On Portfolio Selection under Extreme Risk Measure: the Heavy-tailed ICA Model

Stéphan Clémençon
MODAL'X - Université Paris X Nanterre
LPMA - UMR CNRS 7599 - Universités Paris VI et Paris VII
stephan.clemencon@u-paris10.fr

Skander Slim
Laboratoire Théorie Economique, Modélisation et Applications
UMR CNRS 7536 - Université Paris X Nanterre
skander.slim@u-paris10.fr

Abstract : This paper is devoted to the application of the Independent Component Analysis (*ICA*) methodology to the problem of selecting portfolio strategies, so as to provide against extremal movements in financial markets. A specific ICA model for describing the extreme fluctuations of asset prices is introduced, stipulating that the distributions of the IC's are heavy tailed (*i.e.* with power law behaviour at infinity). An inference method based on conditional maximum likelihood estimation is proposed for our model, which permits to determine practically optimal investment strategies with respect to extreme risk. Empirical studies based on this modelling are carried out to illustrate our approach.

Key words and phrases: independent component analysis, portfolio selection, heavy-tailed distribution, extreme values, conditional MLE, safety first investment strategies

1 Introduction

Borrowing tools and statistical techniques commonly used in the field of insurance for risk assessment, many finance experts have to face questions related to extremal events for handling problems concerning the probable maximal loss of investment strategies (see Embrechts *et al.* (1999), Embrechts (2000) or Chavez-Demoulin & Embrechts (2004) for instance and refer to Embrechts *et al.* (1997) for a comprehensive overview of the applications of extreme values methodology to insurance and finance). As highlighted by the recent turbulences in financial markets such as the Russian financial crisis in 1998 or the burst of the speculative information technology bubble in 2001, extreme price fluctuations may expose at times portfolio managers to high levels of market risk (*cf* Zajdenweber (2000)). This phenomenon is naturally more pronounced for emerging markets, in which such extremal variations are frequent (see Susmel (2001) for instance). When investing in these markets, funds managers have to pay more attention on the distribution of "large" stocks returns values and implement suitable risk management tools to avoid big losses. Some funds managers and other market participants have been driven to adopt such quantitative risk management techniques not only in reaction to their own experience of market turbulence, but also because of regulatory climate (see Chapters 1-4 in Crouhy *et al.* (2000)). Hence, pension funds and insurance companies portfolio managers are not allowed to take high risks and are constrained in their management by some statutory regulatory restrictions like *prudent man rule*[1] or *quantitative portfolio restrictions*[2], with the aim to ensure diversification for limiting investment risks: the primary goal of any prudential control being to ensure that the pension fund or insurance beneficiaries are adequately protected and that they will receive the benefits or the compensation they are entitled to. In this specific context, namely when risk aversion takes precedence of potential gain in an overwhelming fashion (see Roy (1952), Arzac & Bawa (1977) and refer to Kahneman & Tversky (1979) for a discussion on the dissymetry in the perception of gains and losses, see also Rabin (2000) and Rabin & Thaler (2001) for a quantitative description of the degree of risk aversion of investors

[1] a concept whereby instruments are made in such a way that they are considered to be handled " prudently" as someone would do in the conduct of his own business.

[2] a quantitative limitation of a given asset class. Typically, those instruments whose holding is limited are those with high price volatility and low liquidity.

for large stakes), standard methodologies for portfolio allocation, such as approaches based on mean-variance optimization (*cf* Markowitz (1952)), are not convenient any more. Hence, various methods have been recently proposed for addressing the portfolio construction problem in specific nongaussian frameworks. For instance, in Malevergne & Sornette (2001) gaussian copulas combined with a family of modified Weibull distributions are used for modelling the tail of the flow of returns and obtaining as a byproduct the tail behaviour of the return of any portfolio. And in Bradley & Taqqu (2004), the problem of how to allocate assets for minimizing particular quantile based measures of risk is considered via the *structure variable approach*, using the tools of univariate extreme values theory for modelling the tail of the portfolio (see also Jansen *et al.* (2000)).

In this paper, we consider the problem of selecting a portfolio so as to provide against extreme loss. From a specific modelling of extremal lower fluctuations of asset returns based on *Independent Component Analysis* (*ICA*), we show how to quantify the risk of extreme loss of any investment strategy in our specific setting and how diversification should be carried for minimizing this particular extreme risk measure. The outline of the paper is as follows. In section 2, the *tail index* is discussed as a specific risk measure regarding to maximal relative loss for heavy-tailed portfolios. In section 3, the key concepts of ICA are briefly recalled and the applications of this recent statistical methodology for latent variables analysis to mathematical finance are reviewed. Section 4 is devoted to the description of the *heavy-tailed ICA model* we propose for describing the extreme fluctuations of stock prices returns, in which the IC's are assumed to have heavy tails of Pareto's type (*i.e.* power-like lower tails) and may be interpreted as *elementary portfolios* returns. An estimation method based on conditional maximum likelihood is detailed, which provides a numerical procedure for recovering the elementary portfolio strategies and their extreme risk measures as well. The elementary portfolio with the largest left tail index is shown to be optimal regarding to extreme risk with respect to our setting. The relevance of this modelling for financial returns is discussed through several applications: empirical studies are carried out in section 5, which show on some examples how the diversification induced by our model performs. In section 6, some concluding remarks are collected, together with several lines of further research.

