ transformation by Eq.(16) : examples and statistical tests

5.1.1 Multivariate and marginal return distributions of six currencies

In Figures 1,2,4,3,5 are shown the empirical bivariate distributions ($N = 2$) for pairs made of the Japanese Yen (JPY) and one of five other currencies, all quoted in US dollars. Each point $(r_1(t), r_2(t))$ is defined as the daily annualized return

$$r_i(t) = 250 \, \ln \frac{s_i(t+1)}{s_i(t)} \tag{101}$$

where $s_i(t)$ is the price of currency i at day t. We use the index values $i = 2$ for the Japanese Yen and $i = 1$ for the other currencies which are respectively the Swiss Franc (CHF), the British Pound (UKP), the Russian

Ruble (RUR), the Canadian Dollar (CAD) and the Malaysian Ringgit (MYR). The time interval is from Jan. 1971 to Oct. 1998 except for the data with RUR where the time interval is from Jun. 1993 to Oct. 1998.

The contour lines are obtained from Eq.(23) and corresponds to different probability levels. A given level gives the probability that an event (r_1, r_2) falls inside the domain limited by the corresponding contour line. Three probability levels corresponding to 90%, 50% and 10% are shown for each pair except for the pair RUR-JPY where only the 50% and 10% levels are shown. In order to distinguish the contour lines from the data points, only one fourth of the data points (randomly chosen) are shown.

In these figures, we also present the price time series $s_i(t)$ (top and right panels) and the corresponding monovariate distributions $P(r_i(t))$. The scales of the axis for the bivariate and the abcissa of the monovariate distribution plots are chosen identical in order to highlight that monovariate distributions are obtained as projections of the multivariate distribution. The continuous lines correspond to the best fits of the Weibull pdf defined by (5) to the tail for large returns $\mathbf{r}$ of the empirical distributions shown as open circles. It can be compared to the best Gaussian fit shown as an inverted parabola in these semi-log plots. The values for the exponent c and the constant A are given in the figure captions. The fat tail nature of these pdf's is strongly apparent to the eye and is also reflected quantitatively by the fact that the exponents c's of the Weibull pdfs are all found of the order or less than one, while a Gaussian pdf corresponds to $c = 2$. These fits show that the distribution Eq.(5) provides an excellent representation for all six currencies.

It is quite clear that the bivariate distributions are very non-Gaussian, as can be seen from the shape of the countour lines at the 90%, 50% and 10% confidence levels. For Gaussian distributions, the contour lines would be given by the equation $\mathbf{r}'V^{-1}\mathbf{r} = C$, where C is a constant and V is the covariance matrix for the $\mathbf{r}$ variables. Depending on V, the contour lines would take the shape of

- circles for $V_{11} = V_{22}$ and $V_{21} \equiv V_{12} = 0$;
- ellipses with principal axis along the coordinates for $V_{11} \neq V_{22}$ and $V_{21} = 0$;
- ellipses with principal axis tilted with respect to the coordinates for $V_{21} \neq 0$ with the tilt and the eccentricity of the ellipses depending on V_{11}, V_{22}, V_{21}.
- In particular, if $V_{11} = V_{22}$ and $V_{21} \neq 0$, the ellipses are always tilted at ± 45 degrees with the ratio of the small over large principal axis given by $\sqrt{(1-|V_{12}|)/(1+|V_{12}|)}$. For perfect correlations $|V_{12}| = 1$, the ellipse becomes a segment of straight line as expected for the dependence of two perfectly correlated variables.

Consider first the 90% contour level for the CHF-JPY data Fig.1. It is apparent that the data is not described by an ellipse as prescribed by a Gaussian distribution, but rather the contour line takes the shape of a "bean". For the fairly large $\mathbf{r}$ events that fall outside the 90% level, events with different signs of r_1, r_2 are less likely to occur, or equivalently events which have same signs of r_1, r_2 are more likely to occur, compared to an ellipse described by the multivariate Gaussian distribution. At the 50% level, the same "bean"-like structure is found. Similar comments apply to the UKP-JPY pair. This "bean" structure is thus the signature of a stronger correlation of the sign of the returns, i.e. in the trend of the US $ against these other currencies, for relatively small returns compared to the large returns for which this correlation weakens.

For the RUR-JPY data, the limited statistics makes it hard to state anything decisive concerning the large $\mathbf{r}$ bivariate distribution, while the marginal distributions are clearly strongly non-Gaussian. However, the small

r data seems not incompatible with Gaussian behavior, both for the marginal and bivariate distributions. This illustrated in figure (6) that shows a close-up of the countour lines of the bivariate distributions. For small $|\mathbf{r}|(< 0.5\%)$, the contour lines in all five cases appear either as approximate circles or ellipses with principal axis along the coordinates, thereby suggesting *uncorrelated* multivariate Gaussian behavior for small changes of exchange rate.

The CAD-JPY contour lines shown in Fig.4 have a diamond like shape with principal axis along the coordinates. This signals that correlations are very weak and that, compared to an ellipse, the CAD-JPY data shows a higher probability for having events in which one of the returns is small. In other words, either one or the other exhibits a large move but rarely the two together.

The MYR-JPY data looks similar except that the diamond principal axis appear tilted and "twisted". The tilt reflects the fact that the MYR-JPY pairs are correlated, significantly more so than the CAD-JPY pair.

5.1.2 Multivariate Gaussian distributions of the transformed variables y_i

In Figures 7,8,9,10,11, we plot the bivariate distributions $\hat{P}(\mathbf{y})$ defined by Eq.(20), obtained from Figures 1,2,4,3,5 using the transformation Eq.(16). In each figure, we also construct the marginal distributions obtained by the transformation Eq.(16) performed on the empirical data points. The continuous lines represent the parabola (in this semi-log representation) corresponding to a Gaussian of unit variance. It is not surprising to find an excellent agreement between the empirical marginal distributions of the y variables and the Gaussian pdf, as the mapping (16) ensures this correspondence. We note however some discrepancy near $y = 0$, due to the large number of quotes with zero or very small absolute values of the returns and the limited accuracy of the data (the quotes are measured with a finite number of digits, in units of points).

The contour lines for the bivariate distributions of the y variables are defined as in Figures 1,2,4,3,5. Let us mention that, due to the phenomenon just mentioned of the abnormal behavior of $P(r)$ near $r = 0$, there is a high degeneracy of the points near $y = 0$, i.e. one single dot may correspond to many events near $y = 0$.

Since the transformation Eq.(16) ensures that $V_{11} \approx V_{22} \approx 1$, it is easy to show that the contour lines should be ellipses tilted at an angle of ± 45 degrees (for correlated/anticorrelated V_{12} respectively) with a ratio of the small over large pricipal axis of the ellipse given by $\sqrt{(1-|V_{12}|)(1+|V_{12}|)}$. Visual inspection of the Figures 7,8,10,9,11 confirms that the transformation (16), which is exact for each of the marginal distribution, provides an excellent representation of the bivariate distributions.

5.1.3 Goodness of the nonlinear transformation and statistical tests

The transformation Eq.(16) offers a new way for characterizing marginal distributions and their deviations from normality.

In Fig.12, we represent the transformation Eq.(16) for all six currencies. The r-variables are calculated from the empirical data. From each data point, the transformation Eq.(16) provides the corresponding y value. The negative returns have been folded back to the positive quadrant (by taking their absolute values). The double logarithmic representation (log-log plot) is useful to characterize a power law (6), corresponding to marginal distributions in r given by the Weibull pdf (5).

Two straight lines are drawn on the figures. The $y = r$ line corresponds to the exponent $c = 2$ in Eq.(5) and Eq.(6), which would qualify a Gaussian pdf. The other straight line shown for each currency describes the transformation law (6), where the exponents c have been determined independently by fitting the tails of the marginal distributions reported in Figures 1,2,4,3,5. Thus, the Weibull distributions Eq.(5) correspond to exact power laws Eq.(6), translating into straight lines of slope $c/2$ in the log-log representation of Fig.12. Without any adjustment of the exponent, we find an excellent consistency of the description in terms of the modified Weibull distributions for large returns.

For smaller returns, we observe in Figure 12 that the $y(r)$ curve bends down to become approximately parallel to the line $y = r$ suggesting again, in agreement what was found above from the contour lines of the bivariate distributions, that the distributions are approximately Gaussian for small absolute returns.

We now present a statistical test for the reliability and "goodness of fit" of the representation Eq.(20) of empirical multivariate distributions. We plot in thick line in Fig.13a-e the χ^2 cumulative distribution for $N = 2$ degrees of freedom versus the fraction of events as shown in Figures 7,8,9,10,11 within an ellipse of equation $\chi^2 = \mathbf{y}' V^{-1} \mathbf{y}$. If the empirical distributions strictly followed the pdf Eq.(20), one should have a straight line. The thin lines correspond to the 90% confidence levels obtained from 100 Monte Carlo simulations : 100 random realizations with the same number of points as the empirical data, taken from a Gaussian multivariate distribution, are transformed into the representation Eq.(20). The thick line is found close to the diagonal, implying that the empirical $\hat{P}(\mathbf{y})$ is very well approximated by a multivariate Gaussian. There is however a small but statistically significant departure from the diagonal, mostly in the upper parts of the figures, corresponding to the small returns. This seems to tell us that the correlation has a different structure for small and moderate returns compared to the larger returns, again in agreement with the Weibull (resp. Gaussian) representation of large (resp. small) returns. The largest deviation can be seen for the RUR-JPY data, which is to be expected since the statistics is the poorest for this set.

In Fig.13f, we construct a similar plot for the multivariate distributions with the $N = 6$ currencies taken all together. Since the empirical determination of the multivariate distribution requires us to keep only the quotes for all six currencies occurring simultaneously on the same day, this statistics for the multidimensional plot of Fig.13f is necessarily the poorest and is controlled by the poorest of all bivariate distributions which is the RUR-JPY set. Fig.13f shows that the method works well also with more than $N = 2$ assets.

We have checked for the stationarity of these results by dividing the total time window into three sub-windows of equal lengths and have performed the same analysis using Eq.(16). We find that the representation Eq.(20) tested in Fig.13 is extremely robust : no statistically significant differences are found between the three sub-windows. The stationarity of the nonlinear covariance matrix V is quite remarkable.

The same cannot be said with a represention in terms of the return r variables. Fig.14 illustrates this fact by constructing the same tests as shown in Fig.(13) but for the returns r instead of the transformed variables y. One can observe the dramatic departure from the diagonal, with an extremely large statistical significance, confirming the highly non-gaussian structure of the empirical multivariate distribution. This figure is presented to stress (1) how far from Gaussian are the empirical distributions and (2) how good is our transformation Eq.(16) in comparison. The transformed variable y is indeed the correct one to work with to get a robust and stationary representation.

Notice that, in Fig.14, the thick line does not approach the diagonal near abcissa values close to 1, as one could naively expect from our argument that the empirical distribution are not far from Gaussian for small returns. The reason is that the distribution may be close to Gaussian for small returns, but with a *different* covariance matrix. This interpretation is born out by the observed deviations from the diagonal in Fig.14 occurring for small y's.

The fact that the thick line in Fig.14 is below the diagonal for most of the range of χ^2 can be explained as follows. The correlation matrix of the returns is estimated over all returns, including the large realizations. Therefore, for a given value of χ^2, the corresponding ellipse is much larger than the ellipse that would be estimated from the covariance matrix (let us call it R) of the small returns. For small values of r (close to 1 on the abcissa of Fig.14), there are therefore more points inside, i.e. fewer points outside, than one should expect for an ellipse evaluated from the small returns. Another way to see this result is to recall that $y \approx r$ for small r as shown in Fig.12. Since R^{-1} is much smaller than V^{-1} (think of a matrix with only diagonal elements), χ^2 is therefore much smaller for small r-values than for small y-values. This implies that, for small r, the integral on the abcissa stays approximately constant as one decreases the fraction of events outside the ellipse on the ordinate. In comparison in Fig.14, since V^{-1} is much larger than R^{-1}, χ^2 is larger and the integral on the abscissa decreases with the same rate as the fraction of events outside the ellipse on the ordinate.

5.2 Variance and excess kurtosis

Consider one of the five portfolios with two currencies investing a fixed fraction p of its wealth W calculated in US$ in currency 1 and the remaining fraction $1-p$ in currency 2, where 1 and 2 refers to the five pairs of currencies analyzed in the previous figures. Using the historical time series, we construct numerically the time series for the portfolio value $W(t)$ from the recursion

$$W(t+1) = pW(t)s_1(t) + (1-p)W(t)s_2(t) \tag{102}$$

with p fixed. This expression ensures that the fraction of the wealth invested in a given currency is constant. We do not address here the issue of friction and transaction costs that would be involved in this dynamical reallocation but rather focus on the tests of the theory.

The annualized daily return r_W of $W(t)$ is defined by $r_W(t) = 250\ \ln \frac{W(t+1)}{W(t)}$. Fig. 15 shows the dependence as a function of p of the variance

$$c_2 \equiv \langle (r_W - \langle r_W \rangle)^2 \rangle \ , \tag{103}$$

of the excess kurtosis

$$\kappa \equiv \frac{c_4}{c_2^2} = \frac{\langle (r_W - \langle r_W \rangle)^4 \rangle}{\langle (r_W - \langle r_W \rangle)^2 \rangle^2} - 3 \ , \tag{104}$$

and of the sixth-order normalized cumulant

$$\lambda_6 \equiv \frac{c_6}{c_2^3} \ , \tag{105}$$

of the daily portfolio returns. The excess kurtosis and the sixth-order normalized cumulant quantify the deviation from a Gaussian distribution and provide measures for the degree of "fatness" of the tails, i.e. a measure of the "large" risks. Taking into account only the variance and the excess kurtosis and neglecting all higher order cumulants, a distribution can be approximated by the following expression valid for small excess kurtosis [Sornette, 1998]

$$P_(r_W) \simeq \exp\left[-\frac{(r_W - \langle r_W \rangle)^2}{2c_2} \left(1 - \frac{5\kappa}{12} \frac{(r_W - \langle r_W \rangle)^2}{c_2} \right) \right] \ . \tag{106}$$

The negative sign of the correction proportional to κ means that large deviations are more probable than extrapolated from the Gaussian approximation. For a typical fluctuation $|r_W - \langle r_W \rangle| \sim \sqrt{c_2}$, the relative size of the correction in the exponential is $\frac{5\kappa}{12}$. For large values of κ, this approximation (106) break down and the deviation from a Gaussian is much more dramatic.

The most interesting feature of Fig. 15 is that the weight that minimizes the variance does not correspond to the minimum of κ or of λ_6. This illustrates one of the main message of our work, derived in section 4 analytically for non-Gaussian distributions : minimizing the portfolio variance is not optimal with respect to the large risks. The strength of this effect depends on the relative shape of the distributions of the assets constituting the portfolio and their correlations. The strongest effect in Fig. 15 is found for the Malaysian-Yen pair, for which $p \approx 0.5$ minimizes the variance but gives an almost four-fold increase of the excess kurtosis compared to its minimum. Fig. 15 shows that it is possible to do much better and construct a portfolio which has not much more "small" risks (as measured by the variance) while having significantly smaller "large" risks (as measured by the excess kurtosis and six-order normalized cumulant). This is due to the fact that the excess kurtosis and the six-order normalized cumulant have in general a direct or inverted S-shape with a rather steep dependence or narrow well as a function of the asset weight p, while the variance exhibits comparatively a smoother and rather slower dependence. The investor will be well-inspired in using the additional information provided by the higher-order cumulants to control for the portfolio's risks.

It is also very interesting to observe that κ and λ_6 exhibit similar behaviors with minima occurring essentially for the same weight. This confirms a prediction of our theory, according to which the asset weights that minimize λ_{2m}, *irrespective* of the order $2m$ for $m > 1$, are given by Eq.(84). Notice that this prediction holds in principle only for uncorrelated assets, while the correlation coefficients are $\rho(y_1, y_2) = 0.57$ (CHF-JPY); 0.43 (UKP-JPY); -0.06 (RUR-JPY); 0.07 (CAN-JPY); 0.32 (MYR-JPY), where ρ is defined by (64). The case where the correlations are weak (RUR-JPY and CAN-JPY) exhibit the best agreement with our prediction. The agreement is still reasonable for the three other cases with larger correlations. The main lesson learned here is that such correlations are not very important for "fat tail" distributions ($c \leq 1.5$) in the determination of cumulants of large orders (in practice larger than two).

Fig.16a-b and d-e compare the empirical excess kurtosis (fat solid line) shown in Fig. 15 for the five portfolios to our theoretical prediction (40) with (41) (solid line). We use the result for uncorrelated assets as the coefficient of correlations are small for two of the five portfolios. In addition, we find that the empirical exponents c are much less than 2 for which the calibration of the exponents c plays a much more important role in the determination of the high-order cumulants than the correlation. We leave for a future paper the extension of our theory to the case of correlated assets with different exponents c. The exponents c_1 and c_2 are those determined in the fits of the pdf's tail, as given in Fig.12. There is thus *no* adjustable parameters in the comparison shown in Fig.16a-b and d-e. The thin solid lines and dashed lines plot the theoretical formula (40) for values of the exponents $\alpha_i \equiv c_i/2 \pm 0.05$, so as to provide uncertainty brackets for the comparison.

Fig.16c (RUR) shows the empirical excess kurtosis (fat solid line) and the theoretical prediction (40) with the fixed exponents c_1 and c_2 given in Fig.12. The thin solid line gives the predicted excess kurtosis for exponents $c_1 + 0.05, c_2 \pm 0.05$. A better agreement is observed for an exponent $c_1 + 0.05$ for the Russian currency slightly larger than determined from fitting the tail of its pdf. We interpret this result by the fact that this fat tail pdf embodies the effect of much higher-order cumulants, while the excess kurtosis is relatively a still rather low-order cumulant.

Fig.16f compares the empirical excess kurtosis (fat solid line) of the portfolio CHF-JPY (Fig.16a) to the prediction (65) with (63) for correlated assets with the fixed exponents $c_1 = c_2 = 2/3$. The thin solid line correspond to the empirical value $\rho(y1,y2) = 0.57$ while the dashed line is obtained from the same formula with $\rho(y1,y2) = 0$. The lesson we learn here by comparing the dashed and thin solid line together with the solid line of Fig.16a is that the existence or absence of a correlation for “fat tail” distributions is not very important for the determination of the excess kurtosis. Much more important is the correct determination of the tail exponents c of the Weibull distributions.

In summary, we find a good agreement between the empirical excess kurtosis and our prediction with (40) with fixed exponents c_1 and c_2 determined in Fig.12. The other theoretical curves provide the range of uncertainty in the determination of the excess kurtosis coming from measurement errors in the exponents c. In fact, the fits with the predicted excess kurtosis look amazingly good, because even minor effects like the double well structure in Fig.16c near $p = 0$ and in Fig.16d near $p = 1$ are in accordance with theory. These results are very consistent : a bad choice of c leads to a bad fit of the pdf tails of each asset constituting the portfolio and change completely the kurtosis of the portfolio pdf, as can be seen from its large sensitivity under relatively small variations of ± 0.1 of the exponent c.

The main point here is that the theory adequately identifies the set of portfolios which have small excess kurtosis and thus small ‘large risks’ and still reasonable variance (‘small risk’). We stress the importance of such precise analytical quantification to increase the robustess of risk estimators : historical data becomes notoriously unreliable for medium and large risks for lack of suitable statistics.

6 Conclusion

We have presented a novel and general methodology to deal with multivariate distributions with non-Gaussian fat tails and non-linear correlations. In a nutshell, our approach consists in projecting the marginal distributions onto Gaussian distributions, through highly nonlinear changes of variables. In turn, the covariance matrix of these nonlinear variables allows us to define a novel measure of dependence between assets, coined the “nonlinear covariance matrix”, which is specifically adapted to remain stable in the presence of non-gaussian structures. We have then presented the formulation of the corresponding portfolio theory which requires to perform non-Gaussian integrals in order to obtain the full distribution of portfolio returns. We have developed a systematic perturbution theory using the technology borrowed from particle physics of Feynmann diagrams to calculate the cumulants of the portfolio distributions, in the case where the marginal distributions are of the Weibull class. The main prediction is that minimizing the portfolio variance may in general increase the large risks quantified by the higher-order cumulants. Our detailed empirical tests on a panel of six currencies confirm the relevance of the Weibull description and allows us to make precise comparisons with our theoretical predictions. For “fat tail” distributions, we find in particular that the valid determination of large risks, as quantified by the excess kurtosis, are much more sensitive to the correct measurement of the Weibull exponent of each asset than to their correlation, which appears almost negligible.