2 On measuring extreme risks of portfolio strategies

Risk quantification for financial strategies have been the object of intense research, still developping. Given the nongaussian character of financial returns distributions and in consequence the limitation of the variance as an indicator for describing the amount of uncertainty in their fluctuations, various risk measures have been proposed (see Szegö (2004) for a recent survey on this subject, as well as Chapters 5-6 in Crouhy *et al.* (2000)), such as Value at Risk (*VaR*, see Jorion (1997), Duffie & Pan (1997) for instance) or Expected Shortfall (*ES*, one may refer to Acerbi & Tasche (2002)), which are both quantile-based risk measures. Risk measures may be considered in particular for guiding investment behaviour. Once a risk measure is chosen, the matter is then to select an optimal portfolio with respect to this latter. Suppose that there are D (risky) assets available, indexed by $i \in \{1, ..., D\}$. Let us fix a certain (discrete) time scale for observing the fluctuations of asset prices and denote by $X_i(t)$ the price of the i-th asset at time t. Let $r_i(t) = (X_i(t) - X_i(t-1))/X_i(t-1)$ be the return at time $t \geqslant 1$ and denote by $r(t) = (r_1(t), ..., r_D(t))'$ the flow of returns. Consider the portfolio strategy consisting in investing a fixed relative amount $w_i \in [0, 1]$ of the capital in the i-th asset (short sales being excluded), so that $\sum_{i=1}^{D} w_i = 1$ (the portfolio is fully invested). The return of the corresponding portfolio at time t is then

$$R_w(t) = \sum_{i=1}^{D} w_i r_i(t). \tag{1}$$

Hence, if the $r(t)$'s are assumed to be i.i.d., so are the $R_w(t)$'s, with common distribution function F_w. We point out that, although in the case when one does not consider investment strategies involving short sales $R_w(t)$ is bounded below by -1 (like the $r_i(t)$'s), here we classically use an infinite lower tail approximation for modelling extreme lower values of portfolio returns (*i.e.* the left tail of the df F_w). Any risk measure is classically defined as a functional of the portfolio return distribution F_w (a specific quantile or its variance for instance). In this paper, we focus on the maximal relative loss of the portfolio over a large period of time T, that is to say on $m_T = \min_{t=1,...,T} R_w(t)$, of which fluctuations may be characterized by an asymptotic extreme value distribution H in some cases, namely in the cases when F_w is in the domain of attraction of an extreme value distribution H with respect to its lower tail. It is a well-known result in extreme values theory that there are only three types of possible limit distributions for the minimum of i.i.d. r.v.'s under positive affine transformations, depending on the tail behavior of their common density (refer to Resnick (1987) for further details on extreme values theory).

Here we shall consider investment strategies w with distribution functions F_w in the maximum domain of attraction of Fréchet distribution functions Φ_α, $\alpha > 0$ ($MDA(\Phi_\alpha)$ in abbreviated form), which dfs form the prime examples for modelling heavy-tailed phenomena (see Chap. 6 in Embrechts *et al.* (1997) for instance and the references therein). Recall that

$$\begin{aligned} \Phi_\alpha(x) &= \exp(-x^{-\alpha}), \text{ for } x > 0, \\ \Phi_\alpha(x) &= 0, \text{ for } x \leqslant 0, \end{aligned}$$

and that a df $F \in MDA(\Phi_\alpha)$ iff $F(-x)$ is regularly varying with index $-\alpha$, that is $F(-x) = x^{-\alpha}L(x)$, for some measurable function L slowly varying at ∞ (i.e. for some function L such that $L(tx)/L(x) \to 1$ as $x \to \infty$ for all $t > 0$). In the case when the df F behaves as a power law distribution at $-\infty$, the *tail index* α characterizes the extreme lower behaviour of an i.i.d. sequence $(R(t))_{t \geqslant 0}$ drawn from F regarding to its minimum value, in the sense that: for any $x > 0$, $\mathbb{P}(c_T^{-1}(\min_{t=1,\dots,T} R(t)) \leqslant -x) \to 1 - \Phi_\alpha(x)$ as $T \to \infty$, where $c_T = \sup\{x \in \mathbb{R} : F(-x) \leqslant n^{-1}\}$ (recall that $\xi = \alpha^{-1}$ is also known as the *shape parameter* of the extreme value df in this case). As shown in several empirical studies (see for instance Hogg & Klugman (1984) for such a modelling in the domain of insurance, and Guillaume *et al.* (1997), Longin (1996), Loretan & Phillips (1994) or Pisarenko & Sornette (2004) for empirical studies assessing the pertinence of such assumption for daily log-returns in finance and in Gabaix *et al.* (2003) a testable theory for the origin of power law tails in price fluctuations is proposed) and illustrated in Fig. 1 by the *quantile plot* and the *mean excess plot* related to the returns of the portfolio obtained by allocating 25% of the capital in each of the following financial indexes, Dow Jones, IGPA, TSX and Taiwan SE index from 01/1987 to 09/2002 (the pertinence of *QQ-plots* and *ME-plots* as exploratory graphical tools for extremes is discussed in § 6.2 in Embrechts *et al.* for instance), this class of dfs F contains left heavy-tailed distributions, usually called *generalized Pareto* or *power law-like* dfs, that may be appropriate for modelling large lower fluctuations of returns. Let us observe that the smaller the tail index α is, the heavier the left tail of F is. Hence, when modelling the lower tail behaviour of the distribution of portfolio returns this way, the tail index α may appear as a legitimate measure of extreme risk for the portfolio strategy (refer to Hyung & de Vries (2005) for a thorough discussion about the relevance of this specific safety first criterion, when managing the downside risk of portfolios is the matter).