Plenty of works remain to be done to explore further this approach.

- The case of assets with different exponents c have been treated only for uncorrelated assets and the corresponding problem of heterogeneous c’s in the correlated case is relevant for a precise comparison with empirical data. Furthermore, we have not studied assets with large exponents $c \geq 1.5$. The relevance of correlations increases with increasing c and we expect a precise determination of the correlation matrix to become more important as $c \to 2$.

- We have focused our analysis on the risk dimension of the problems by studying symmetric distributions, i.e. assets which are not expected to exhibit long-term trends. A natural and relevant extension of our theory is to treat the case where the mean return is non-zero and different from asset to asset.
- The next level of complexity is to have non-symmetric distributions, with variable Weibull exponents c and with correlations.
- The perturbation theory in terms of Feynmann diagrams can be used for other classes of distributions and it would be interesting to explore in details other potentially useful classes.
- We have assumed and found to be reasonably verified that the nonlinear covariance matrices are stationary. There is however no conceptual difficulty in generalizing and adapting the ARCH [Engle, 1982] and GARCH [Bollerslev et al., 1992] models of time-varying covariance to this formulation in terms of effective y variables.
- Our empirical tests have been performed on small portfolios with two and six assets, with the purpose of a pedagogical exposition and easier tests of our theory. It is worthwhile to extend our work to larger and more heterogeneous portfolios.

These goals all seem within reach and we intend to address these questions in future works.

7 APPENDIX A: Consistency condition for Elliptic distributions

Elliptic multivariate distributions $P(X)$ are defined by

$$P(X) = F\left((X - X^0)^T V^{-1} (X - X^0)\right) . \tag{107}$$

X^0 is the unit column vector of the average returns and V is a dependence matrix proportional to the covariance matrix when it exists. The function F is kept a priori arbitrary (non-negative and normalized). If F is an exponential, (107) retrieves the normal distribution and V becomes the covariance matrix.

The proposition that the results of the CAPM extends to elliptical distributions [Owen and Rabinovitch, 1983; Ingersoll, 1987] relies on the "consistency" condition, according to which any unconditional marginal distribution is of the form $F(V_{ii}^{-1}(X_i - X_i^0)^2)$ with the same function F. In particular, this leads to the fact that the distribution of the portfolio is a function solely of $W'VW$ (where W is the column vector of the asset weights) with a functional form *independent* of the weights W and the number N of assets in the portfolio.

Kano [1994] has shown that this is not true for most marginal distributions. For instance, for $F(x) = C\exp[-\sqrt{|x|}]$ where C is a normalizing constant, the unconditional variance of a single variable calculated using (107) is equal to $(N+1)/2$ times the variance obtained for a single variable using the same functional form. The result is thus not independent of N. This inconsistency implies that the overall shape of the distribution will change as a function of the asset weights in the portfolio. As a consequence, $W'\Sigma W$ is not the sole measure of the portfolio risks. The conclusions of Owen and Rabinovitch [1983] and Ingersoll [1987] thus apply only to a restricted class of elliptic distributions for which the consistency condition apply.

Let us show now explicity that

1. the dependence of the portfolio distribution on its wealth variation δS can be expressed solely in terms of the ratio $(\delta S)^2/W'VW$, but
2. the distribution itself has a functional form which, in general, is still dependent on the asset weights W constituting the portfolio.

For this, we write its variation during a unit time step as

$$\delta S(t) = \sum_{a=1}^{N} w_a \delta x^a(t) = W'X . \tag{108}$$

The density distribution $P(\delta S)$ can be written as

$$P(\delta S) = \int dX F\left((X - X^0)^T V^{-1} (X - X^0)\right) \delta(\delta S - W'X) , \tag{109}$$

where $\delta(x)$ is the Dirac distribution. To estimate this integral, we isolate one of the assets $x_1 = x_1^0 + y_1$ and write, using $Y = X - X^0$,

$$Y^T V^{-1} Y = V_{11}^{-1} y_1^2 + 2(v^T y) y_1 + y^T V^{-1} y , \tag{110}$$

where V_{ij}^{-1} is the element ij of the matrix V^{-1}, y is the unit column vector $(y_2, y_3, ..., y_N)'$ of dimension $N-1$, v is the unit column vector $(V_{21}^{-1}, V_{31}^{-1}, ..., V_{N1}^{-1})^T$ of dimension $N-1$, and V^{-1} is the square matrix of dimension $N-1$ by $N-1$ derived from V^{-1} by removing the first row and first column. The factor 2 in $2(v^T y)y_1$ comes from the symmetric structure of the matrix V^{-1}.

We can now express the condition $\delta(\delta S - W'X)$ in the integral (109) :

$$\frac{1}{w_1}\delta(y_1 - \frac{1}{w_1}(\delta S - W'X_0 - W'y)) \ , \tag{111}$$

where W is the unit column vector $(w_2, w_3, ..., w_N)'$ of dimension $N-1$. The integration over the variable y_1 cancels out the Dirac function and we obtain the argument of the function F under a quadratic form in the variables S and y. Using the identity

$$X'V^{-1}X + X'Y = \hat{X}'V^{-1}\hat{X} - \frac{1}{4}Y'VY \ , \tag{112}$$

where $\hat{X} = X + VY$, we obtain

$$P(\delta S) = \int d\hat{y} F\left(\hat{y}' M^{-1}\hat{y} + \frac{\delta S^2}{W^T V W}\right) \ , \tag{113}$$

where the integral is carried out over the space of vectors $\hat{y}$ of dimension $N-1$ and

$$M^{-1} \equiv V^{-1} - \frac{2}{w_1}(v - \frac{V_{11}^{-1}}{2w_1}W)W' \ . \tag{114}$$

We can finally write

$$P(\delta S) = F\left(\frac{\delta S^2}{W'VW}\right) \ , \tag{115}$$

where $F(x)$ is defined by (113).

This confirms that the typical volatility of the portfolio is controlled by the quasi-variance $W'VW$ as for the normal case. It is then natural to optimize the portfolio using this measure of the risk. However, it is clear that, for arbitrary functional forms of F, the specific functional form of the distribution $P(\delta S)$ is in general a function of the asset weights constituting the portfolio. Minimizing only $p^T V p$ may thus be insufficient because it may be linked to a dangerous deformation of $F(x)$ in the tail.

8 APPENDIX B: Derivation of the approximate stability in family of the Weibull distributions

Consider an asset with daily returns distributed according to (5) with $c < 1$. What is the distribution of returns over T days? It is convenient to use the representation of the distribution in terms of its cumulants defined by the derivatives of the logarithm of its characteristic function :

$$c_n = (-i)^n \frac{d^n}{dk^n} \log \hat{P}(k) \bigg|_{k=0} \ . \tag{116}$$

The characteristic function thus reads

$$\hat{P}(k) = \exp\{\sum_n^\infty \frac{c_n}{n!}(ik)^n\}. \tag{117}$$

Assuming the absence of correlations between successive daily returns, the distribution of returns over T days is the distribution of the sum of T independent random variables. Its cumulants are thus T times the cumulants of the distribution of the random variables [4]. Adapting the results derived in Appendix C to the case of a single asset with "nonlinear" variance $d = \chi^{\frac{2}{q}}$ according to (38) with $c = 2/q$, we obtain the expression of the cumulants of the returns over T days as

$$c_{2r}(T) = T c_{2r}(1) = T\ C(r,q)\ d^{rq}\ , \tag{118}$$

where $C(r,q)$ is a numerical factor given by (41) in the main text and in Appendix C. This expression (118) is valid for any real value q, i.e. describe the case of a Weibull exponential with arbitrary real exponent $c = 2/q$.

The deviation from the normal distribution is quantified by the normalized cumulants

$$\lambda_{2r}(T) = \frac{c_{2r}(T)}{[c_2(T)]^r} = \frac{C(r,q)}{[C(1,q)]^r}\ \frac{1}{T^{r-1}}\ . \tag{119}$$

The cumulants of order $2r$ larger than 2 decay to zero as T increases to infinity. This constitutes a signature of the central limit theory according to which the distribution of returns for the sum of $T \to \infty$ i.i.d. random variables tends to the normal law. In practice, the distribution of returns over a finite T is not exactly normal because the convergence to the normal law is rather slow. For instance, the excess kurtosis decays to zero only as $1/T$, starting from a rather large value $\kappa(T = 1) = 7.2$.

For finite $T > 1$, the distribution of returns over the time scale T may be approximated by a Weibull pdf with an apparent exponent $c_T = 2/q_T$ larger than the exponent c determined at the daily time scale. For this, we approximate the first cumulants by their expression for a Weibull distribution with adjustable nonlinear variance d_T and different exponent $c_T = 2/q_T$. This amounts to look for an approximate representation of $T\ C(r,q)\ d^{qr}$ by $C(r,q_T)\ d_T^{q_T r}$. Since there are only two variables d_T and q_T to optimize, they can be determined from two conditions, that we take to be the correspondence of the variance and excess kurtosis

$$T\ C(1,q)\ d^q = C(1,q_T)\ d_T^{q_T}\ , \tag{120}$$

[4]This stems from the fact that the distribution of the sum is a convolution integral and its Fourier transform is then the product of the Fourier transforms of each individual pdf's. From the definition (117), the result follows.

$$T\ C(2,q)\ d^{2q} = C(2,q_T)\ d_T^{2q_T}\ . \tag{121}$$

Eliminating d_T between the two equations gives q_T as the solution of

$$T = \frac{C(1,q_T)}{C(1,q)}\ \frac{C(2,q)}{C(2,q_T)}\ . \tag{122}$$

For a distribution of daily returns given by a Weibull distribution with exponent $c_1 = 2/3$ ($q = 3$), we find that the monthly ($T \approx 25$) returns are distributed according to an exponential ($c_{25} \approx 1$). This approximation illustrates the very slow convergence to the normal law since the value $c_{25} \approx 1$ is still very far from the asymptotic normal value $c_\infty = 2$.

The effective variables d_T and q_T can be determined by more global conditions such that the weighted sum of the square of the differences $T\ C(r,q)\ d^{qr} - C(r,q_T)\ d_T^{q_T r}$ be minimum over a certain set of r's, this set controlling how far in the tail the approximation is valid.

We test this idea by the following synthetic tests. Let us call $c_1 = 2/q_1 = 2/3$ the exponent of the Weibull pdf P_1 of the returns r_1 at the daily time scale. We construct the pdf P_T of the returns r_T over T days by taking the characteristic function of P_1 to the T-th power and then taking the inverse Fourier transform [5]. Let us now test whether P_T can be approximated by a Weibull distribution with an effective exponent $c_T = 2/q_T$ and determine its value as a function of T.

For this, we perform the change of variable $r_T \to y_T(r_T)$ given by (2) with (4), with a given choice for the exponent c_T and using (120,122) to get $\chi_T = d_T^q$. If the T-fold convolution distribution P_T of the Weibull distribution P_1 is approximately a Weibull, this change of variable should lead to an approximate gaussian with unit variance for the correct choice of c_T. We check the consistency of this program for $T = 1$ for which we do retrieve, as expected, an exact Gaussian with unit variance independent of c and χ.

Figs.17,18,19,20 plots the pdf's $P_T(y)$ as a function of $z \equiv y^2$ so that a Gaussian (in the y variable) is qualified as a straight line (dashed line on the plots). Thus, from the series of transformations, a straight line qualifies a Weibull distribution. We show the cases $T = 2,4,8$ and 20 for which the best c_T are respectively $c_2 = 0.73, c_4 = 0.80, c_8 = 0.90$ and $c_{20} \approx 1.05$. The other curves allow one to estimate the sensitivity of the representation of P_T in terms of a Weibull as a function of the choice of the exponent c_T. These simulations confirm convincingly our proposal that a Weibull distribution remains quasi-stable for many orders of convolutions, once the exponent c_T is correspondingly ajusted. We observe on Fig.17,18,19,20 that the Weibull representation is accurate over more than five orders of magnitude of the pdf P_T. Only for the largest time scale $T = 20$, we observe significant departure from the Weibull representation.

[5] We use the theorem that the Fourier transform of the convolution of two functions is the product of the Fourier transforms.

9 APPENDIX C: Calculation of the cumulants of $P(\delta S)$ in the diagonal case

The integral

$$I_i = \int_{-\infty}^{+\infty} du \; e^{-\frac{u^2}{2d_i} + ikw_i u^q} \; . \tag{123}$$

can be perturbatively expanded as

$$\begin{aligned} I_i &= 2\sum_{m=0}^{\infty} \frac{(-k^2 w_i^2)^m}{(2m!)} \int_0^{+\infty} du \; e^{-\frac{u^2}{2d_i}} u^{2qm} \qquad (124) \\ &= (2d_i)^{\frac{1}{2}} \sum_{m=0}^{\infty} \frac{(-2^q k^2 w_i^2 d_i^q)^m}{(2m!)} \int_0^{+\infty} dt \; e^{-t} t^{qm-\frac{1}{2}} \\ &= (2d_i)^{\frac{1}{2}} \sum_{m=0}^{\infty} \frac{(-2^q k^2 w_i^2 d_i^q)^m}{(2m!)} \Gamma\left(qm + \frac{1}{2}\right) \\ &= \sqrt{2\pi d_i} \sum_{m=0}^{\infty} \frac{(2qm-1)!!}{(2m)!} (w_i^2 d_i^q)^m (-k^2)^m . \end{aligned}$$

We recall the definition $\Gamma(x) = \int_0^\infty dt \, e^{-t} t^{x-1}$ and the property $\Gamma(n+1/2) = \pi^{1/2} \frac{(2n-1)!!}{2^n}$ when n is an integer and $(2n-1)!! = (2n-1)(2n-3)(2n-5)...5.3.1$.

The density $\hat{P}_S$ therefore becomes

$$\begin{aligned} \hat{P}_S(k) &= 1 - k^2 \frac{(2q-1)!!}{2!} \sum_i w_i^2 d_i^q \\ &\quad + k^4 \left[\frac{(4q-1)!!}{4!} \sum_i (w_i^2 d_i^q)^2 + \left(\frac{(2q-1)!!}{2!} \right)^2 \sum_{i<j} (w_i^2 d_i^q)(w_j^2 d_j^q) \right] + \cdots \end{aligned} \tag{125}$$

that can be exponentiated to extract the cumulants $c_m(q)$

$$\begin{aligned} \hat{P}_S(k) &= \exp\left[\sum_m \frac{c_m}{m!} (ik)^m \right] \qquad (126) \\ &= \exp\left[-k^2 \frac{(2q-1)!!}{2!} \sum_i (w_i^2 d_i^q) + k^4 \left[\frac{(4q-1)!!}{4!} - \frac{1}{2} \left(\frac{(2q-1)!!}{2!} \right)^2 \right] \sum_i (w_i^2 d_i^q)^2 + \cdots \right] . \end{aligned}$$

Pushing the calculation to the next orders, we get the sixth and eigth cumulants

$$\begin{aligned} \frac{c_6(q)}{6!} &= \left\{ \frac{(6q-1)!!}{6!} - \frac{(4q-1)!!}{4!} \frac{(2q-1)!!}{2!} + \frac{1}{3} \left[\frac{(2q-1)!!}{2!} \right]^3 \right\} \sum_i (p_i^2 d_i^q)^3 \qquad (127) \\ \frac{c_8(q)}{8!} &= \left\{ \frac{(8q-1)!!}{8!} - \frac{(6q-1)!!}{6!} \frac{(2q-1)!!}{2!} + \frac{(4q-1)!!}{4!} \left[\frac{(2q-1)!!}{2!} \right]^2 - \frac{1}{4} \left[\frac{(2q-1)!!}{2!} \right]^4 \right\} \times \\ &\quad \sum_i (w_i^2 d_i^q)^4 . \end{aligned}$$

The general cumulant is found by recurrence :

$$\frac{c_{2r}(q)}{(2r)!} = \left\{ \sum_{n=0}^{r-2} (-1)^n \frac{[2(r-n)q-1]!!}{(2r-2n)!} \left[\frac{(2q-1)!!}{2!} \right]^n - \frac{(-1)^r}{r} \left[\frac{(2q-1)!!}{2!} \right]^r \right\} \sum_i (w_i^2 d_i^q)^r. \tag{128}$$

Note that the above computation is valid even when q is real and the interaction term is proportional to $\mathrm{sign}(u_i)|u_i|^q$. In this case the interaction is still an odd function of u and the derivation goes through exactly with the same combinatorics as above. The result is

$$\frac{c_{2r}(q)}{(2r)!} = 2^{qr} \left\{ \sum_{n=0}^{r-2} (-1)^n \frac{\Gamma\left((r-n)q+\frac{1}{2}\right)}{(2r-2n)!\pi^{1/2}} \left[\frac{\Gamma\left(q+\frac{1}{2}\right)}{2!\pi^{1/2}} \right]^n - \frac{(-1)^r}{r} \left[\frac{\Gamma\left(q+\frac{1}{2}\right)}{2!\pi^{1/2}} \right]^r \right\} \sum_i (w_i^2 d_i^q)^r. \tag{129}$$

10 APPENDIX D: Generalization of the extreme deviation theorem of Frisch and Sornette [1997] to obtain the tail structure of $P(\delta S)$ in the diagonal case for $c > 1$

We start from the definition $\delta S = \sum_{i=1}^{N} w_i \delta x_i$ and the corresponding equation for its probability density function :

$$P_N(\delta S) = \underbrace{\int \cdots \int}_{N} e^{-\sum_{i=1}^{N} f_i(\delta x_i)} \, \delta\left(\delta S - \sum_{i=1}^{N} w_i \delta x_i\right) d\delta x_1 \cdots d\delta x_N \ , \tag{130}$$

where we have used the parameterization

$$P_i(\delta x_i) \equiv e^{-f_i(\delta x_i)} \ , \tag{131}$$

with

$$f_i(\delta x_i) = (\frac{\delta x_i}{\chi_i})^c \ , \quad \text{with} \ \ c > 1 \ . \tag{132}$$

All integrals in (130) are from $-\infty$ to $+\infty$. The delta function expresses the constraint on the sum.

We need the following conditions on the functions f_i (see [Frisch and Sornette, 1997] for precisions) :

(i) $f_i(\delta x_i) \to +\infty$ sufficiently fast to ensure the normalization of the pdf's.

(ii) $f_i''(\delta x_i) > 0$ (convexity), where f'' is the second derivative of f.

(iii) $\lim_{x\to\infty} x^2 f''(x) = +\infty$.

Under these assumptions, the leading-order expansion of $P_N(\delta S)$ for large δS and finite $N \geq 1$ is obtained by a generalization of the Laplace's method which here amounts to remark that the set of δx_i^*'s that maximize the integrant in (130) are solution of

$$f_i(\delta x_i^*) = C_N(\delta S) \ , \tag{133}$$

where $C_N(\delta S)$ is independent of i. In words, the leading behavior of $P_N(\delta S)$ is obtained by the set of δx_i's that occurs with the same probability. The δx_i^* obey

$$\sum_{i=1}^{N} w_i \delta x_i^* = \delta S \ . \tag{134}$$

Expanding $f_i(\delta x_i)$ around δx_i^* yields

$$f_i(\delta x_i) = f_i(\delta x_i^*) + a_i h_i + b_i h_i^2 + \ldots \tag{135}$$

where $a_i \equiv f_i'(\delta x_i^*)$, $b_i = \frac{1}{2} f_i''(\delta x_i^*)$ and $h_i \equiv \delta x_i - \delta x_i^*$ obey the condition

$$\sum_{i=1}^{N} w_i h_i = 0 \ . \tag{136}$$

We ignore the terms of order higher than two as they do not contribute to the leading order. We rewrite

$$a_i h_i + b_i h_i^2 = b_i (h_i + \frac{a_i}{2b_i})^2 - \frac{a_i^2}{4b_i} = \frac{b_i}{w_i^2} (H_i + \frac{a_i w_i}{2b_i})^2 - \frac{a_i^2}{4b_i} \ , \tag{137}$$

where the $H_i = w_i h_i$ verify $\sum_{i=1}^N H_i = 0$. Expression (130) then becomes

$$P_N(\delta S) = e^{-N\ C_N(\delta S)}\ e^{\sum_{i=1}^N \frac{a_i^2}{4b_i}} \underbrace{\int \cdots \int}_{N-1} e^{-\sum_{i=1}^N \frac{b_i}{w_i^2}(H_i + \frac{a_i w_i}{2b_i})^2}\ dH_1 \cdots dH_{N-1} \ . \tag{138}$$

The integral in (138) is evaluated by setting $y = \sum_{j=1}^N \frac{a_j w_j}{2b_j}$ and $\lambda_j = \frac{b_j}{w_j^2}$ in the identity

$$\underbrace{\int \cdots \int}_{N-1} e^{-\sum_{j=1}^{N-1} \lambda_j h_j^2 - \lambda_N (y - h_1 - \cdots - h_{N-1})^2}\ dh_1 \cdots dh_{N-1} = \pi^{\frac{N-1}{2}}\ \sqrt{\frac{\Lambda}{\prod_{j=1}^N \lambda_j}}\ e^{-\Lambda y^2} \ , \tag{139}$$

where Λ is defined by

$$\frac{1}{\Lambda} = \sum_{j=1}^N \frac{1}{\lambda_j} \ . \tag{140}$$

This identity is obtained by viewing y as the sum of the N Gaussian variables y_i's.