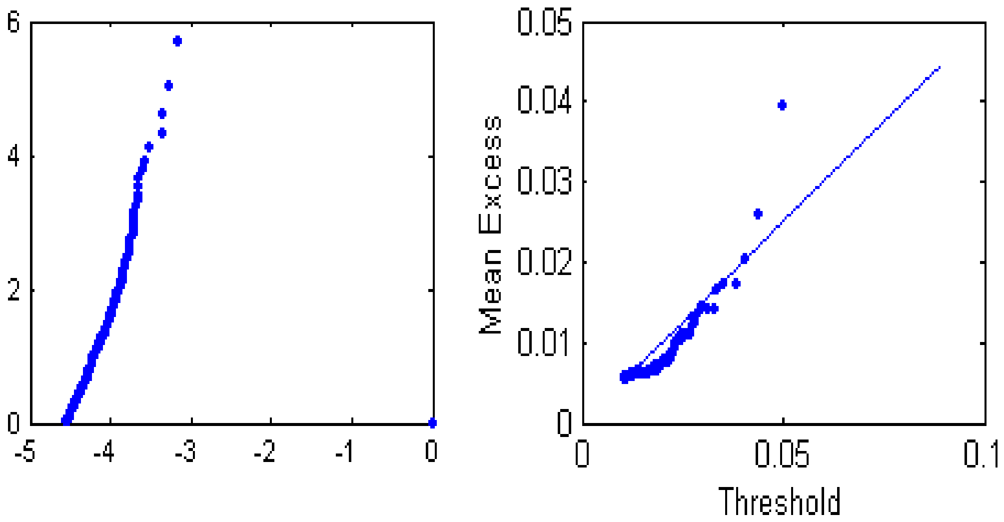

Fig. 1: Graphical analysis of returns data for a uniform portfolio: (a) QQ-plot of the lower returns data *vs* the Pareto df on $\mathbb{R}^-$ with tail index 2 (b) the empirical ME function of the lower returns data compared with the theoretical ME function of the Pareto df on $\mathbb{R}^-$ with tail index 2.

Remark 2.1 *Moreover, the tail index* α *rules the asymptotic behaviour of the excess-of-loss df* $F^{(u)}(x) = \mathbb{P}(R < -u - x \mid R < -u)$ *below the threshold* $-u$ *related to* $F(x) = \mathbb{P}(R \leqslant x) \in MDA(\Phi_\alpha)$ *for large* $u > 0$, *as shown by the following limit distributional approximation (see de Haan (1984) for further details on the convergence of normalized excesses over large thresholds to Generalised Pareto Distributions): for* $1 + x/\alpha > 0$,

$$\lim_{u \to \infty} F^{(u)}(xa(u)) = (1 + \frac{x}{\alpha})^{-\alpha},$$

where $a(u)$ *is a measurable positive function such that* $a(u)/u \to \alpha^{-1}$ *as* $u \to \infty$.

The set of strategies w of which returns have dfs F_w in $MDA(\Phi_\alpha)$ is thus naturally equipped with a complete preference relation, as follows.

Definition 2.1 *Suppose that* w_1 *and* w_2 *are two portfolio strategies such that* $F_{w_i} \in MDA(\Phi_{\alpha_i})$ *with* $\alpha_i > 0$ *for* $i = 1, 2$. *We shall say that strategy* w_1 *is riskier (respectively, strictly riskier) than strategy* w_2 *regarding to extreme relative loss iff* $\alpha_1 \leqslant \alpha_2$ *(resp.,* $\alpha_1 < \alpha_2$*).*

3 ICA: some basics.

Multivariate data are often viewed as multiple indirect measurement of underlying factors or components, which cannot be directly observed. In some cases, one may hope that a few of these latent factors are responsible for the essential structure we see in the observed data and correspond to interpretable causes. *Latent Variables Analysis* aims to provide tractable theoretical framework and develop statistical methods for identifying these underlying components. The general setting of latent variables modelling stipulates that

$$X = AY, \tag{2}$$

where X is an observable D-dimensional random vector and Y is a vector of d unobserved latent variables converted to X by the linear transform A, classically called the *mixing matrix.* Note that in the case when $d \leqslant D$ and the mixing matrix is of full rank d, there is no loss (but some redundancy on the contrary, when $d < D$) in the information carried by X. As a matter of fact, one may then invert the relation (2) and write