For the case (132) with $c > 1$, the condition (133) together with (134) yields

$$f_i(\delta x_i^*) = (\frac{\delta x_i^*}{\chi_i})^c = (\frac{\delta S}{\chi})^c \ , \quad \text{i.e.} \qquad x_i^* = \delta S\ \frac{\chi_i}{\chi} \ , \tag{141}$$

where

$$\chi = \sum_{j=1}^N w_j \chi_j = \sum_{j=1}^N w_j d_j^{\frac{q}{2}} \tag{142}$$

(with $c = 2/q$) and

$$\frac{a_i^2}{4b_i} = \frac{c}{2(c-1)}\ (\frac{\delta S}{\chi})^c \ . \tag{143}$$

We also have

$$y \equiv \sum_{j=1}^N \frac{a_j w_j}{2b_j} = \sum_{j=1}^N \frac{w_j f'(\delta x_j^*)}{f''(\delta x_j^*)} = \frac{\chi \delta S}{c-1} \ . \tag{144}$$

$$\frac{1}{\Lambda} = \frac{c(c-1)}{2X^2} \left(\frac{\delta S}{\chi}\right)^{c-2} \ , \tag{145}$$

where

$$X^2 \equiv \sum_{j=1}^N w_j^2 \chi_j^2 \ . \tag{146}$$

This yields the contribution

$$e^{-\Lambda y^2} = \exp\left(-\frac{c}{2(c-1)} \frac{\chi^2}{X^2} (\frac{\delta S}{\chi})^c\right) \ . \tag{147}$$

Regrouping all terms in (138) leads to

$$P_N(\delta S) = \pi^{\frac{N-1}{2}} \; \frac{1}{X \; \prod_{j=1}^{N} w_j \chi_j} \left[\frac{2}{c(c-1)} (\frac{\delta S}{\chi})^{2-c} \right]^{\frac{N-1}{2}} \exp\left(-\frac{N}{2(c-1)} (\frac{\delta S}{\hat{\chi}})^c \right) , \tag{148}$$

where

$$\hat{\chi}^c \equiv \frac{\chi^c}{c\frac{\chi^2}{NX^2} - (2-c)} = \frac{(\sum_{j=1}^{N} w_j \chi_j)^c}{c\frac{(\sum_{j=1}^{N} w_j \chi_j)^2}{N\sum_{j=1}^{N} w_j^2 \chi_j^2} - (2-c)} \; . \tag{149}$$

These results are valid for $c > 1$. The extreme tail of the portfolio wealth distribution is thus controlled completely by $\hat{\chi}^c$ which is its characteristic decay value.

Note that $\frac{\chi^2}{NX^2} = 1$ for identical assets $\chi_i = \chi$ when all weights p_i are equal to $1/N$. In this case, the exponential term in (148) simplifies into

$$P_N(\delta S) \sim \exp\left(-N (\frac{\delta S}{\chi})^c \right) \; . \tag{150}$$

11 APPENDIX E: Computation of the characteristic function defined by eq.(151)

The characteristic function of the distribution $P_S(\delta S)$ of portfolio wealth variations for correlated assets with Weibull distributions is given by

$$\hat{P}_S(k) = \frac{1}{(2\pi)^{N/2} \det V^{1/2}} \prod_{i=1}^{N} \left(\int du_i \right) e^{-\frac{1}{2} U'V^{-1}U + ik \sum_{i=1}^{N} w_i u_i^q} . \tag{151}$$

We recall the definition of the functional generator [Sornette, 1998]

$$\hat{P}_S^q(k, J_i) = \frac{1}{(2\pi)^{N/2} \det V^{1/2}} \int \left(\prod_i^N du_i \right) e^{-\frac{1}{2} uV^{-1}u + ik \sum_i w_i u_i^q + \sum_i J_i u_i} . \tag{152}$$

When the integral is a Gaussian ($k = 0$), we get

$$\begin{aligned} \hat{P}_S^q(0,0) &= 1 \\ \hat{P}_S^q(0,J_i) &= e^{\frac{1}{2} JVJ} . \end{aligned} \tag{153}$$

With the property

$$f\left(\frac{\delta}{\delta J_i}\right) \int \left(\prod_i^N du_i \right) e^{-\frac{1}{2} UV^{-1}U + \sum_i J_i u_i} = \int \left(\prod_i^N du_i \right) e^{-\frac{1}{2} UV^{-1}U + \sum_i J_i u_i} f(u_i) , \tag{154}$$

we can formally express the characteristic function as

$$\hat{P}_S^q(k) = e^{ik \sum_i p_i \frac{\delta^q}{\delta J_i^q}} e^{\frac{1}{2} JVJ} \Big|_{J_i = 0} . \tag{155}$$

We first consider the case $q = 3$. The first non-vanishing perturbative contribution for this case is obtained by expanding the formal expression above up to second order in k

$$\hat{P}_S^3(k) = \left[1 + ik \sum_i w_i \frac{\delta^3}{\delta J_i^3} - \frac{k^2}{2} \sum_{i,j} w_i w_j \frac{\delta^3}{\delta J_j^3} \frac{\delta^3}{\delta J_i^3} \right] e^{\frac{1}{2} JVJ} \Big|_{J=0} . \tag{156}$$

The first order vanishes because

$$\frac{\delta^3}{\delta J_i^3} e^{\frac{1}{2} JVJ} = e^{\frac{1}{2} JVJ} \left\{ (VJ)_i^3 + 3V_{ii} (VJ)_i \right\} \tag{157}$$

is zero when $J = 0$. [6] The second order contribution comes from

$$\begin{aligned} \frac{\delta^3}{\delta J_j^3} \frac{\delta^3}{\delta J_i^3} e^{\frac{1}{2} JVJ} &= e^{\frac{1}{2} JVJ} \times \\ & \left\{ (VJ)_j^3 (VJ)_i^3 + 9V_{ij} (VJ)_j^2 (VJ)_i^2 + 3V_{ii} (VJ)_j^3 (VJ)_i + 3V_{jj} (VJ)_j (VJ)_i^3 + \right. \\ & 9V_{ij} V_{ii} (VJ)_j^2 + 9(2V_{ij}^2 + V_{ii} V_{jj}) (VJ)_j (VJ)_j + 9V_{ij} V_{jj} (VJ)_i^2 + \\ & \left. 6V_{ij}^3 + 9V_{ii} V_{ij} V_{jj} \right\} . \end{aligned} \tag{158}$$

[6] We adopt here the compact notation $JVJ = \sum_{ij} J_i V_{ij} J_j$ and $(VJ)_i = \sum_l V_{il} J_l$.

By putting to zero the source term J, this expression leads to

$$\hat{P}_S^3(k) = 1 - \frac{k^2}{2} \sum_{ij} \left(6 w_i \left(V_{ij}\right)^3 w_j + 9 w_i V_{ii} V_{ij} V_{jj} w_j \right) \ . \tag{159}$$

This result can be usefully represented with diagrams in the following way. Let us associate to each factor V_{ij} the propagator diagram and to each factor $i g_3 w_j$ the vertex diagram as shown in Fig.21, where we have defined the coupling constant

$$g_3 = 3! k \ . \tag{160}$$

The two contributions in (159) can thus be represented by propagators connecting vertices as in Fig.22. Having defined the coupling constant g_3 in (160) allows us to interpret easily the coefficient in front of each diagram corresponding to the two terms in the expression

$$\hat{P}_S^3(k) = 1 - g_3^2 \left(\frac{1}{3!2} \sum_{ij} w_i \left(V_{ij}\right)^3 w_j + \frac{1}{2^3} \sum_{ij} w_i V_{ii} V_{ij} V_{jj} w_j \right) \ . \tag{161}$$

The coefficient in front of each term (diagram) is equal to $1/S$ where S is the symmetry factor of each diagram with respect to permutation of lines and vertices. The first diagram has a symmetry under exchange of 3 propagators and 2 vertices, therefore we have $S = 3! \times 2!$. The second diagram has a symmetry under exchange of the 2 lines forming the loop of each vertex, giving a contribution of $2! \times 2!$. It also has the symmetry under permutation of the two vertices. The total symmetry factor is therefore $S = 2! \times 2! \times 2!$.

A systematic way to keep under control the symmetry factor is to compute it diagramatically as follows. The second-order derivative operator with respect to J_i and J_j is represented by the two vertices in the left hand side of Fig.23.

The J-independent term is given by pairwise combining each leg of each vertex in all the possible topologically inequivalent way, taking into account the multiplicity of each configuration. Fig.23 shows how to proceed. The coefficient in front of the vertices of the left hand side comes from the perturbative expansion.

As a first step, let us consider the first leg of the first vertex. We can either contract it with another leg (two possibile contractions) of the same vertex or with a leg of the second vertex (three possible contractions). This is summarized in the first equality of Fig.23. The first contribution of the second equality is the result of the contraction with multiplicity 3 of the residual leg of the first vertex of the first diagram of the line above. Considering the possible contractions of the second contribution of the first equality generate the other two diagrams. At the end of the combining procedure, we have two inequivalent diagrams, with multiplicities 3! and 3^2. The resulting coefficient in front of each diagram can now be interpreted in terms of the symmetries of the diagram as anticipated.

These rules can be generalized to higher orders and other "interaction" terms, i.e. $\sum_i p_i u_i^q$ for general value of q or even generic functions $f(u_i)$ which admit expansions in power series of u_i. As an example, we give the first order correction to the characteristic function for generic value of q odd.

The perturbative expansion is

$$\hat{P}_S^q(k) = \left. \left[1 + i \frac{g_q}{q!} \sum_i w_i \frac{\delta^q}{\delta J_i^q} + \frac{1}{2} \left(\frac{i g_q}{q!} \right)^2 \sum_{i,j} w_i w_j \frac{\delta^q}{\delta J_j^q} \frac{\delta^q}{\delta J_i^q} \right] e^{\frac{1}{2} JVJ} \right|_{J=0} \tag{162}$$

where we have defined the auxiliary coupling constant $g_q = q!k$. Now, instead of taking explicitely the derivatives and keep the term independent on J as the result, let us use the diagrammatic formalism. At this order, we have two vertices each with q legs, representing the partial derivatives with respect to J_i and J_j and to which we associate the value $ig_q w_i$ and $ig_q w_j$ respectively. Now, let us combine pairwise each leg of the two vertices to form all the possible topologically inequivalent diagrams. We have collected in Fig.24 the sequence of all the diagrams. Each diagram is characterized by the number l of loops on each vertex and the number $q-2l$ of lines connecting the two vertices giving therefore a contribution

$$(ig_q)^2 \frac{1}{S_l} \sum_{i,j} w_i \, (V_{ii})^l \, (V_{ij})^{q-2l} \, (V_{jj})^l \, w_j \; , \tag{163}$$

where each loop around vertex i contributes to a factor V_{ii} and each propagator connecting the vertices i and j gives a factor V_{ij}. It is now easy to determine the symmetry factor: we have l symmetries under the exchange of the two lines of each loop at each vertex $((2!)^l \times (2!)^l)$, the symmetry of the l loops at each vertex $((l!) \times (l!))$, one symmetry under the exchange of the $q-2l$ internal propagators $((2q-l)!)$, and the reflection symmetry under the exchange of the two vertices $(2!)$. The total factor is therefore $S_n = (q-2l)!(2!)^{2l+1}(l!)^2$ and the characteristic function up to second order in k reads

$$\hat{P}^q_S(k) = 1 - g_q^2 \sum_{l=0}^{(q-1)/2} \frac{1}{(q-2l)!} \frac{1}{(2!)^{2l+1}} \frac{1}{(l!)^2} \sum_{i,j} w_i \, (V_{ii})^l \, (V_{ij})^{q-2l} \, (V_{jj})^l \, w_j \; . \tag{164}$$

The diagrammatic expansion becomes very useful for calculating the higher cumulants. A well known result of diagrammatic perturbation theory tells us that neglecting the disconnected diagrams, which do not occur at second order, corresponds to compute the logarithm of characteristic function. Therefore, the set of connected diagrams at m-th order give us directly the m-th cumulant coefficient c_m.

Let us give as an explicit example the computation of the fourth cumulant $c_4(3)$. In fig.25, the result of the contraction procedure is shown, where we keep only the connected diagrams. It is explicitely

$$\begin{aligned} c_4(3) \;\; = \;\; & 4!(3!)^4 \sum_{i_1,i_2,i_3,i_4} w_{i_1} w_{i_2} w_{i_3} w_{i_4} \left\{ \frac{1}{2^4} V^2_{i_1 i_2} V_{i_1 i_3} V_{i_2 i_4} V_{i_3 i_3} V_{i_4 i_4} + \right. \\ & \frac{1}{2^3} V^2_{i_1 i_2} V_{i_1 i_3} V_{i_2 i_3} V_{i_3 i_4} V_{i_4 i_4} + \frac{1}{2^4} V^2_{i_1 i_2} V_{i_1 i_3} V_{i_2 i_4} V^2_{i_3 i_4} + \\ & \left. \frac{1}{3!2^3} V_{i_1 i_2} V_{i_1 i_3} V_{i_1 i_4} V_{i_2 i_2} V_{i_3 i_3} V_{i_4 i_4} + \frac{1}{4!} V_{i_1 i_2} V_{i_1 i_3} V_{i_1 i_4} V_{i_2 i_3} V_{i_2 i_4} V_{i_3 i_4} \right\} \; . \end{aligned} \tag{165}$$

We can thus formally generalize the result for the m-th cumulant as

$$c_m(q) = m!(q!)^m \sum_{i_1,\ldots,i_m} w_{i_1} \cdots w_{i_m} \sum_{G_m(q)} \frac{1}{S(\{l_r\},\{n_{rs}\})} \prod_{r=1}^{m} (V_{i_r i_r})^{l_r} \prod_{r<s=1}^{m} (V_{i_r i_s})^{n_{rs}} \tag{166}$$

where $G_m(q)$ is the set of all the topologically inequivalent connected diagrams with m vertices of q legs. Each diagram is characterized by the number of loops at each vertex, $\{l_r\}$, and the number of lines connecting each couple of vertices, $\{n_{rs} = n_{sr}\}$. These numbers have to satisfy the constraints

$$2l_r + \frac{1}{2} \sum_{s \neq r} n_{rs} = q \; , \tag{167}$$

which embody the fact that each vertex has q legs. The symmetry factor S is obviously the most difficult part to determine. The safest procedure is to compute it with the contraction rule case by case. For each diagram, it is of the form

$$S(\{l_r\},\{n_{rs}\}) = (2!)^{\sum_{r=1}^{m} l_r} \prod_{r=1}^{m} l_r! \prod_{r<s=1}^{m} n_{rs}! S_v(\{l_r\},\{n_{rs}\}) \ . \tag{168}$$

Each factorial comes from the various symmetries under the exchange of the propagators, and we have isolated the contribution of the residual symmetries of the diagram under exchange of the vertices. Fig.26 summarizes the vertex symmetry factors for the diagrams contributing to $c_4(3)$.

REFERENCES:

R. Arad Wiener, The implications of a long-tailed distribution structure to portfolio selection and capital asset pricing, PhD thesis, Princeton University (1975).

V. Bawa, E.J. Elton and M.J. Gruber, Journal of Finance, vol. XXXIV, 1041-1047 (1979).

T. Bollerslev, R.Y. Chous and K.F. Kroner, ARCH modeling in Finance - A review of the theory and empirical evidence, J. Econometrics 52, 5-59 (1992).

J.-P. Bouchaud, D. Sornette, C. Walter and J.-P. Aguilar, Taming large events: Optimal portfolio theory for strongly fluctuating assets, International Journal of Theoretical and Applied Finance 1, 25-41 (1998).

S. Cambanis, S. Huang and G. Simons, On the theory of elliptically contoured distributions, Journal of Multivariate Analysis 11, 368-385 (1981).

P. Embrechts, C. Klüppelberg and T. Mikosh, Modelling extremal events (Springer-Verlag, Applications of Mathematics 33, 1997).

R.F. Engle, Econometrica 50, 987 (1982).

E. Fama, Management Science 11, 404-419 (1965).

U. Frisch and D. Sornette, Extreme deviations and applications, J. Phys. I France 7, 1155-1171 (1997).

C.C. Géczy, Some generalized tests of mean-variance efficiency and performance, working paper

B.V. Gnedenko and Kolmogorov, A.N., Limit distributions for sum of independent random variables, Addison Wesley, Reading MA (1954).

E.J. Gumbel, *Statistics of extremes* (Columbia University Press, New York, 1960).

M.J.R. Healy, Matrices for statistics (Clarendon Press, Oxford, 1986).

J.E. Ingersoll, Jr., Theory of financial decision making (Totowa, N.J. : Rowman & Littlefield, 1987).

B. Jorgensen, Exponential dispersion models, J. R. Statist. Soc. B 49 (2), 127-162 (1987).

Y. Kano, Consistency property of elliptical probability density functions, Journal of Multivariate Analysis 51, 139-147 (1994).

D. Karlen, Using projections and correlations to approximate probability distributions, Computer in Physics 12, 380-384 (1998).

J. Laherrère and D. Sornette, Stretched exponential distributions in Nature and Economy: "Fat tails" with characteristic scales, European Physical Journal B 2, 525-539 (1998)

P. Lévy, Théorie de l'addition des variables aléatoires (Gauthier Villars, Paris, 1937-1954).

R. Litterman and K. Winkelmann, Estimating covariance matrices, Risk Management Series, Goldman Sachs (1998).

B.B. Mandelbrot, Journal of Business, 36, 394 (1963)

B.B. Mandelbrot, Fractals and scaling in finance : discontinuity, concentration, risk : selecta volume E (New

York : Springer, 1997).

H. Markovitz, Portfolio selection : Efficient diversification of investments (John Wiley and Sons, New York, 1959).

R. C. Merton, Continuous-time finance, (Blackwell, Cambridge,1990).

S. Mittnik and S.T. Rachev, Reply to comments on 'Modelling asset returns with alternative stable distributions' and some extensions, Econometric Reviews 12, 347-389 (1993).

J. Owen and R. Rabinovitch, On the class of elliptical distributions and their applications to the theory of portfolio choice, The Journal of Finance 38, 745-752 (1983).

C. R. Rao, Linear statistical inference and its applications, 2d ed. (New York, Wiley, 1973).

P.A. Samuelson, J. Financial and Quantitative Analysis Vol.2, pp. 107-122 (1967).

D. Sornette, Large deviations and portfolio optimization, Physica A 256, 251-283 (1998).

J. von-Neumann and O. Morgenstern, Theory of games and economic behavior (Princeton, Princeton University Press, 1944)

M. Veltman, Diagrammatica, the path to Feynman diagrams (Cambridge University Press, 1995).

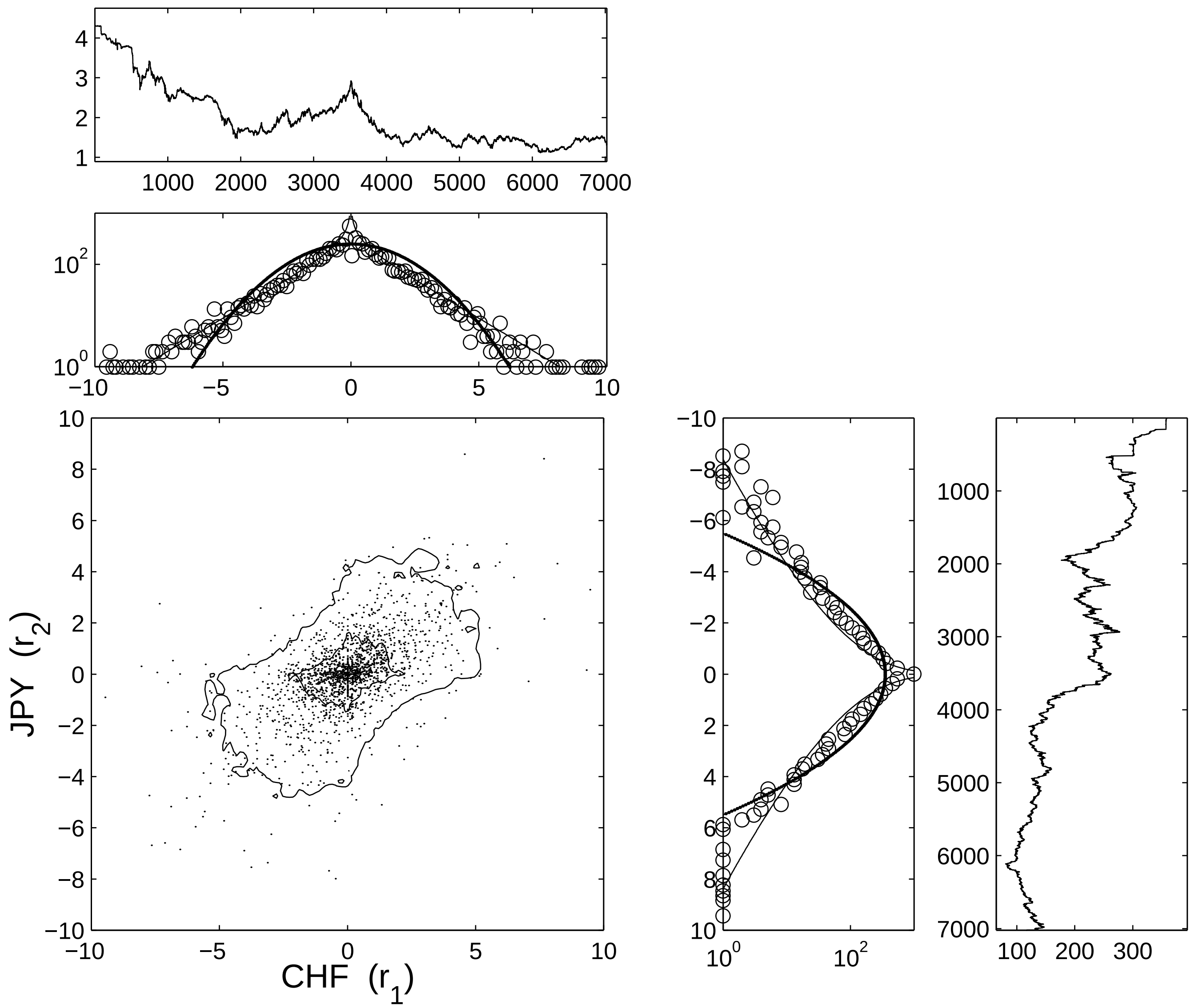

Figure 1: Bivariate distribution of the daily annualized returns of the CHF in US \$ ($i = 1$) and of JPY in US \$ ($i = 2$) for the time interval from Jan. 1971 to Oct. 1998. One fourth of the data points are represented for clarity of the figure. The contour lines define the probability confidence level of 90% (outer line), 50% and 10%. Also shown are the time series and the marginal distributions in the panels at the top and on the side. The parameters for the fit of the marginal pdf's are: CHF in US \$: $A_1 = 250, c_1 = 1.14, r_{01} = 2.13$ and JPY in US \$: $A_2 = 350, c_2 = 0.8, r_{02} = 1.25$.

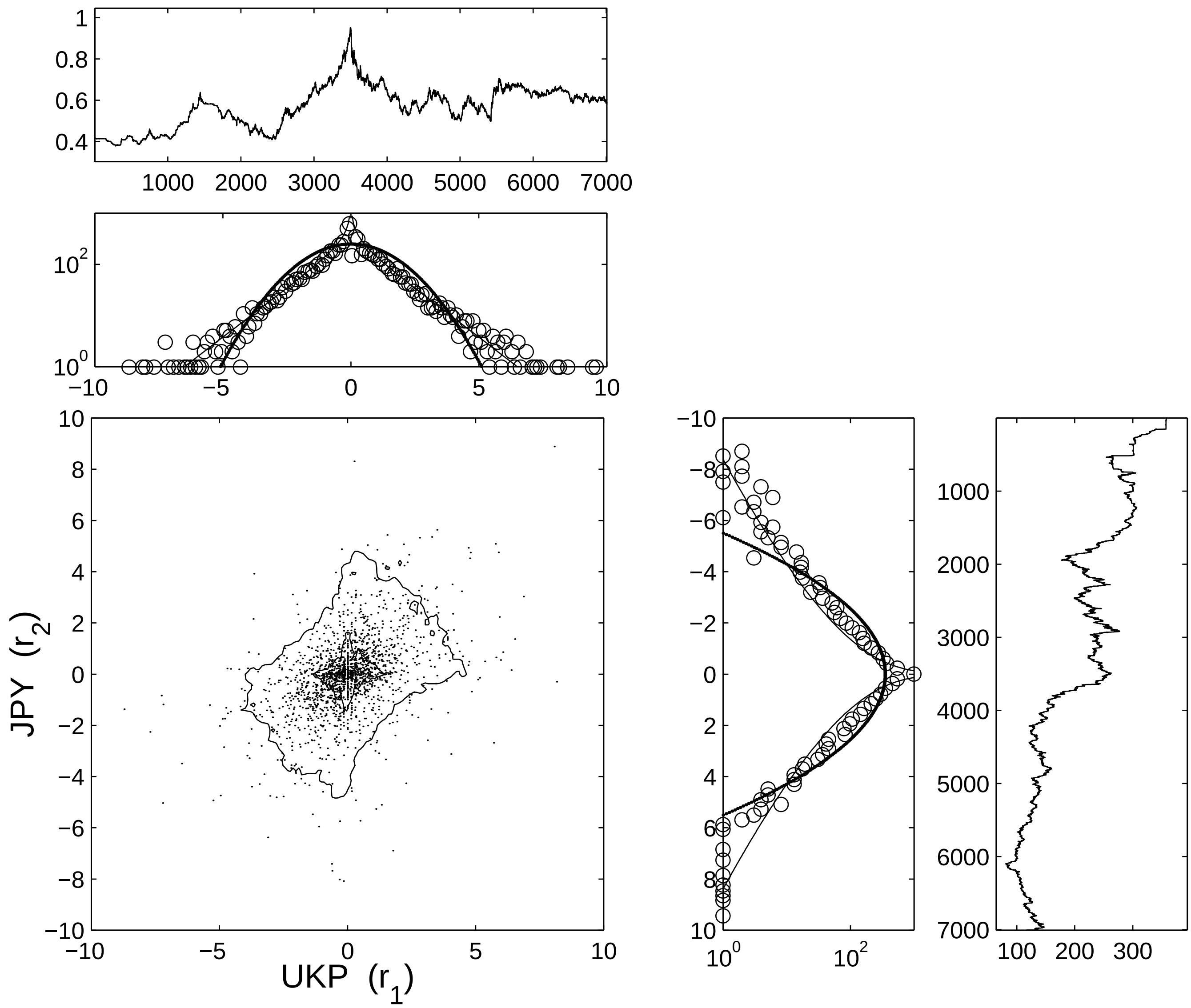

Figure 2: Bivariate distribution of the daily annualized returns of the UKP in US \$ ($i = 1$) and of JPY in US \$ ($i = 2$) for the time interval from Jan. 1971 to Oct. 1998. One fourth of the data points are represented for clarity of the figure. The contour lines define the probability confidence level of 90% (outer line), 50% and 10%. Also shown are the time series and the marginal distributions in the panels at the top and on the side. The parameters for the fit of the marginal pdf's are: UKP in US \$: $A_1 = 250, c_1 = 1.14, r_{01} = 1.67$ and JPY in US \$: $A_2 = 350, c_2 = 0.8, r_{02} = 1.25$.

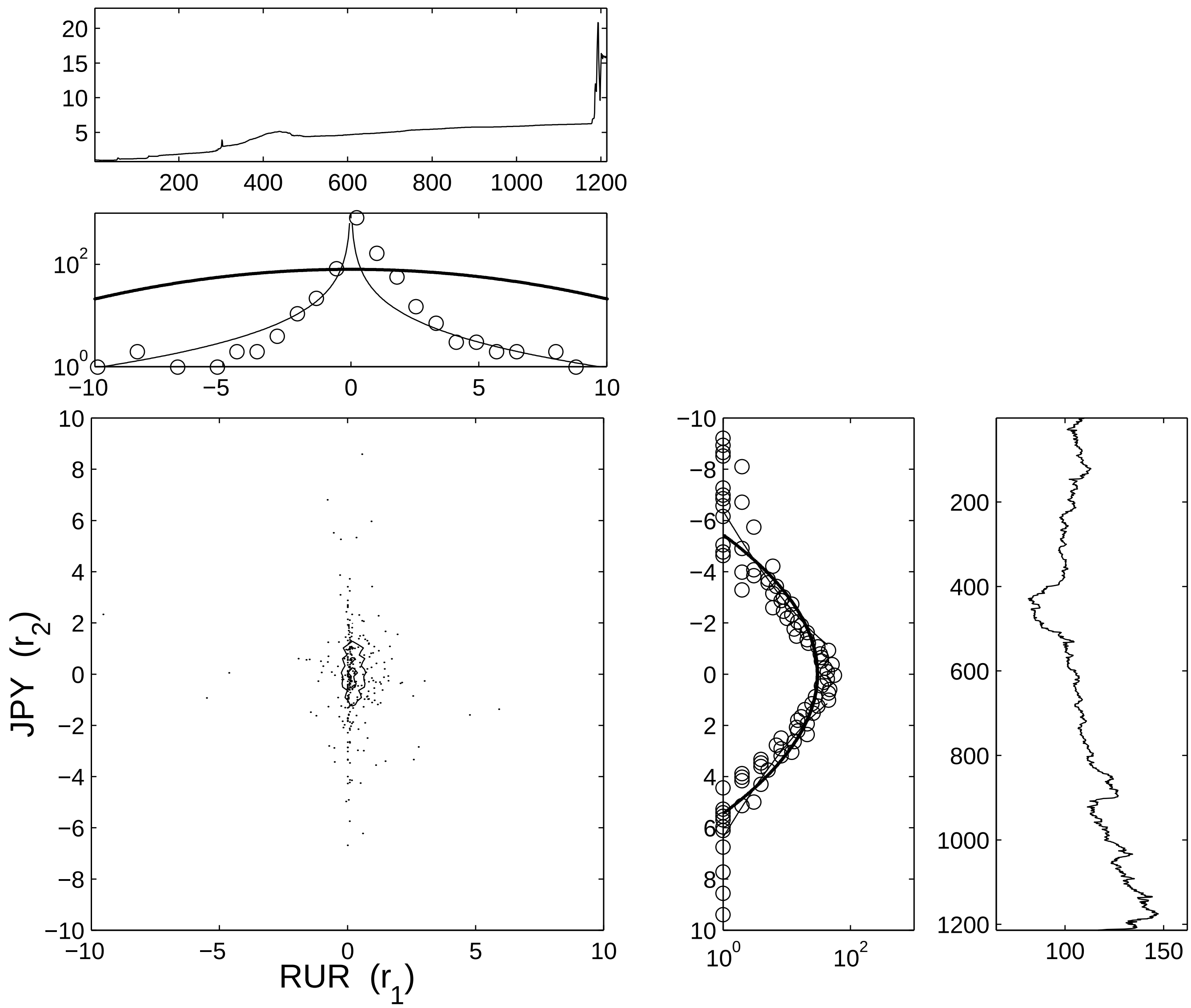

Figure 3: Bivariate distribution of the daily annualized returns of the RUR in US \$ ($i = 1$) and of JPY in US \$ ($i = 2$) for the time interval from Jun. 1993 to Oct. 1998. One fourth of the data points are represented for clarity of the figure. The contour lines define the probability confidence level of 50% (outer line) and 10%. Also shown are the time series and the marginal distributions in the panels at the top and on the side. The parameters for the fit of the marginal pdf's are: RUR in US \$: $A_1 = 80, c_1 = 0.38, r_{01} = 0.83$ and JPY in US \$ $A_2 = 120, c_2 = 0.8, r_{02} = 1.25$.

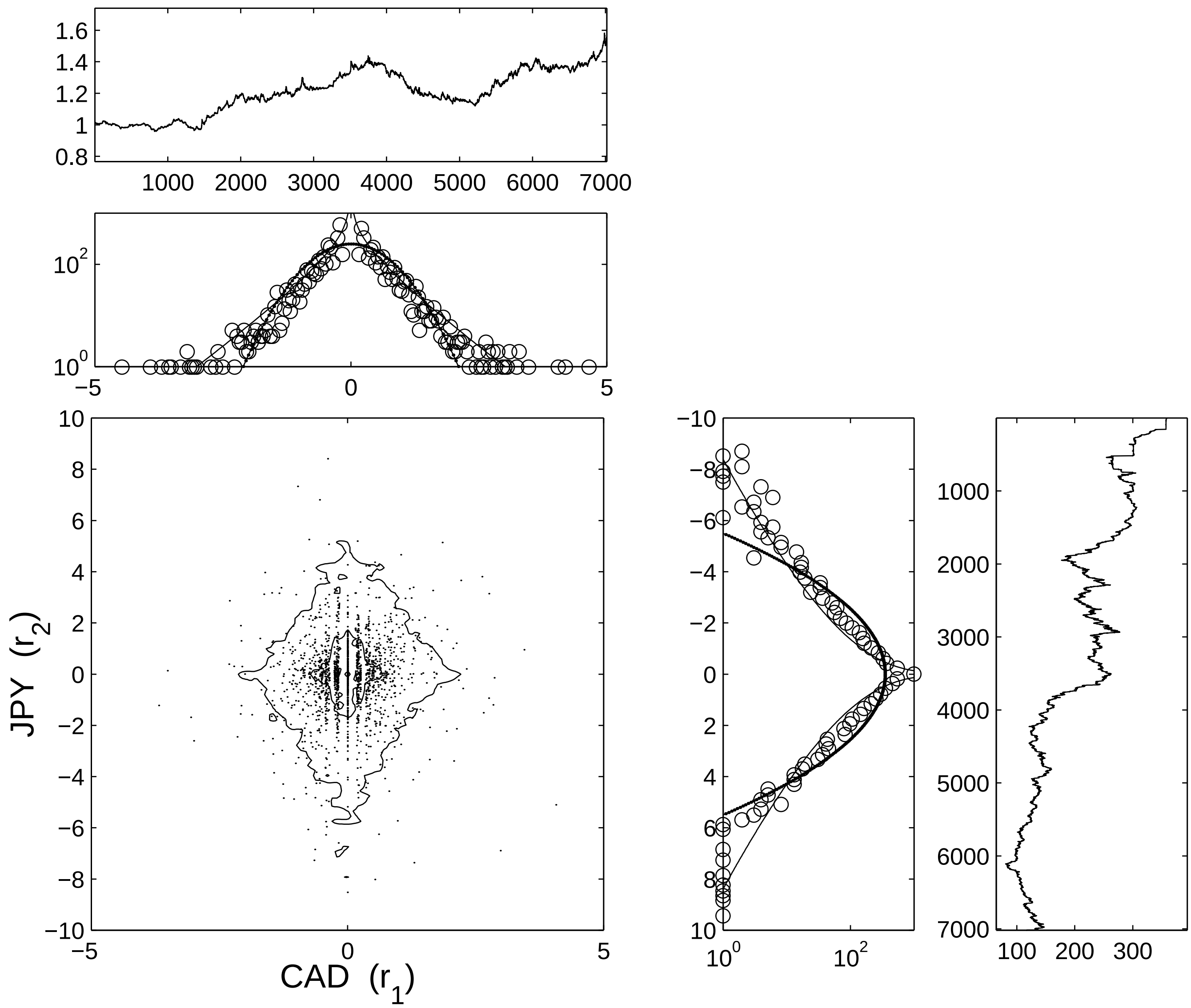

Figure 4: Bivariate distribution of the daily annualized returns of the CAD in US \$ ($i = 1$) and of JPY in US \$ ($i = 2$) for the time interval from Jan. 1971 to Oct. 1998. One fourth of the data points are represented for clarity of the figure. The contour lines define the probability confidence level of 90% (outer line), 50% and 10%. Also shown are the time series and the marginal distributions in the panels at the top and on the side. The parameters for the fit of the marginal pdf's are: CAD in US \$: $A_1 = 250, c_1 = 0.98, r_{01} = 0.59$ and JPY in US \$: $A_2 = 350, c_2 = 0.8, r_{02} = 1.25$.

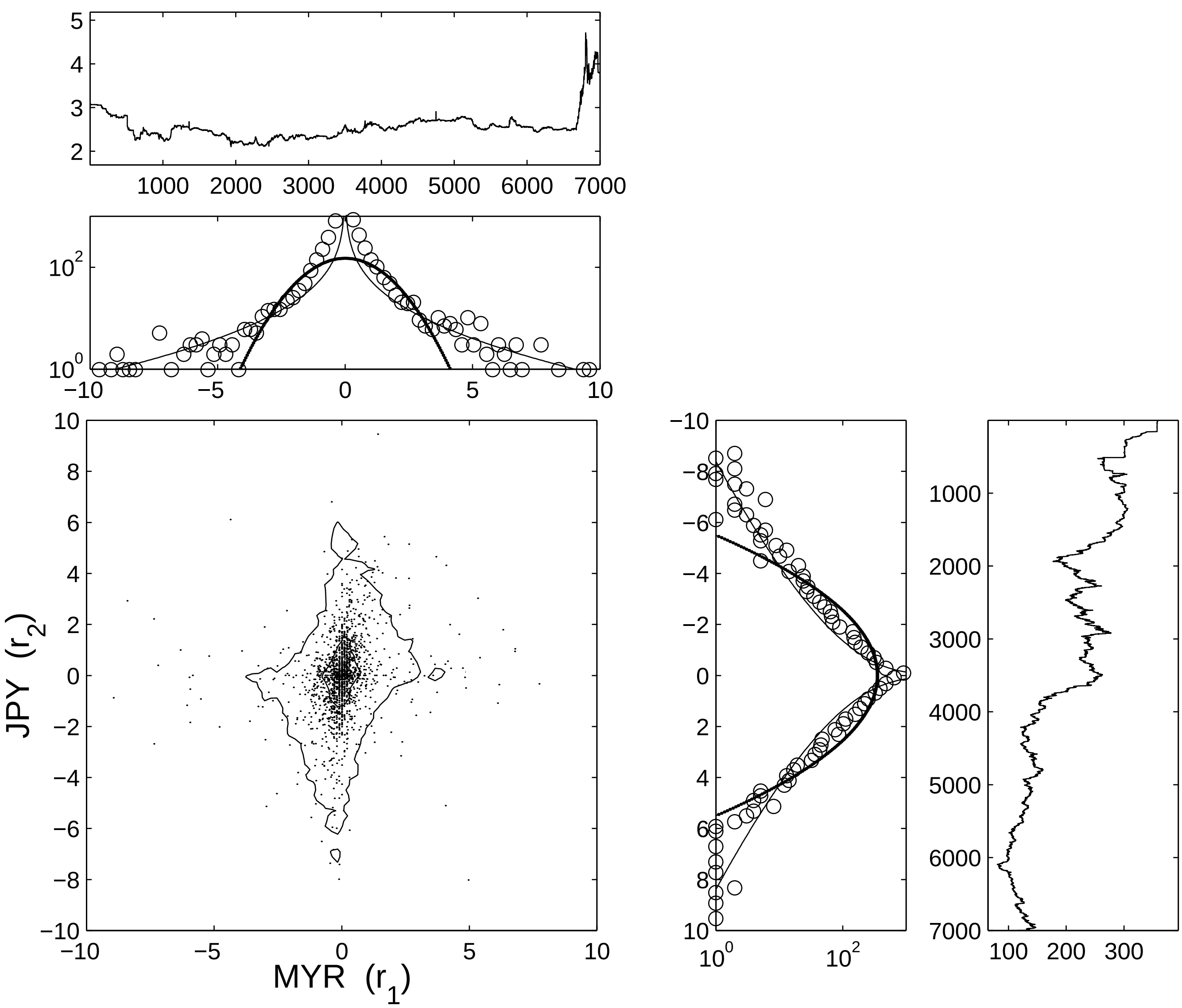

Figure 5: Bivariate distribution of the daily annualized returns of the MYR in US \$ ($i = 1$) and of JPY in US \$ ($i = 2$) for the time interval from Jan. 1971 to Oct. 1998. One fourth of the data points are represented for clarity of the figure. The contour lines define the probability confidence level of 90% (outer line), 50% and 10%. Also shown are the time series and the marginal distributions in the panels at the top and on the side. The parameters for the fit of the marginal pdf's are: MYR in US \$: $A_1 = 150, c_1 = 0.56, r_{01} = 1.00$ and JPY in US \$: $A_2 = 350, c_2 = 0.8, r_{02} = 1.25$.