$$Y = \Omega X, \tag{3}$$

where the *de-mixing matrix* Ω is any generalized pseudo-inverse of A: whereas in the case $d = D$, Ω is uniquely determined by (3) and is simply the inverse A^{-1} of the mixing matrix, additional identifiability constraints are necessary for guaranteeing unicity when $d < D$. *Principal Component Analysis* (*PCA*) and *Factor Analysis* form part of a family of statistical techniques for recovering Y and A on the basis of an observed sample drawn from X, that are typically designed for normal distributions, which assumption clearly limited practical application of such modellings. In spite of the considerable evidence for non-gaussianity of financial returns or log-returns, and on the basis of theoretical market modelling such as CAPM or APT, Factor Analysis and PCA have been nevertheless extensively used by practicioners for gaining insight to the explanatory structure in observed returns and finding interpretable factors such as market confidence that may be hard to measure. In the mid-90's, the methodology of *Independent Component Analysis* (refer to Hyvärinen *et al.* (2001) or Everson & Roberts (2001) for a comprehensive presentation of ICA), comprising highly successfull new algorithms (mainly introduced in the field of signal processing for *Blind Source Separation* (BSS)) has emerged as a serious competitor to PCA and Factor Analysis, and is based, on the contrary to these latter techniques, on the non-normal nature of the latent components. As a matter of fact, ICA only relies on the crucial assumption that the underlying factors Y_1, $Y_2, \ldots$ are statistically independent (and are naturally called *independent components* for this reason), which hypothesis is much stronger than uncorrelation when Y is nongaussian (heuristically, uncorrelation determines second-order cross moments, whereas statistical independence determines all of the cross-moments). Stipulating further identifiability conditions apart from independence and non-normality of the IC's, many valid statistical methods have been proposed for estimating the ICA model, based on entropy, maximum likelihood, mutual information or tensorial methods (see chapters 8 to 11 in Hyvärinen *et al.* (2001)). The nongaussian character of financial returns or log-returns being now carried unanimously, several contributions to the application of ICA to finance have been recently made. Finance as a field of application of ICA is now under intensive investigation, so that no exhaustive list of works in that direction can be displayed at present (the interested reader is advised to consult *www.cis.hut.fi/projects/ica* or *www.tsi.enst.fr/icacentral* for a listing of the papers written so far). The modelling of the fluctuations of financial returns and the search for independent factors through ICA thus gave rise to several works, among which Back & Weigend (1997), Kiviluoto & Oja (1998) and Chan & Cha (2001) (see also Chapter 24 in Hyvärinen *et al.* (2001) and the references therein). Vassereau (2000) proposed a specific ICA model based on neural networks for determining factors generating financial returns in the U.S. market in the APT context and showed it provides an evaluation of these latter, sharper than by using classical factor analysis, especially when high idiosyncratic risks are present. In Chin *et al.* (1999) gaussian mixtures modelling is combined with ICA for computing the market risk of non-normal portfolios. Besides, Moody & Wu (1997) applied the ICA methodology in the context of state space models for interbank foreign exchange rates to improve the separation between noise and "true prices". Malaroiu *et al.* (2000) show how to apply ICA for forecasting financial series. And in Capobianco (2002), ICA combined with a *matching pursuit* algorithm is used as an exploratory tool for investigating the structure in (high-frequency) Nikkei index data. Besides, a study of implicit volatility dynamic through ICA is carried out in Ané & Labidi (2002) with the aim to identify the essential structure of time-varying volatility surfaces.

4 Our proposal: the heavy-tailed ICA model

We now present the specific ICA model we propose for describing the extreme lower fluctuations of asset returns, which we call the *heavy-tailed ICA model*, and give some insight in the latter. Here and throughout, we suppose there are D securities indexed by $i \in \{1, ..., D\}$ and let $X_i(t)$ be the price of the i-th security at time t. The returns of the i-th security are defined by

$$r_i(t) = (X_i(t) - X_i(t-1))/X_i(t-1),\ t \geqslant 1.$$

4.1 Modelling the extremal events via ICA

The hypotheses of the heavy-tailed ICA model are laid out as follows. Let us suppose that the flow of daily returns of the D assets are i.i.d. realizations of a random vector $r = (r_1, ..., r_D)'$ with components that are linear combinations of D independent *elementary portfolios* returns $R_1, ..., R_D$, so that

$$r = AR, \tag{4}$$

where $R = (R_1, ..., R_D)'$ and $A = (a_{ij})$ is a D by D matrix of full rank, of which inverse Ω belongs to the (compact and convex) set of parameters

$$\mathcal{B} = \{\Omega = (\omega_{ij})_{1 \leqslant i \leqslant, j \leqslant D} / \ \omega_{ij} \geqslant 0, \sum_{k=1}^{D} \omega_{ik} = 1 \text{ for } 1 \leqslant i, \ j \leqslant D\}. \tag{5}$$

We thus have

$$R = \Omega r. \tag{6}$$

We assume moreover that the R_i's have heavy-tailed distributions, furthermore lower-tails are supposed to be Pareto-like below some (unknown) thresholds:

$$G_i(y) = \mathbb{P}(R_i < -y) = C_i y^{-\alpha_i}, \text{ for } y \geqslant s_i, \tag{7}$$

with strictly positive constants α_i, C_i and s_i, $1 \leqslant i \leqslant D$. In addition, we suppose that the α_i's are distinct and, with no loss of generality, that the IC's are indexed so that $\alpha_1 > ... > \alpha_D$ (the elementary portfolios are thus sorted by increasing order of their riskiness regarding to Definition 1), so as to ensure that the statistical model is identifiable.

In this framework, we have the following result, which shows that an optimal strategy with respect to extreme relative loss is straightforwardly available from our specific ICA model.

Theorem 4.1 *The elementary portfolio strategy $\omega_1 = (\omega_{11}, ..., \omega_{1D})$ is optimal with respect to the extreme risk measure.*

Proof. Theorem 4.1 straightforwardly derives from the following lemma concerning convolution products of regularly varying densities (refer to Embrechts, Goldie & Veraverbeke (1979) for the technical proof).