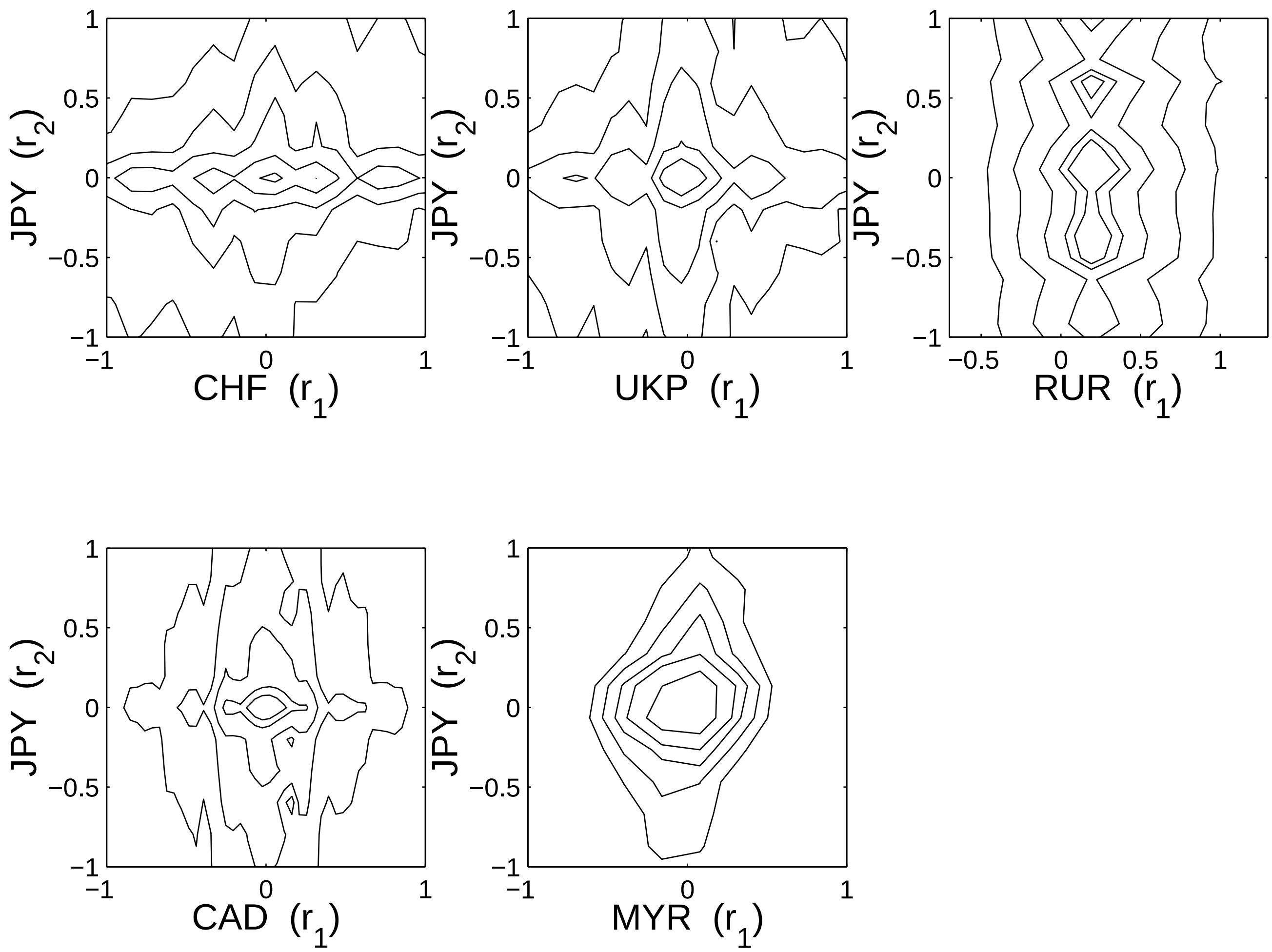

Figure 6: Close-up of the countour lines of the bivariate distributions. One can observe that the countour lines at the center are not far from elliptical while they depart more and more from ellipses for larger levels. The corresponding probability levels are:
CHF: 0.44, 0.24, 0.13, 0.07, 0.04
UKP: 0.48, 0.30, 0.17, 0.09, 0.07
RUR: 0.66, 0.56, 0.43, 0.36, 0.34
CAD: 0.60, 0.37, 0.22, 0.12, 0.10
MYR: 0.42, 0.32, 0.25, 0.20, 0.17

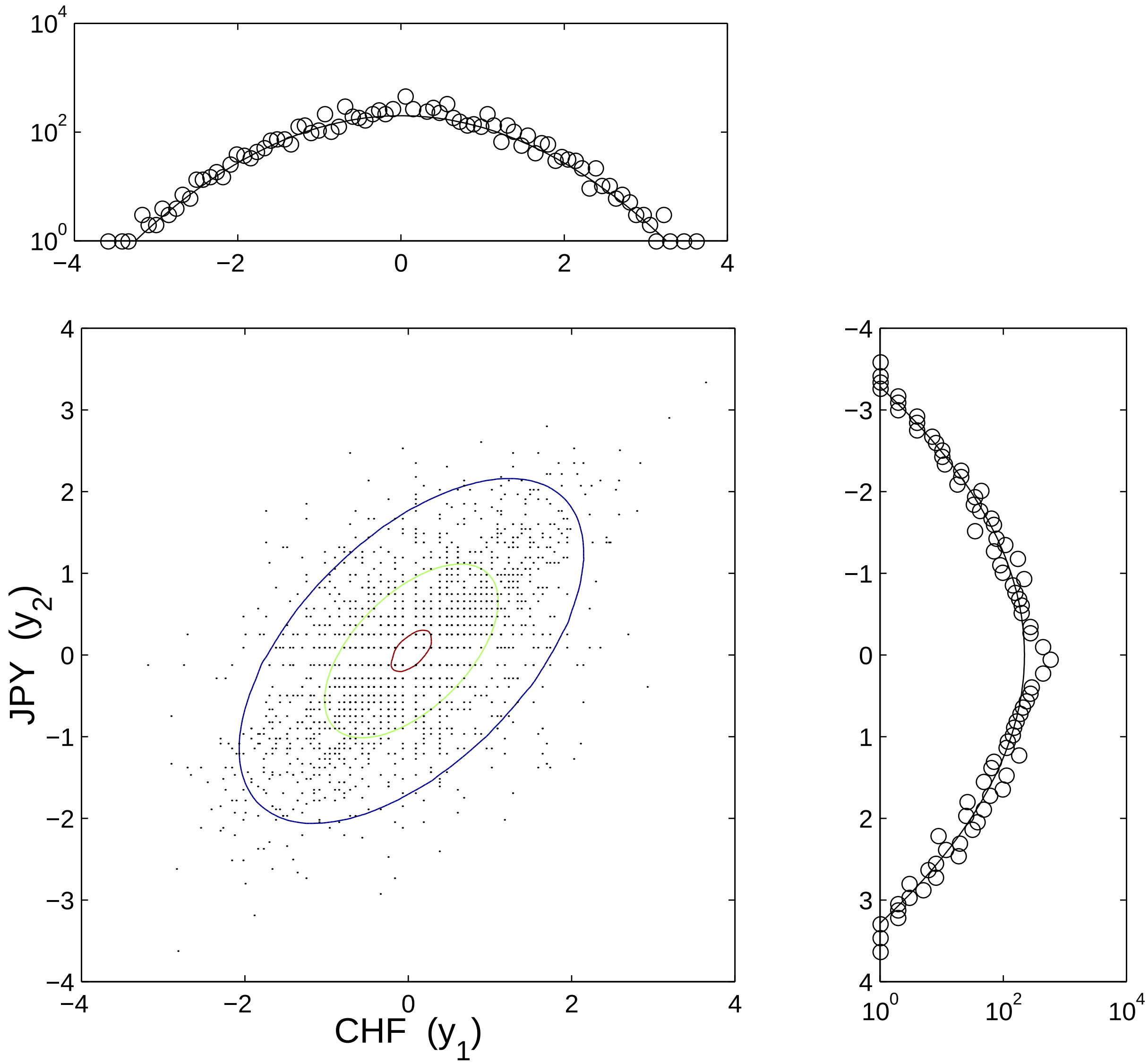

Figure 7: Bivariate distribution $\hat{P}(\mathbf{y})$ obtained from Fig.JPYCHF(r) using the transformation Eq.(16). The contour lines are defined as in Fig.JPYCHF(r). The upper and right diagrams show the corresponding projected marginal distributions, which are Gaussian by construction of the change of variable Eq.(16). The solid lines are fits of the form $A\exp(-|y|^2/2)$ with $A_1 = 200, A_2 = 220$.

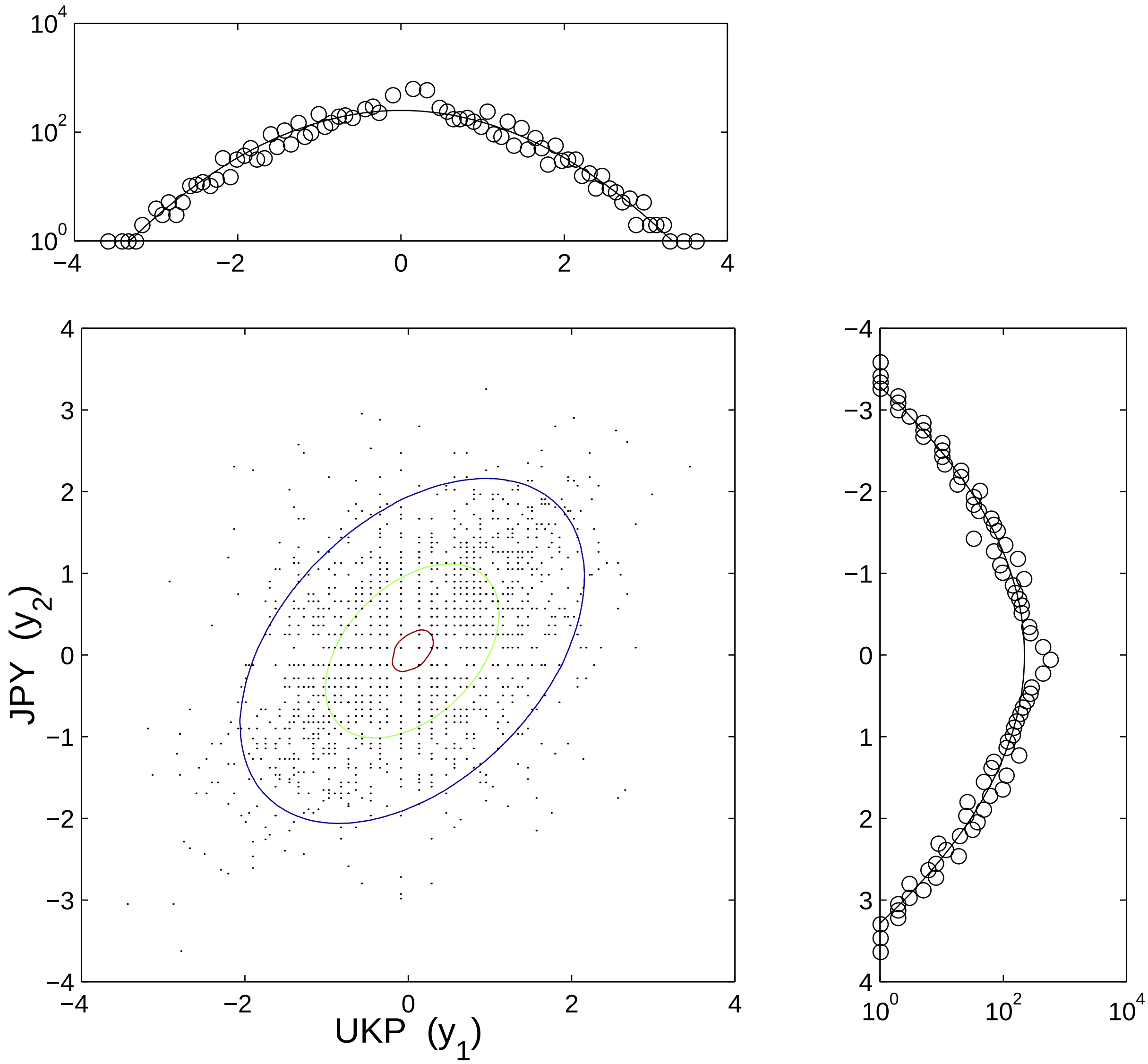

Figure 8: Bivariate distribution $\hat{P}(\mathbf{y})$ obtained from Fig.2 using the transformation Eq.(16). The contour lines are defined as in Fig.2. The upper and right diagrams show the corresponding projected marginal distributions, which are Gaussian by construction of the change of variable Eq.(16). The solid lines are fits of the form $A\exp(-|y|^2/2)$ with $A_1 = 250, A_2 = 220$.

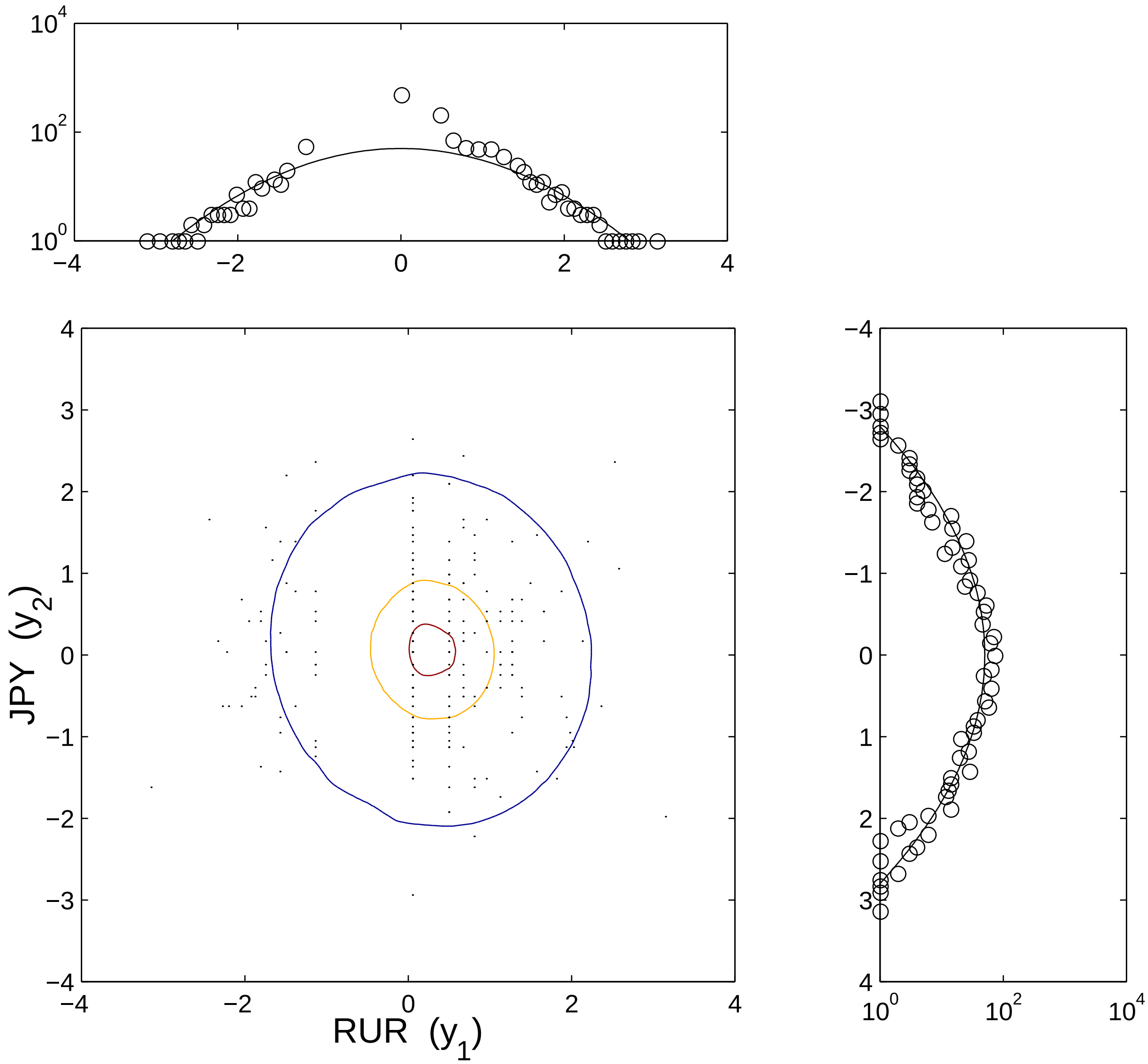

Figure 9: Bivariate distribution $\hat{P}(\mathbf{y})$ obtained from Fig.3 using the transformation Eq.(16). The contour lines are defined as in Fig.3. The upper and right diagrams show the corresponding projected marginal distributions, which are Gaussian by construction of the change of variable Eq.(16). The solid lines are fits of the form $A\exp(-|y|^2/2)$ with $A_1 = 50, A_2 = 50$.

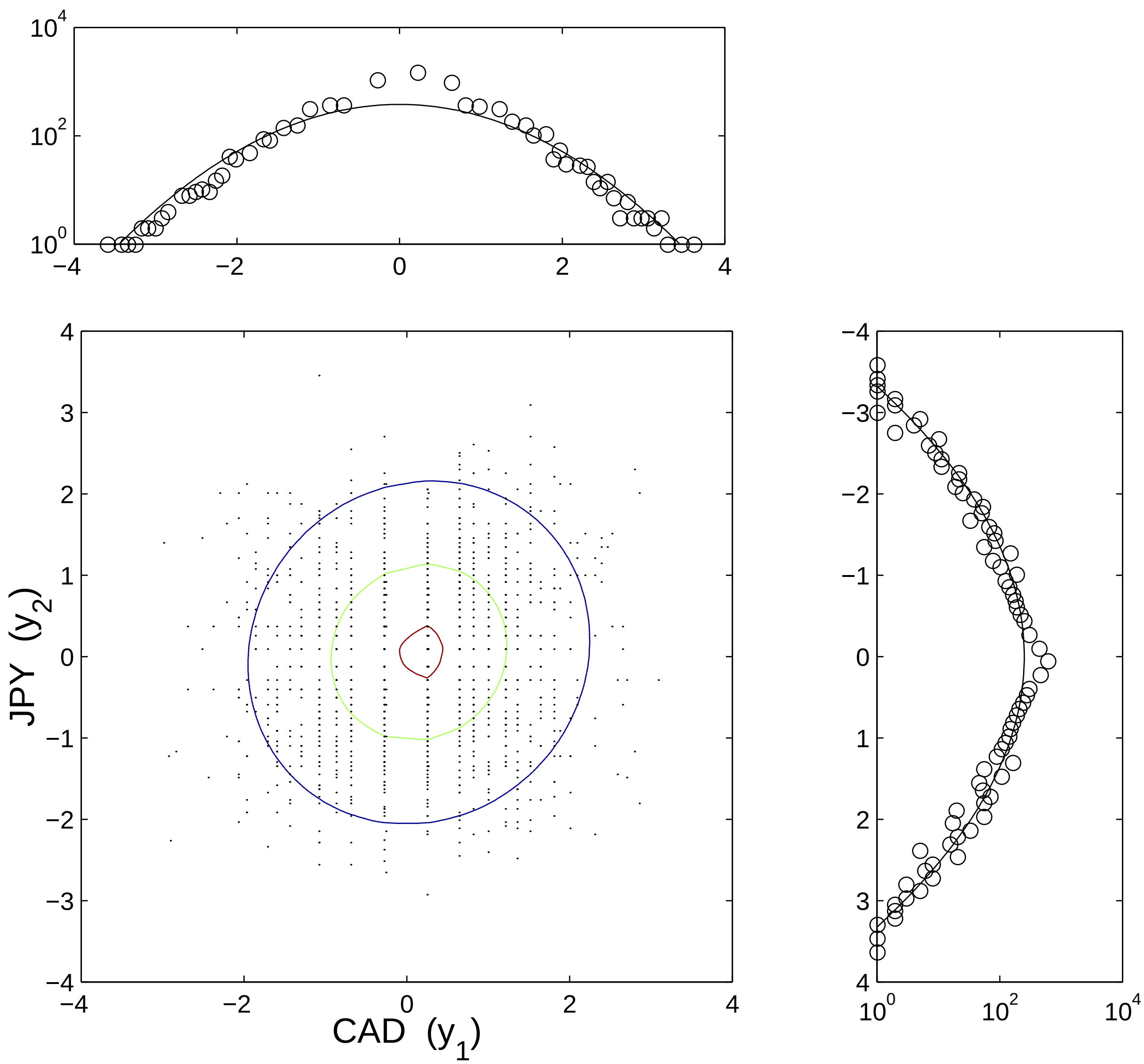

Figure 10: Bivariate distribution $\hat{P}(\mathbf{y})$ obtained from Fig.4 using the transformation Eq.(16). The contour lines are defined as in Fig.4. The upper and right diagrams show the corresponding projected marginal distributions, which are Gaussian by construction of the change of variable Eq.(16). The solid lines are fits of the form $A\exp(-|y|^2/2)$ with $A_1 = 380, A_2 = 220$.

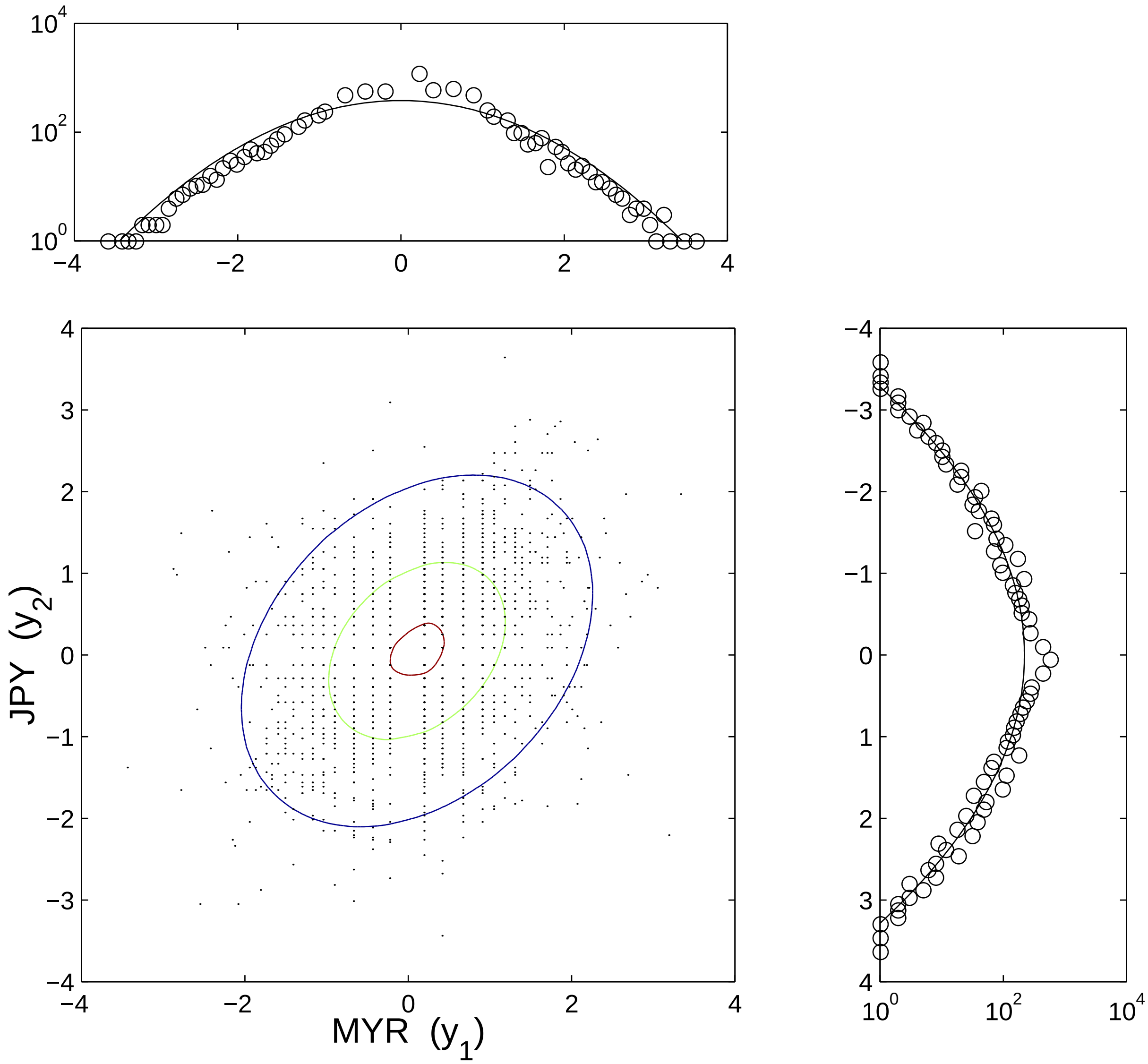

Figure 11: Bivariate distribution $\hat{P}(\mathbf{y})$ obtained from Fig.5 using the transformation Eq.(16). The contour lines are defined as in Fig.5. The upper and right diagrams show the corresponding projected marginal distributions, which are Gaussian by construction of the change of variable Eq.(16). The solid lines are fits of the form $A\exp(-|y|^2/2)$ with $A_1 = 380, A_2 = 220$.

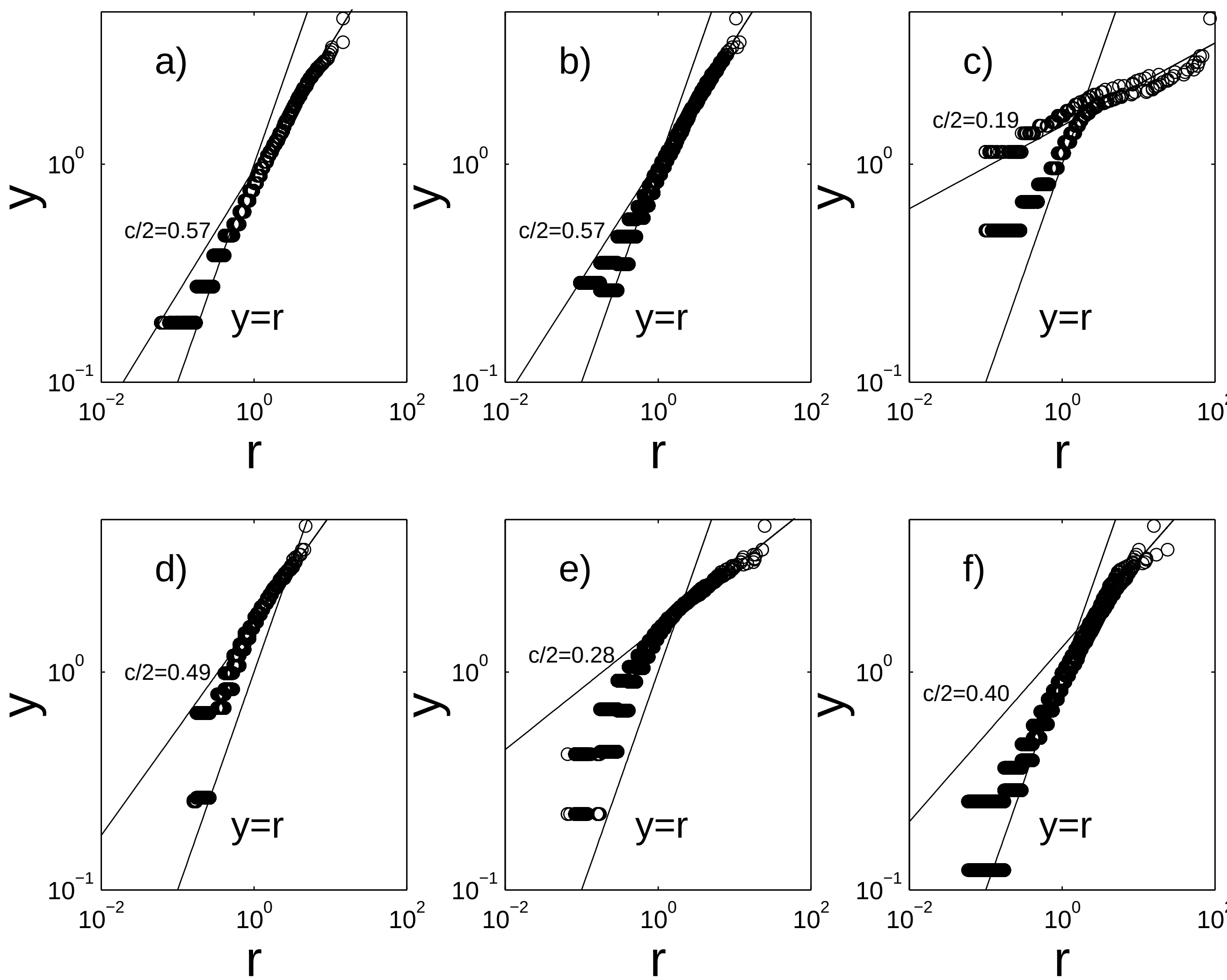

Figure 12: $y(r)$-transformation defined by Eq.(16) for a) CHF, b) UKP, c) RUR, d) CAD, e) MYR and f) JPY. The negative returns have been folded back to the positive quadrant.

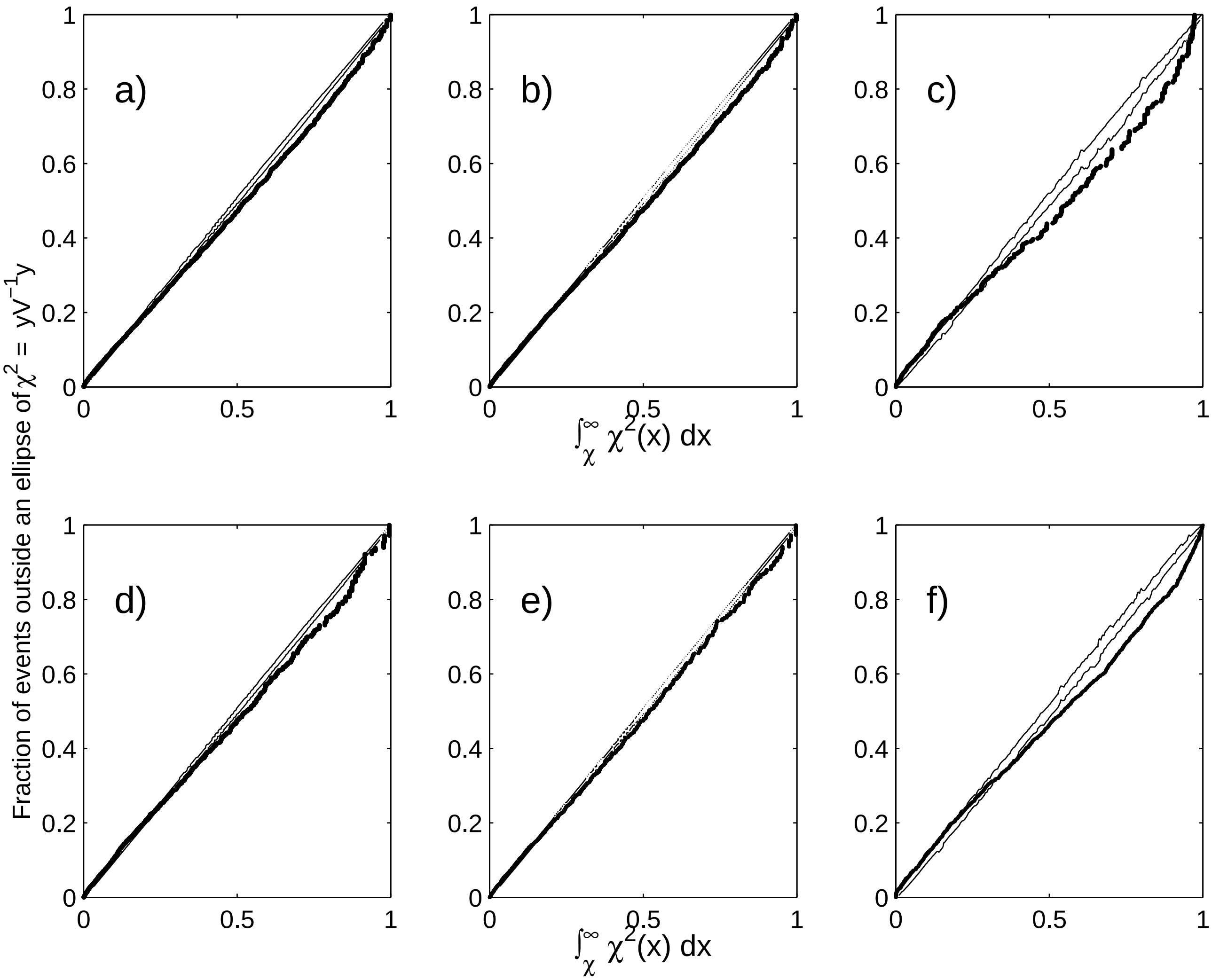

Figure 13: a-e: χ^2 cumulative distribution for $N = 2$ degrees of freedom versus the fraction of events shown in Figures 7,8,10,9, 11 outside an ellipse of equation $\chi^2 = \mathbf{y}' V^{-1})\mathbf{y}$. a) CHF-JPY, b) UKP-JPY, c) RUR-JPY, d) CAD-JPY, e) MYR-JPY. f) same plot as a)-e) but for $N = 6$ degrees of freedom for the data set CHF-UKP-RUR-CAD-MYR-JPY.

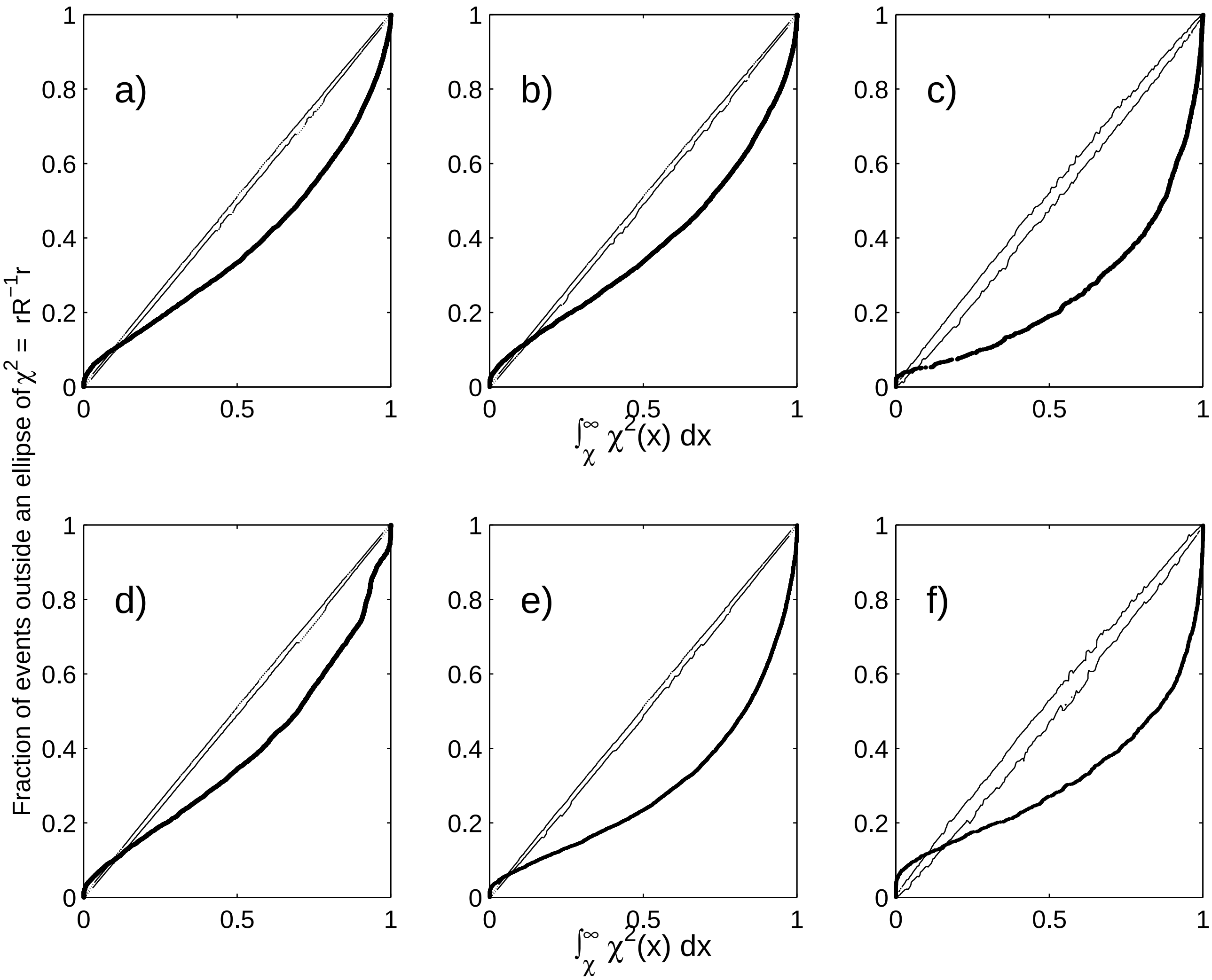

Figure 14: Same as Fig.(13) but for the returns r. a-e: χ^2 cumulative distribution for $N = 2$ degrees of freedom versus the fraction of events shown in Figures 1,2,4,3, 5 outside an ellipse of equation $\chi^2 = \mathbf{r}' V^{-1})\mathbf{r}$. a) CHF-JPY, b) UKP-JPY, c) RUR-JPY, d) CAD-JPY, e) MYR-JPY. f) same plot as a)-e) but for $N = 6$ degrees of freedom for the data set CHF-UKP-RUR-CAD-MYR-JPY.