Lemma 4.1 *Suppose that F_1 and F_2 are probability distributions such that $1 - F_1(x) = o(1 - F_2(x))$ as $x \to \infty$, then $1 - F_1 * F_2(x) \sim 1 - F_2(x)$ as $x \to \infty$.*

Therefore, according to the proposed ICA model, any convex linear combination $R_w = wr$ of the asset returns, is also a linear combination of the elementary portfolio returns ($R_w = \widetilde{w}R$ with $\widetilde{w} = Aw$ and notice that necessarily $\sum_{1 \leqslant i \leqslant D} \widetilde{w}_i = 1$). The result thus immediately follows from the assumptions related to the independence structure of the R_i's and their tail behaviour. ∎

Remark 4.1 *We emphasize that obvious modifications of the heavy-tailed ICA model described above permits to deal with stakes of completely different nature, namely to find strategies w maximizing extreme relative profits* $\max_{t=1,...,T} R_w(t)$, *which issue may be somehow considered as "in duality" with the problem adressed in this article and typically concerns market participants such as hedge funds. In this dual setting, the upper tail behaviour of portfolio dfs F_w are modelled by Pareto-like distributions $\mathbb{P}(R_w(t) > x) = 1 - F_w(x) = L(x)x^{-\alpha}$, with $\alpha > 0$ and L a slowly varying function at infinity: the smaller the tail index α is, the more frequent extreme profit values are encountered: we then have $\mathbb{P}(c_T^{-1} \max_{t=1,...,T} R_w(t) \leqslant x) \to \Phi_\alpha(x)$ as $T \to \infty$, for any $x > 0$, where $c_T = \inf\{x \in \mathbb{R} : F(x) \geqslant 1 - n^{-1}\}$. An ICA model for identifying preferable investment strategies regarding to extreme profit may be derived by simply replacing condition* (7) *by the assumption that the R_i's are distributed so that $\mathbb{P}(R_i > x) = C_i x^{\alpha_i}$, for x over some unknown threshold s_i. Experimental studies based on this model will be carried out in a forthcoming paper (see also Skander (2005)).*

4.2 Statistical inference

Several objects must be estimated, the tail indexes $\alpha_1, ..., \alpha_D$, the constants $C_1, ..., C_D$, as well as the matrix Ω of elementary strategies (the inverse of the mixing matrix A of the ICA model). The statistical method we shall now detail for the parametric ICA model described above is based on conditional MLE, as the classical Hill inference procedure for tail index estimation (refer to Hill (1975)), which is widely used in applications related to risk assessment (see Zajdenweber (1996) or Koedijk & Kool (1992) for instance).

4.2.1 Conditional likelihood of the heavy-tailed ICA model

Let us first introduce some additional notation. For $i \in \{1, ..., D\}$, denote by $R_i(1), ..., R_i(N)$ an i.i.d. sample drawn from R_i and by $R_i(\sigma_i(1)) \leqslant ... \leqslant R_i(\sigma_i(N))$ the corresponding order statistics. Recall that the basic *Hill estimator* for the tail index α_i based on this sample is:

$$\widehat{\alpha}_{i,k}^H = (\frac{1}{k}\sum_{l=1}^{k} \ln(\frac{R_i(\sigma_i(l))}{R_i(\sigma_i(k))}))^{-1}, \tag{8}$$

with $1 \leqslant k \leqslant N$ such that $R_i(\sigma_i(k)) < 0$, while C_i is estimated by $\widehat{C}_i = \frac{k}{N}(-R_i(\sigma_i(k)))^{\widehat{\alpha}_i}$. These estimates are (weakly) consistent, as soon as $k = k(N)$ is picked such that $k(N) \to \infty$ and $N/k(N) \to \infty$ as $N \to \infty$ (*cf* Mason (1982)) and are strongly consistent if furthermore $k(N)/\ln\ln N \to \infty$ as $N \to \infty$ (see Deheuvels *et al.* (1988)). They are classically interpreted as a conditional maximum likelihood estimators based on maximization of the joint density $f_{i,k}(y_1, ..., y_k)$ of $(-R_i(\sigma_i(1)), ..., -R_i(\sigma_i(k)))$ conditioned on the event $\{R_i(\sigma_i(k)) \leqslant -s_i\}$:

$$f_{i,k}(y_1, ..., y_k) = \frac{N!}{(N-k)!}(1 - C_i y_k^{-\alpha_i})^{N-k} C_i^k \alpha_i^k \prod_{l=1}^{k} y_l^{-(\alpha_i+1)}, \tag{9}$$

for $0 < s_i \leqslant y_1 \leqslant ... \leqslant y_k$ and $f_{i,k}(y_1, ..., y_k) = 0$ otherwise. Hence the conditional likelihood based on R is

$$\prod_{i=1}^{D} f_{i,k}(-R_i(\sigma_i(1)), ..., -R_i(\sigma_i(k))). \tag{10}$$

Remark 4.2 *We would like to stress at this point that the practical choice of the number k of (lower) order statistics used is crucial when implementing the Hill estimation procedure. The so-called Hill-plot $\{(k, \widehat{\alpha}_{i,k}^H) : 2 \leqslant k \leqslant N\}$ is a graphical tool, that may be used efficiently for solving practically this problem. Regarding our framework, this issue is discussed in § 4.2.3.*

Now it is not difficult to derive the conditional likelihood of the heavy-tailed ICA model from the observation of a sample of length N of asset returns $r_{(N)} = (r(1), ..., r(N)) = ((r_i(1))_{1\leqslant i\leqslant D}, ..., (r_i(N))_{1\leqslant i\leqslant D})$. For all $1 \leqslant i \leqslant D$, sort the return vector observations $r(l)$, $1 \leqslant l \leqslant N$, so that $\omega_i r(\sigma_i(1)) \leqslant ... \leqslant \omega_i r(\sigma_i(N))$ (observe that the permutation σ_i depends on ω_i: $\omega_i r(\sigma_i(l)) = R(\sigma_i(l))$, $1 \leqslant l \leqslant N$). Hence, the likelihood function based on the observations $\{r(\sigma_i(l)),\ 1 \leqslant l \leqslant k,\ 1 \leqslant i \leqslant D\}$ and conditioned on the event $\{\omega_i r(\sigma_i(k)) \leqslant -s_i,\ 1 \leqslant i \leqslant D\}$ is

$$L_k(r_{(N)}, \Omega, \alpha, C) = |\det \Omega|^k \prod_{i=1}^{D} f_{i,k}(-\omega_i r(\sigma_i(1)), ..., -\omega_i r(\sigma_i(k))). \qquad (11)$$

Note that, for any given $r_{(N)}$, the functional $L_k(r_{(N)}, .\,, .\,, .)$ is continuous and piecewise differentiable on $\mathcal{B} \times \mathbb{R}_+^* \times \mathbb{R}_+^*$. Furthermore, as previously recalled, for any fixed $\Omega \in \mathcal{B}$ and for all $i \in \{1, ..., D\}$, $f_{i,k}(-\omega_i r(\sigma_i(1)), ..., -\omega_i r(\sigma_i(k))$ is maximum for $\alpha_i = \widehat{\alpha}_i$ and $C_i = \widehat{C}_i$ with

$$\widehat{\alpha}_i = (\frac{1}{k}\sum_{l=1}^{k} \ln(\frac{\omega_i r(\sigma_i(l))}{\omega_i r(\sigma_i(k))}))^{-1}, \qquad (12)$$

$$\widehat{C}_i = \frac{k}{N}(-\omega_i r(\sigma_i(k)))^{\widehat{\alpha}_i}. \qquad (13)$$

For any $\Omega \in \mathcal{B}$, $L(r^{(N)}, \Omega, \alpha, C)$ is thus maximum for $\alpha = \widehat{\alpha} = (\widehat{\alpha}_i)_{1 \leqslant i \leqslant D}$ and $C = \widehat{C} = (\widehat{C}_i)_{1 \leqslant i \leqslant D}$ and we denote this maximal value by

$$\widetilde{L}_k(\Omega) = L_k(r^{(N)},\ \Omega,\ \widehat{\alpha},\ \widehat{C}). \qquad (14)$$

Here, conditional MLE reduces then to maximizing the multivariate scalar function $\widetilde{L}_k(\Omega)$ over $\Omega \in \mathcal{B}$, which may be easily shown as equivalent to maximizing over $\Omega \in \mathcal{B}$:

$$l_k(\Omega) = |\det \Omega|^k \exp\left(-\sum_{i=1}^{D}\{k \ln(\sum_{l=1}^{k} \ln(\frac{\omega_i r(\sigma_i(l))}{\omega_i r(\sigma_i(k))})) + \sum_{l=1}^{k} \ln(-\omega_i r(\sigma_i(l))\}\right). \qquad (15)$$

4.2.2 Algorithms for conditional MLE

In our setting, estimating the ICA model (6) thus boils down to the task of finding $\widehat{\Omega}$ in the closed convex set $\mathcal{B}$ such that

$$l_k(\widehat{\Omega}) = \max_{\Omega \in \mathcal{B}} l_k(\Omega).$$

In addition to the theoretical estimation principle described above, a numerical method for maximizing the objective function $l_k(\Omega)$ (or $\widetilde{L}_k(\Omega)$) subject to the linear matrix constraint $\Omega \in \mathcal{B}$ is required. Various optimization methods, among which the popular *subgradient-type learning algorithms,* have been introduced for solving approximatively such a constrained optimization problem from a practical viewpoint (refer to Shor (1985) or Kiwiel (1985) for an extensive discussion of several classes of algorithms). As a matter of fact, the objective function $l_k(\Omega)$ is continuous and piecewise differentiable on $\mathcal{M}_D(\mathbb{R})$: its gradient $\nabla l_k(\Omega)$ is well-defined at each point $\Omega = (\omega_{i,j})_{1\leqslant i,j\leqslant D}$ such that $\det\Omega \neq 0$ and $\omega_i r(\sigma_i(k-1)) < \omega_i r(\sigma_i(k)) < \omega_i r(\sigma_i(k+1))$ for all $i \in \{1, ..., D\}$, we have

$$\begin{aligned} \frac{\partial l_k}{\partial \omega_{i,j}}(\Omega)/l_k(\Omega) &= k\gamma_{i,j}(\Omega)/\det\Omega + \sum_{l=1}^{k} r_j(\sigma_i(l))/\omega_i r(\sigma_i(l)) \\ &\quad - k\frac{\sum_{l=1}^{k}\{r_j(\sigma_i(l))/\omega_i r(\sigma_i(l)) - r_j(\sigma_i(k))/\omega_i r(\sigma_i(k))\}}{\sum_{l=1}^{k}\ln(\omega_i r(\sigma_i(l))/\omega_i r(\sigma_i(k)))}, \end{aligned} \tag{16}$$

where $\gamma_{i,j}(\Omega)$ denotes the cofactor of $\omega_{i,j}$ in Ω, $1 \leqslant i,\ j \leqslant D$ (the subdifferential $\partial l_k(\Omega)$ is then easily determined through equation (16) at any point $\Omega \in \mathcal{M}_D(\mathbb{R})$). Hence, the classical *projected subgradient method* (see Chap. 6 in Bertsekas (1995) for instance) allows to estimate the Heavy-tailed ICA model and to solve numerically the portfolio selection problem knowing the sample $r_{(N)}$, as illustrated by the following simple example, using two independent components with Pareto distributions.