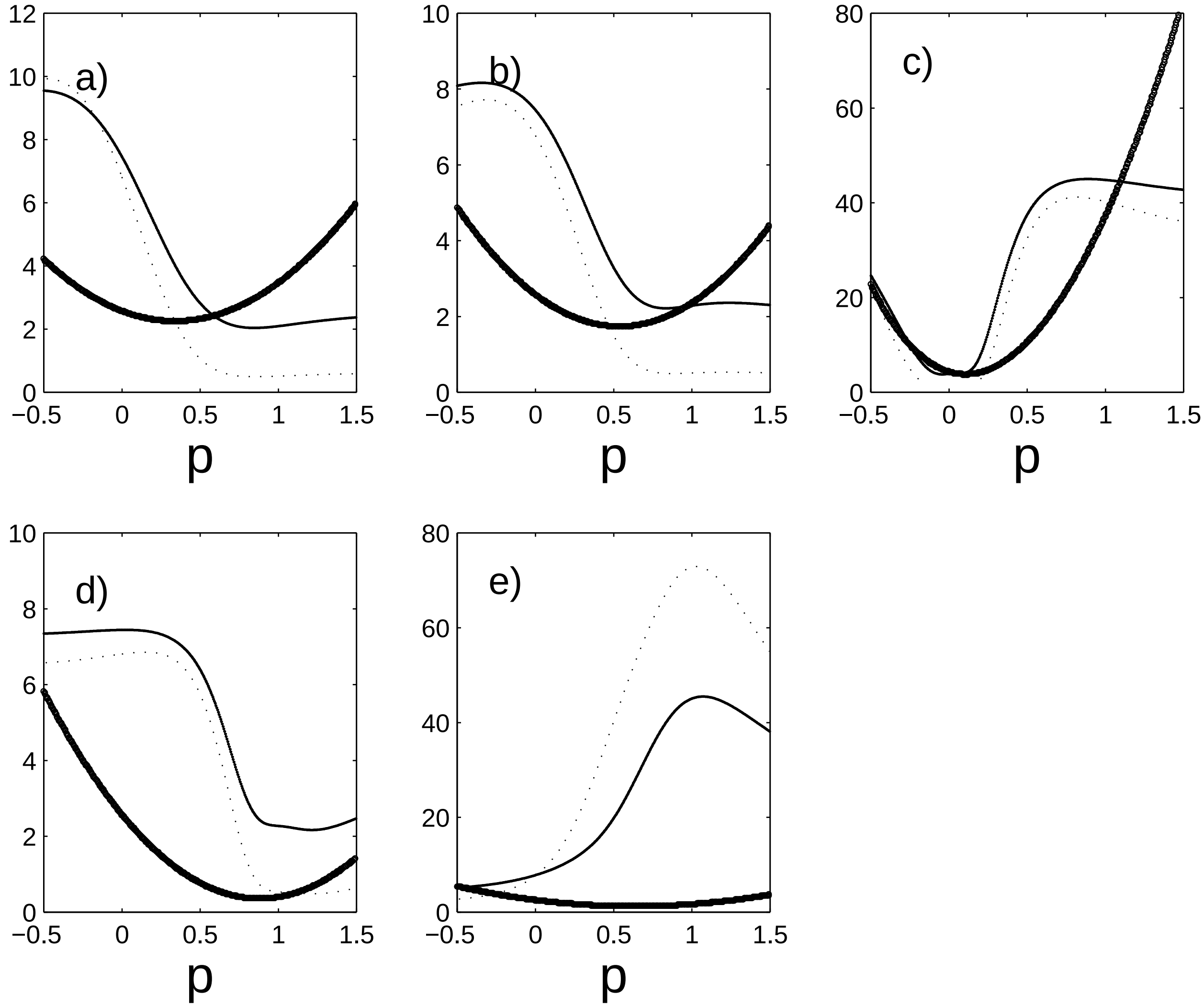

Figure 15: Variance (thick solid line), excess kurtosis κ (thin solid line) and sixth-order normalized cumulant λ_6 as a function of the weight p of currency 1 (the weight of currency 2 is $1-p$). for the data sets: a) CHF-JPY, b) UKP-JPY, c) RUR-JPY, d) CAD-JPY and e) MYR-JPY. κ is divided by 2 and λ_6 is divided by 300.

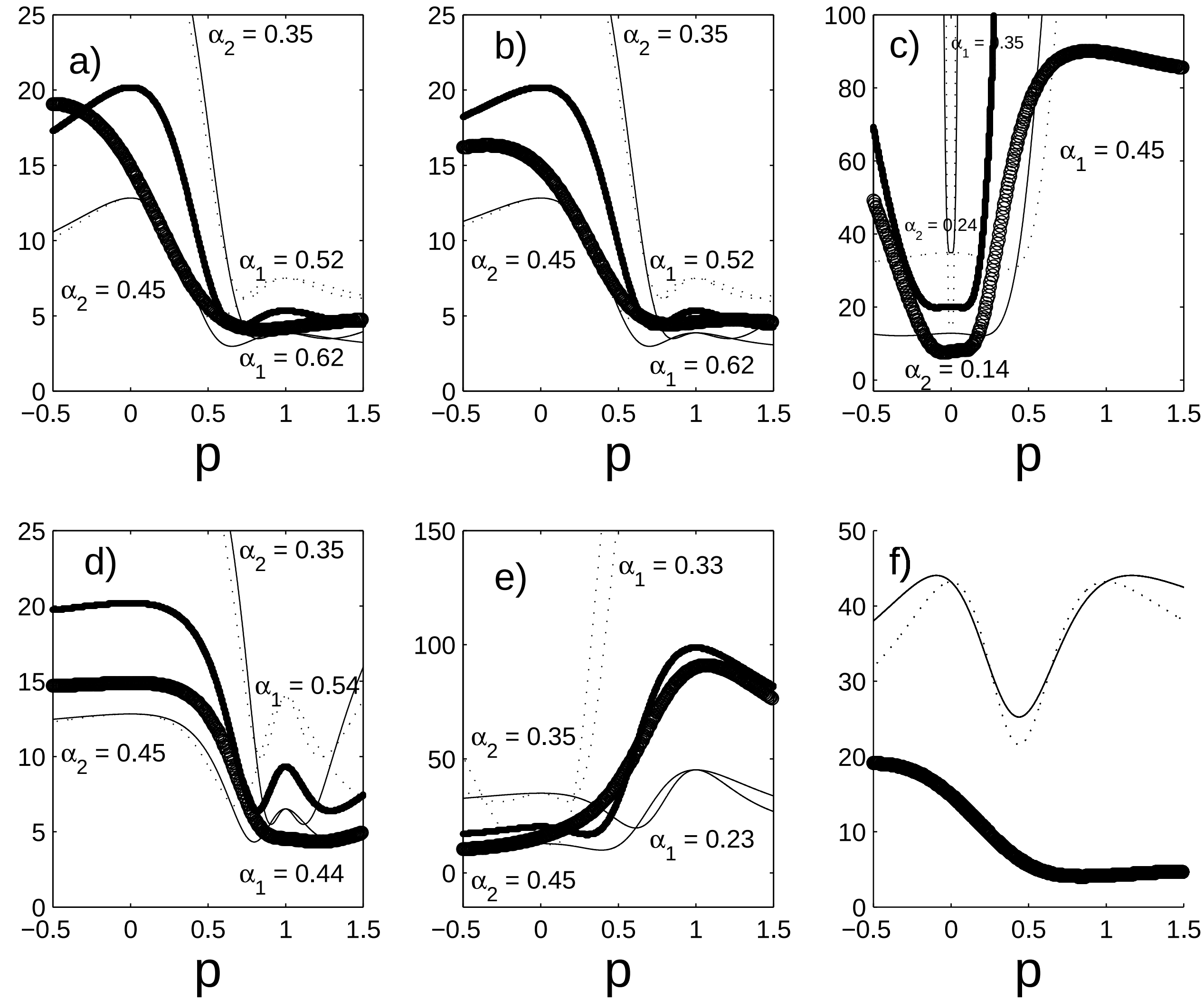

Figure 16: Comparison of the empirical excess kurtosis (fat solid line) shown in Fig. 15 for the five portfolios to our theoretical prediction (40) with (41) for uncorrelated assets (solid line). The exponents c_1 and c_2 are those determined in the fits of the pdf's tail, as given in Fig.12. The thin solid lines and dashed lines plot the theoretical formula (40) for values of the exponents $\alpha_i \equiv c_i/2 \pm 0.05$. Figure c (RUR) is the same as a-b and d-e but the thin solid line gives the predicted excess kurtosis for exponents $c_1 + 0.05, c_2 \pm 0.05$. Figure f compares the empirical excess kurtosis (fat solid line) of the portfolio CHF-JPY (Figure a) to the prediction (65) with (63) for correlated assets with the fixed exponents $c_1 = c_2 = 2/3$. The thin solid line correspond to the empirical value $\rho(y1, y2) = 0.57$ while the dashed line is obtained from the same formula with $\rho(y1, y2) = 0$.

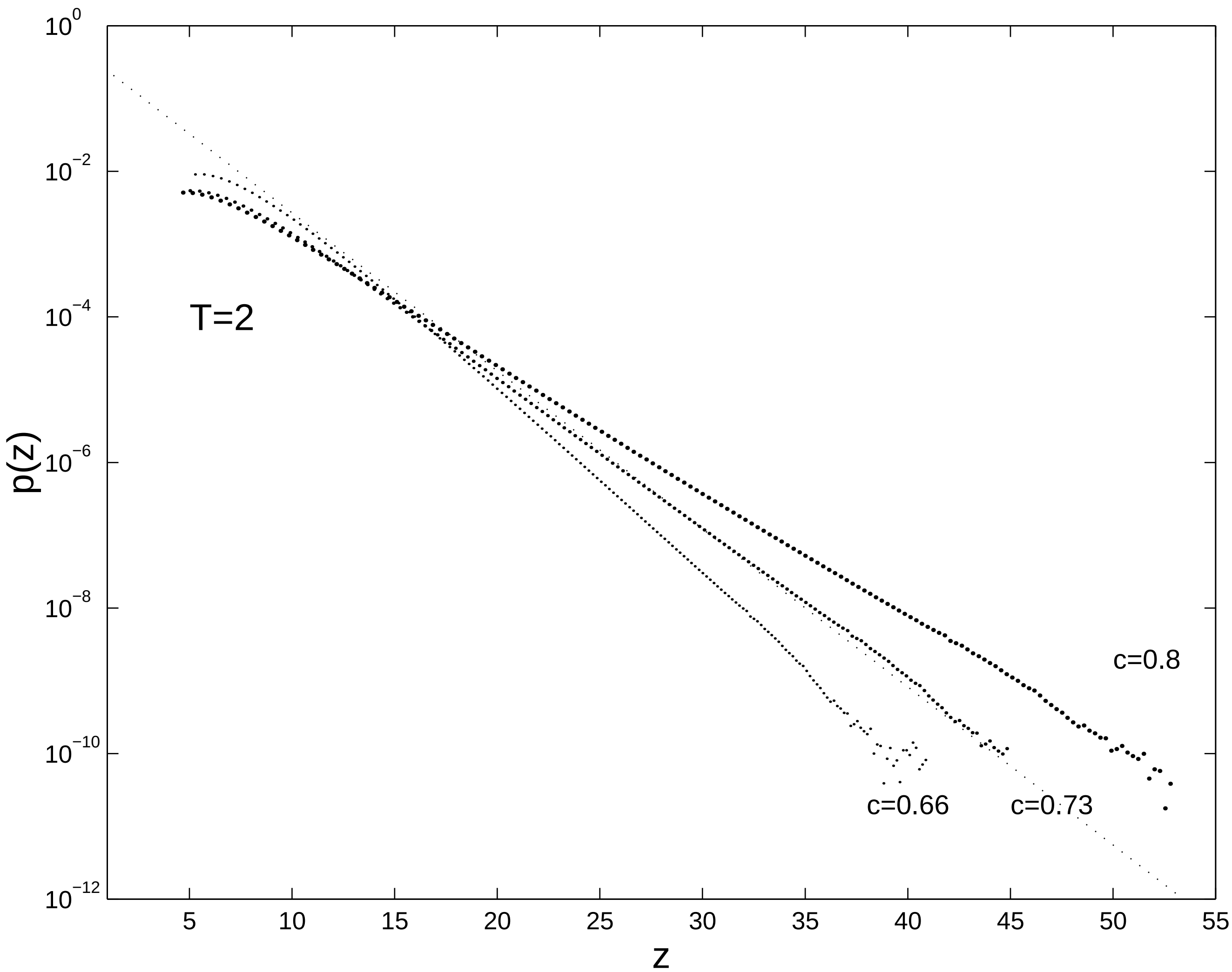

Figure 17: Plots of the pdf's $P_T(y)$ as a function of $z \equiv y^2$ so that a Gaussian (in the y variable) is qualified as a straight line (dashed line). In turn, by the construction explained in the text, a straight line qualifies a Weibull distribution. Here is shown the case $T = 2$ for which the best c_T is $c_2 = 0.73$. The other curves allow one to estimate the sensitivity of the representation of P_T in terms of a Weibull as a function of the choice of the exponent c_T. The curves have been normalized by a coefficient A_T^T, with $A_2 = 10$ for $c_2 = 0.66$ and $A_2 = 8$ for $c_2 = 0.73$ and 0.8.

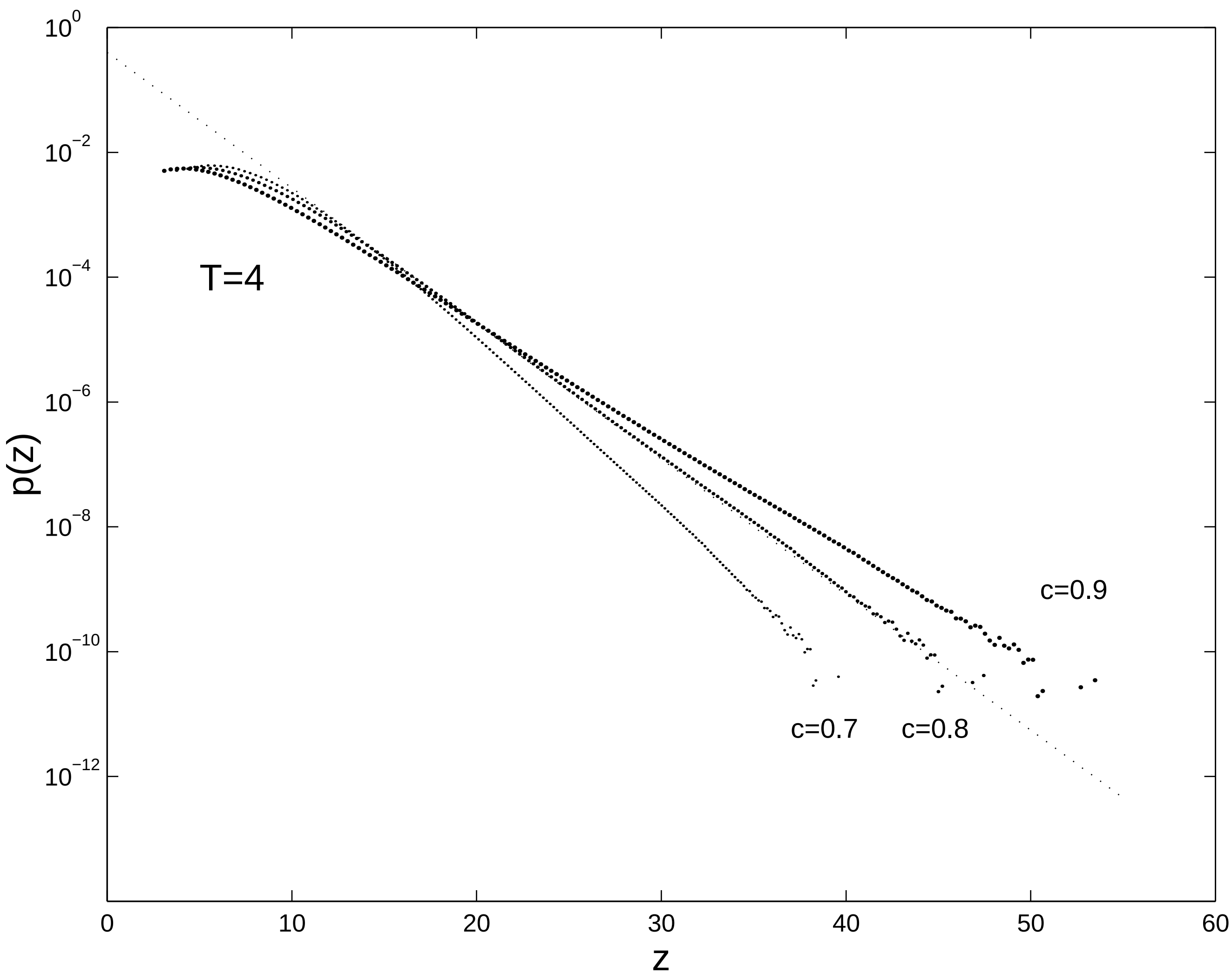

Figure 18: Plots of the pdf's $P_T(y)$ as a function of $z \equiv y^2$ so that a Gaussian (in the y variable) is qualified as a straight line (dashed line). In turn, by the construction explained in the text, a straight line qualifies a Weibull distribution. Here is shown the case $T = 4 forwhichthebest$c$_T$ is $c_4 = 0.80$. The other curves allow one to estimate the sensitivity of the representation of P_T in terms of a Weibull as a function of the choice of the exponent c_T. The curves have been normalized by a coefficient A_T^T, with $A_4 = 25$ for all c_4's.

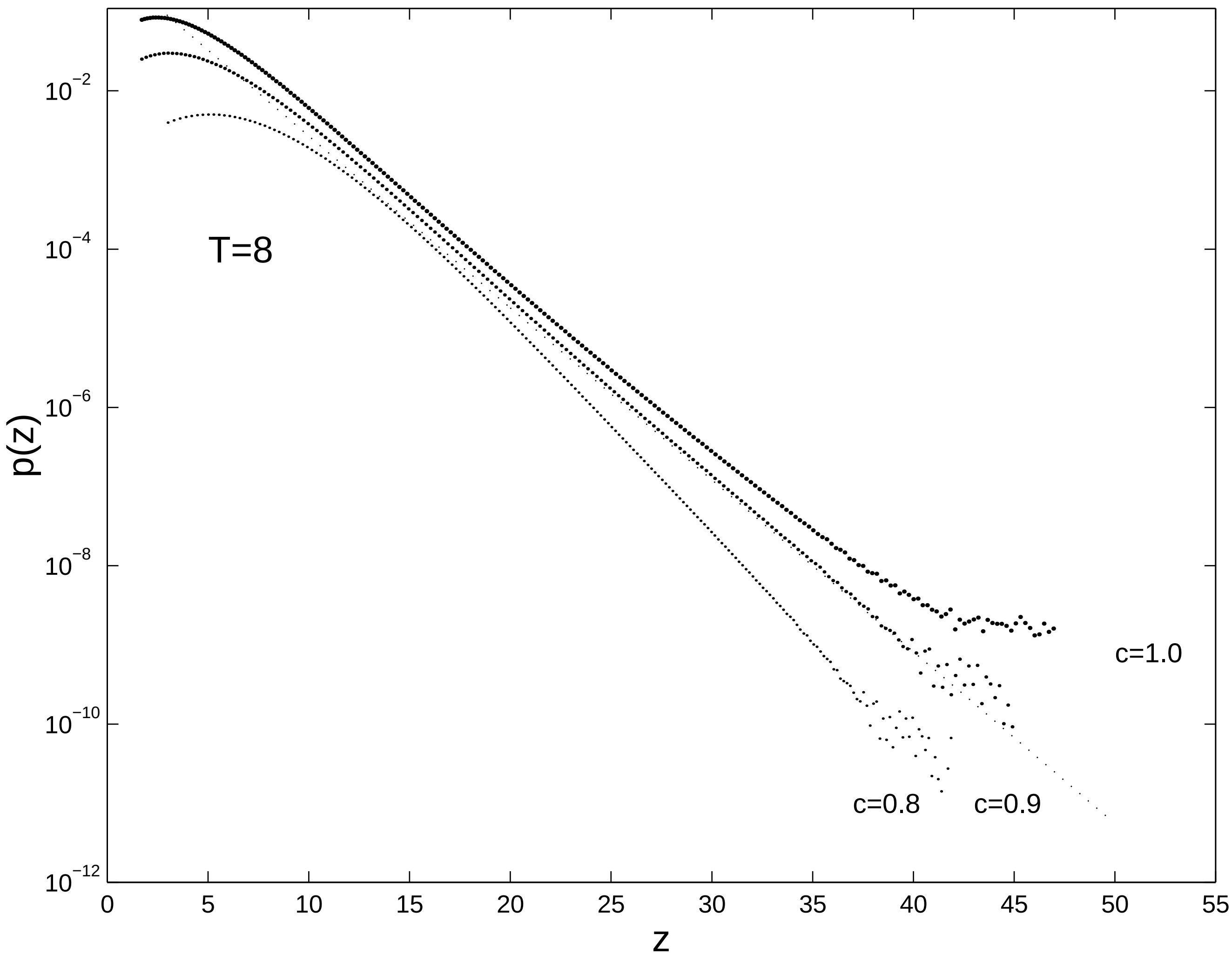

Figure 19: Plots of the pdf's $P_T(y)$ as a function of $z \equiv y^2$ so that a Gaussian (in the y variable) is qualified as a straight line (dashed line). In turn, by the construction explained in the text, a straight line qualifies a Weibull distribution. We show here the case $T = 8$ for which the best c_T is $c_8 = 0.90$. The other curves allow one to estimate the sensitivity of the representation of P_T in terms of a Weibull as a function of the choice of the exponent c_T. The curves have been normalized by a coefficient A_T^T, with $A_8 = 40$ for $c_8 = 0.8$ and 1.0 and $A_8 = 360$ for $c_8 = 0.9$.