Example (mixtures of two independent Pareto distributions) In this experiment (see graphpanel in Fig. 2), data are two mixtures of independent Pareto r.v.'s on $]-\infty, -1[$ with respective parameters $\alpha_1 = 4$ and $\alpha_2 = 3$, where the coefficients of the de-mixing matrix are $\omega_{1,1} = 1-\omega_{1,2} = 0.7$ and $\omega_{2,1} = 1 - \omega_{2,2} = 0.1$. Figures 2a) and 2b) show the two independent underlying i.i.d. series (sources) of length $N = 1000$. The observed bivariate data series is plotted in Fig. 2c) together with the directions of the IC's and the likelihood function is displayed in Fig. 2d). As shown by Fig. 2e), the projected gradient algorithm converged correctly to the maximum likelihood solution in $m = 38$ iterations: the resulting estimates are $\widehat{\alpha}_1 = 3.8631$, $\widehat{\alpha}_2 = 2.8828$, $\widehat{\omega}_{1,1} = 0.6995$ and $\widehat{\omega}_{2,1} = 0.0822$.

4.2.3 Practical considerations

When applying the estimation algorithm above on financial data, practical questions naturally arise and must be handled. We now discuss the latter issues.

On choosing the number k of lower order statistics As for calculating Hill estimates, the estimation method proposed in § 4.2.1 for the Heavy-tailed ICA model requires to choose the number k of lower order statistics used in the likelihood computation. It is well-known that the Hill estimator is very sensitive to the choice of k. This problem has been the subject of intense

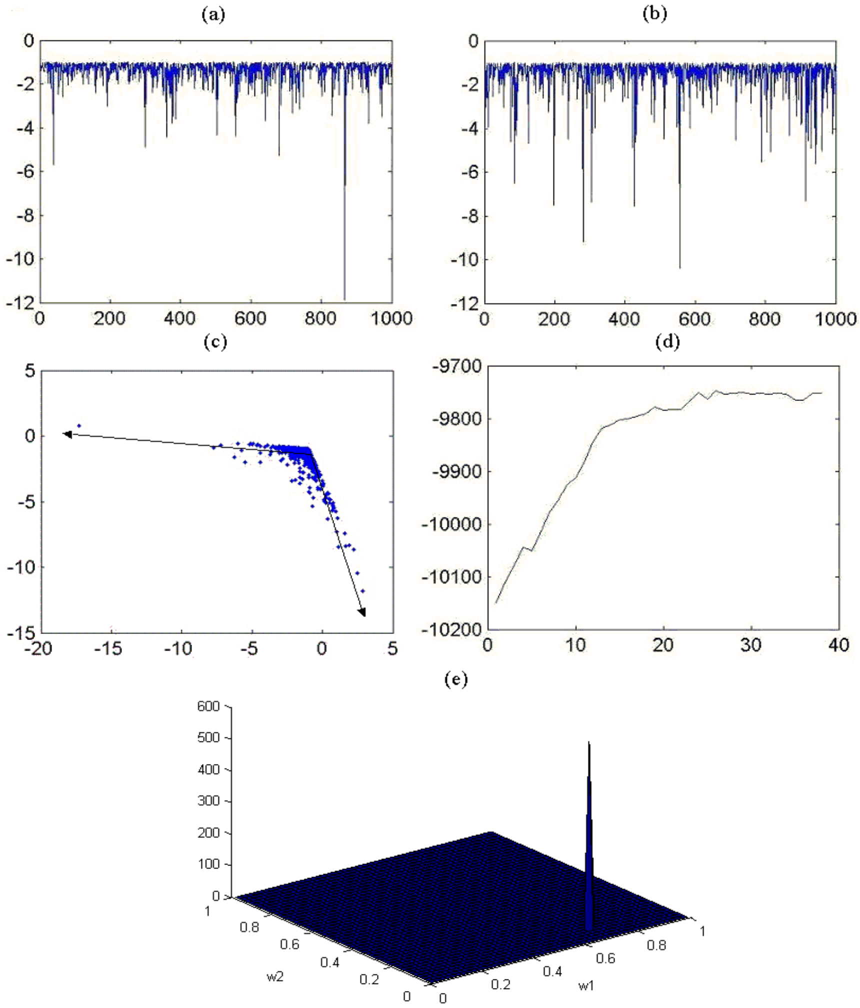

Fig. 2: Heavy-tailed ICA model estimation from simulated data: (a) and (b) the two underlying independent Pareto sources with tail indexes 4 and 3 , (c) observed mixed signal with the IC directions, (d) the likelihood value shown as a function of iteration count, (e) plot of the likelihood function.

research in theoretical statistics, still developping. For instance, one may try to pick an optimal k that minimizes the asymptotic mean square error (refer to Danielsson *et al.* (2001) or Drees & Kaufmann (1998)). However, in practice one either uses as a rule of thumb the lower 5% say of the sample in the calculation or else plots the Hill estimate based on k lower order statistics against k and find a suitable k in a stable region of the graph. In our framework, given the strong technical difficulties for studying the limit distribution of the estimator $\widehat{\alpha} = (\widehat{\alpha}_1, ..., \widehat{\alpha}_D)$ (since the latter depends not only on the observed data but also on the estimate $\widehat{\Omega}$) as n and k tend to ∞, we preferably use empirical graphical methods for selecting k. As in most cases encountered in practice the aim of such a modelling is to determine an optimal investment strategy ω_1 and the corresponding risk level α_1, a possible heuristic method would then consist in implementing the estimation procedure described in § 4.2.1 for several choices of k and plotting the maximum tail index $\widehat{\alpha}_1$ against k only and infering then a value of k such that the maximum tail index estimate appears to be stable.