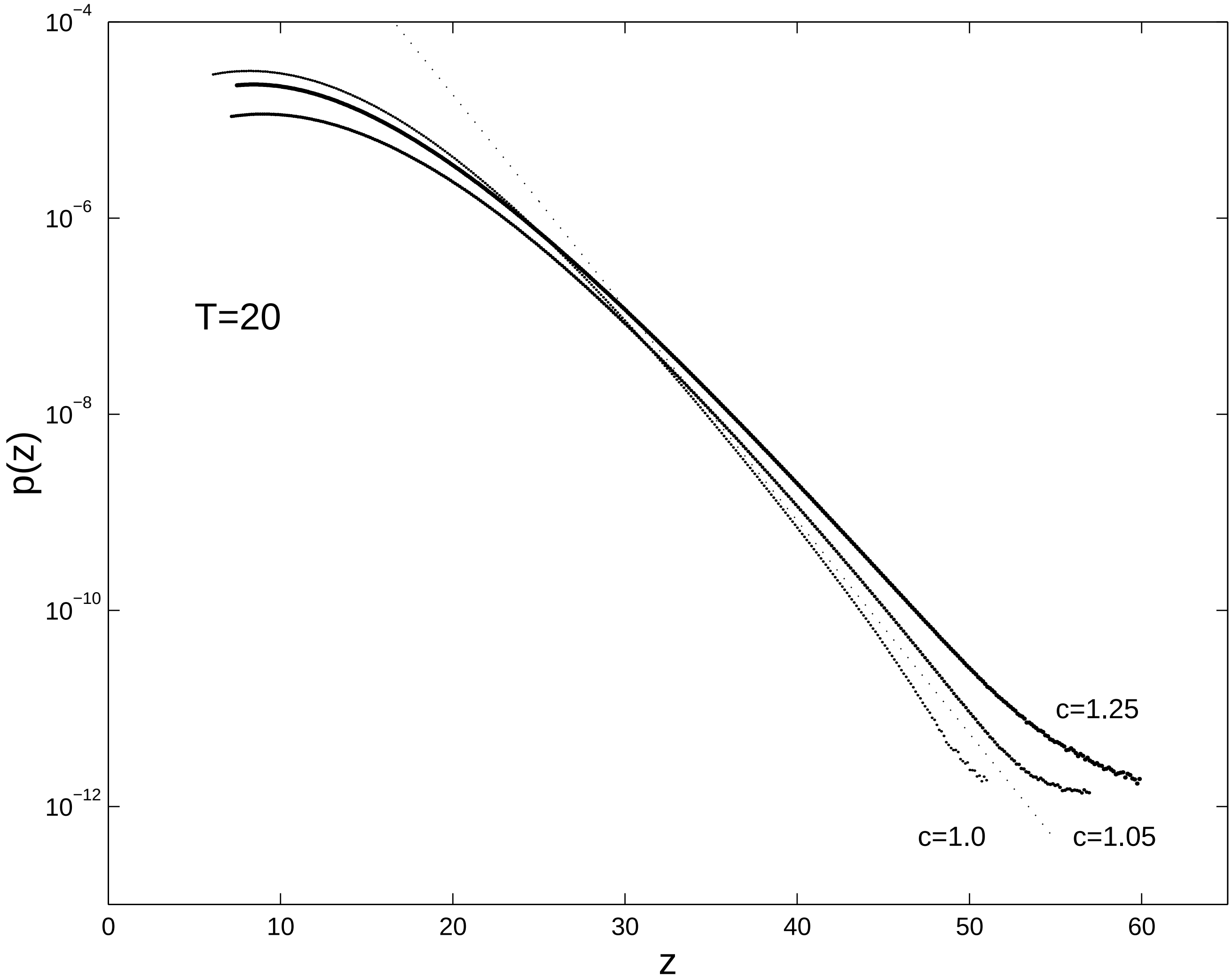

Figure 20: Plots of the pdf's $P_T(y)$ as a function of $z \equiv y^2$ so that a Gaussian (in the y variable) is qualified as a straight line (dashed line). In turn, by the construction explained in the text, a straight line qualifies a Weibull distribution. We show here the case $T = 20$ for which the best c_T is $c_{20} \approx 1.05$. The other curves allow one to estimate the sensitivity of the representation of P_T in terms of a Weibull as a function of the choice of the exponent c_T. The curves have been normalized by a coefficient A_T^T, with $A_{20} = 4.7\ 10^4$ for $c_{20} = 1.0$, $A_{20} = 3.5\ 10^5$ for $c_{20} = 1.05$ and $A_{20} = 7\ 10^7$ for $c_{20} = 1.25$.

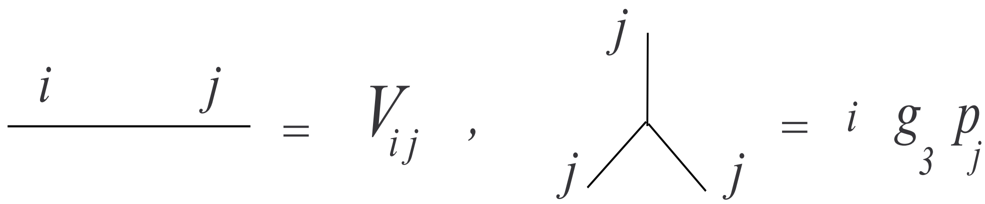

Figure 21: We associate to each factor V_{ij} the propagator diagram and to each factor $ig_3 w_j$ the vertex diagram as shown in this figure, where we have defined the coupling constant $g_3 = 3!k$.

Figure 22: The two contributions in (159) are represented by propagators connecting the vertices as shown in the figure.

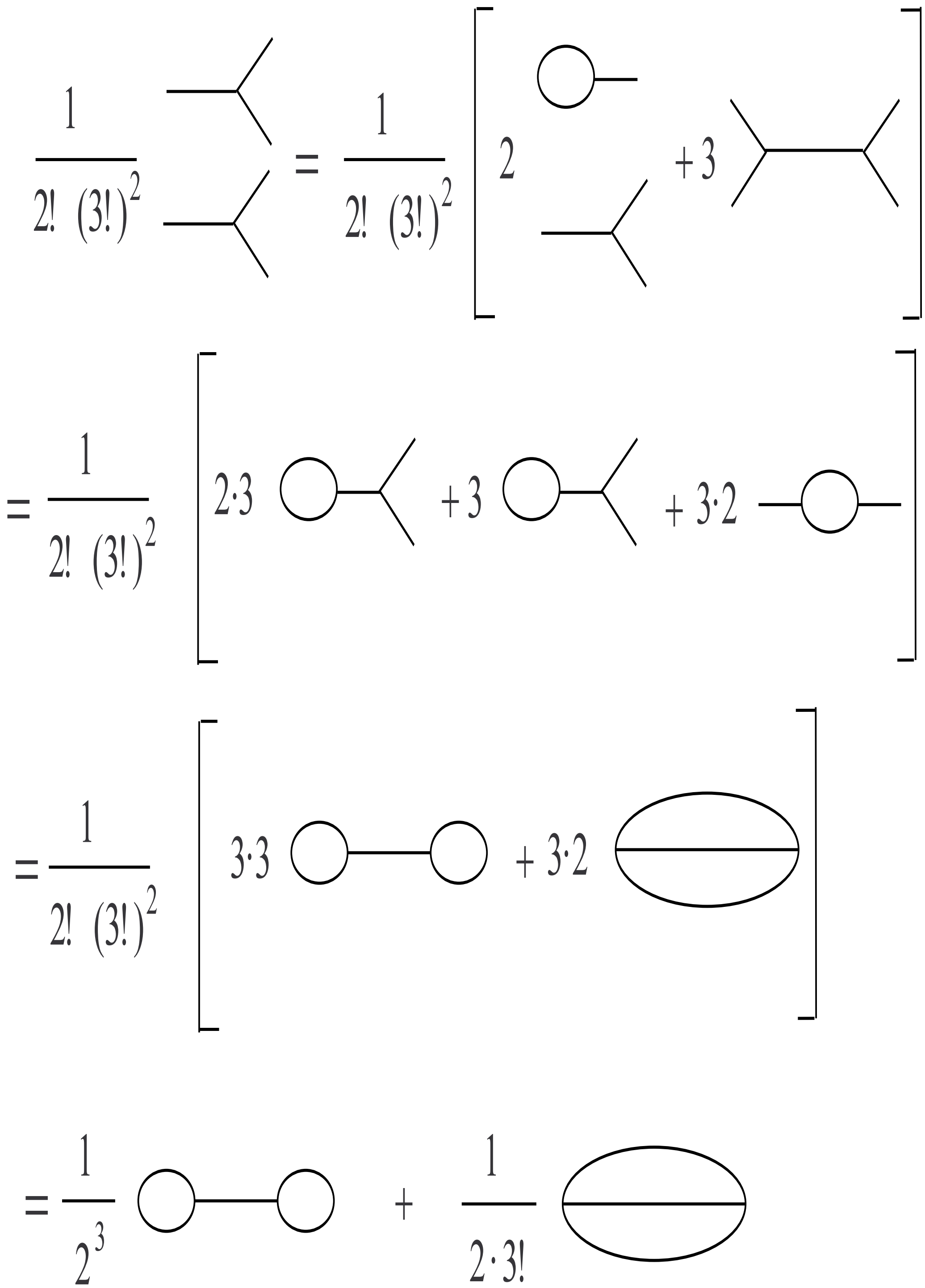

Figure 23: A systematic way to keep under control the symmetry factor is to compute it diagramatically as shown in this figure. The second-order derivative operator with respect to J_i and J_j is represented by the two vertices in the left hand side of the figure. The J-independent term is given by pairwise combining each leg of each vertex in all the possible topologically inequivalent way, taking into account the multiplicity of each configuration. The figure shows how to proceed. The coefficient in front of the vertices of the left hand side comes from the perturbative expansion. As a first step, let us consider the first leg of the first vertex. We can either contract it with another leg (two possibile contractions) of the same vertex or with a leg of the second vertex (three possible contractions). This is summarized in the first equality of the figure.

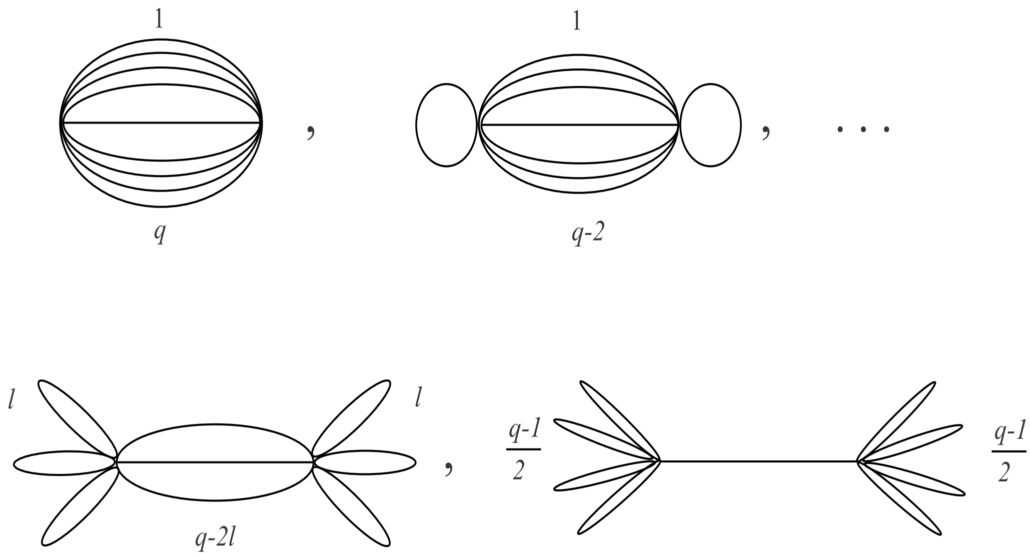

Figure 24: Sequence of all diagrams obtained by combining pairwise each leg of two vertices to form all the possible topologically inequivalent diagrams. Each diagram is characterized by the number l of loops on each vertex and the number $(q-2)l$ of lines connecting the two vertices giving the contribution (163) where each loop around vertex i contributes to a factor V_{ii} and each propagator connecting the vertices i and j gives a factor V_{ij}.

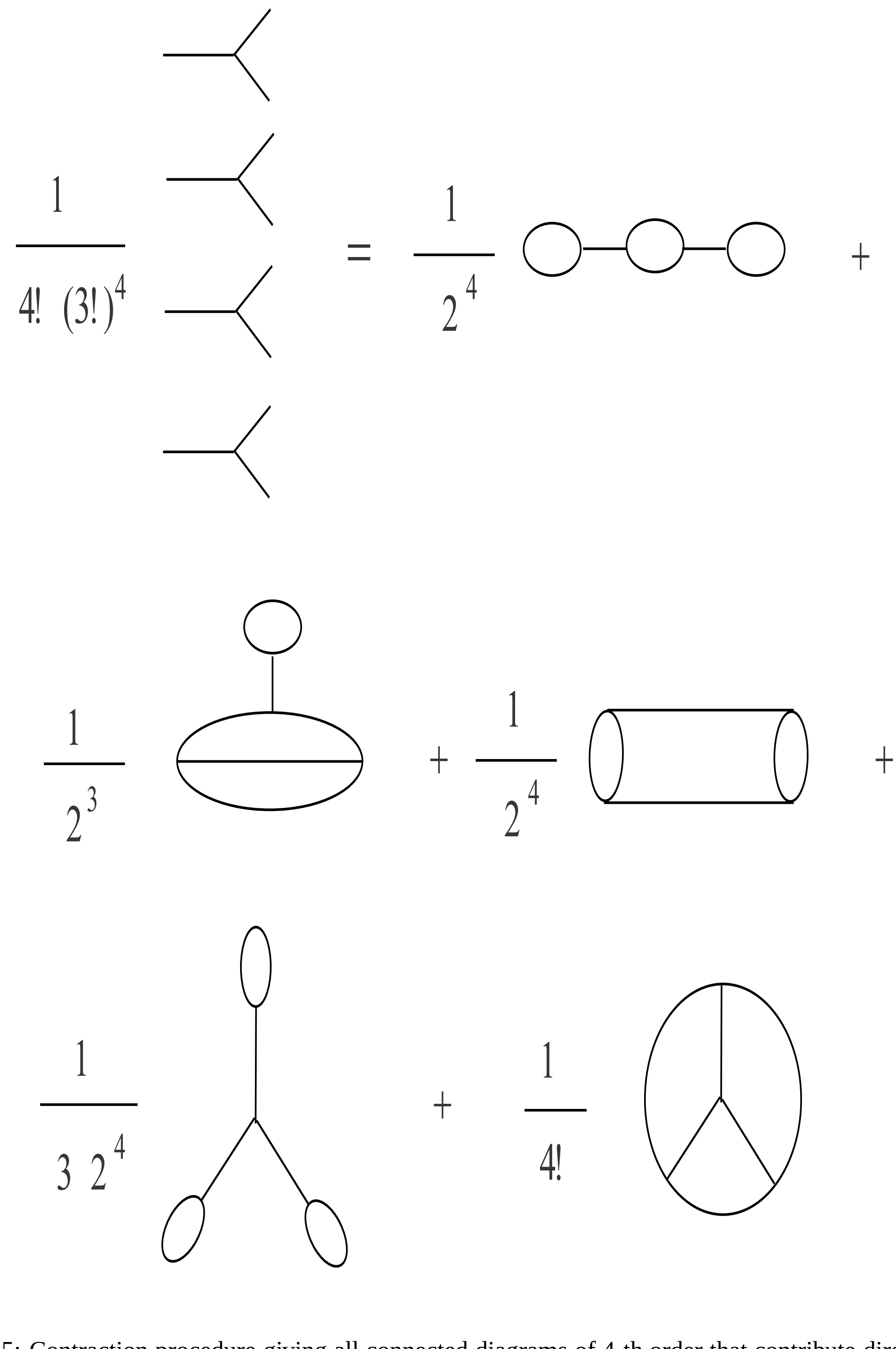

Figure 25: Contraction procedure giving all connected diagrams of 4-th order that contribute directly to the 4-th cumulant coefficient $c_4(3)$ (165).

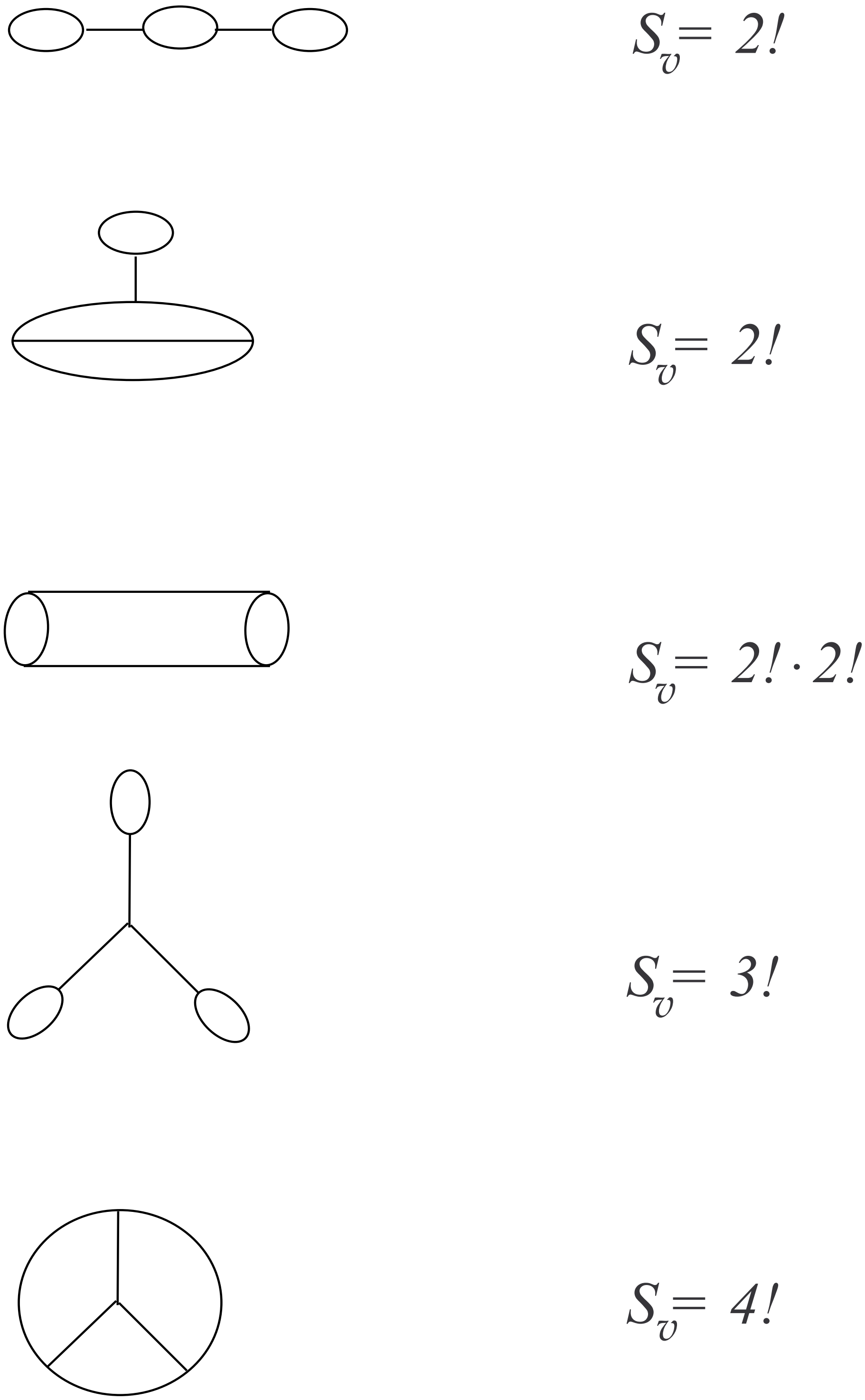

Figure 26: Summary of the vertex symmetry factors for the diagrams contributing to $c_4(3)$.