On-line implementation In practice, financial data are available "on-line": asset returns keep on being observed daily and the size of the available sample $r_{(N)} = (r(1), ..., r(N))$ grows thus with time. In such a setting, it is naturally possible to adapt the estimation method described in § 4.2.2 and run an on-line version of the latter by indefinitely pursuing the optimization procedure and using at each step the whole available data set for updating the calculation of the conditional likelihood and its subgradient. Furthermore, the underlying data generating process might also evolve through time (the mixing-matrix and the tail parameters might be slowly varying) and fast tracking is needed then: an on-line adaptive version of the projected subgradient algorithm may be implemented by using a forward rolling data history of fixed length $m \geqslant 1$, *i.e.* by performing each step of the algorithm with the m latest observed values $r(N-m+1), ..., r(N))$ for tracking these structural changes (see § 5.3 below).

5 Applications - Empirical studies

We now carry out several applications of the Heavy-tailed ICA modelling, illustrating in particular its role in portfolio selection and asset allocation when the aim is to manage the downside risk.

5.1 Two examples

As explained above, the statistical method we present here aims to search for independent portfolio strategies and to estimate their Pareto left tail indexes as well. It also allows us to recover as a byproduct a portfolio strategy with a maximum left tail index (see Theorem 4.1).

Example 1. As a first illustration of the Heavy-tailed ICA model, we apply the latter for analyzing the daily return series of $D = 11$ international equity indexes of (developed or developing) financial markets over the period running from 02-January-1987 to 22-October-2002 (the length of the financial time series studied is thus $N = 4096$) listed in Table 1.

Table 1. Extreme value statistics based on daily returns from 11 international equity markets for the period 02/01/87 - 22/10/02. For each market index, the number k of values used for computing the Hill estimate $\widehat{\alpha}_g$ of the left tail index is selected using the AMSE approach as in Danielsson *et al.* (1997). The empirical Mean Excess function is computed at levels 1%, 2%and 3%.

Country (Index)	k	$\widehat{\alpha}_g$	Mean Excess (%)			Min (%)
			1%	2%	3%	
Canada (TSX)	99	2.77	0.81	1.08	1.43	-11.32
Chile (Igpa)	152	2.93	0.68	0.86	1.18	-12.50
Germany (Dax30)	150	2.98	0.94	1.12	1.60	-13.71
Hong Kong (Hang Sang)	195	2.43	1.10	1.17	1.43	-40.54
Korea (Kospi)	196	2.55	1.32	1.44	1.45	-12.80
Japon (Nikkey225)	165	3.44	1.02	0.97	1.02	-16.14
Malaisia (KLSE)	240	2.08	1.25	1.66	1.97	-24.15
Singapore (Straits Times)	193	2.41	1.03	1.42	2.16	-29.19
Taiwan (SE)	200	2.95	1.55	1.51	1.57	-10.29
U.S. (Nasdaq Comp.)	141	3.09	1.29	1.29	1.36	-12.05
U. K. (FTSE100)	118	2.91	0.77	1.05	1.02	-13.03

The results presented in Table 2 indicate the allocations corresponding to the 11 *independent elementary portfolios*, as well as their left tail index estimate, sorted by increasing order of their extreme risk measure (as quantified in section 2), obtained by implementing the statistical procedure described in § 4.2 with the $k = 200$ lowest values (representing roughly the lowest 5% values). Descriptive statistics related to the lower tail behaviour of each elementary portfolio are also displayed in Table 2: minimum return values, empirical estimates of the probability of excess (*EPE* in short), $\mathbb{P}(R_i < -u)$, that the i-th elementary portfolio looses more than u% of its value (at a one day horizon) are calculated at various threshold levels u over the time period considered, as well as the empirical counterpart of the mean excess function (*ME* in abbreviated form), $e_i(u) = -\mathbb{E}(R_i + u \mid R_i < -u)$, traditionally referred to as the *expected shortfall* in the financial risk management context. For comparison purpose, inferential and descriptive statistics concerning the lower tail of single equity indexes are also given in Table 1. In a general fashion, by looking at these indicators one can see that the lower tails of the single assets are globally much heavier than the ones of the least risky elementary portfolios we obtained. For instance, the maximum relative loss suffered by the optimal elementary portfolio (PF1) over the period of interest is 3.86% , while the minimum values of single market indexes range from

-10.29% to -40.54%. As expected, except for the FTSE100 index, zero or small weights in PF1 correspond to the market indexes with the most heavier left tails (namely, the Hong Kong, Malaisia and Singapore indexes, which correspond to emerging financial markets that are presumably very interdependent, for instance the plot in Fig. 3a) shows strong patterns of left-tail dependence for the returns of the Hong Kong and Singapore market indexes), whereas on the contrary the latter indexes have the largest weights for PF11. These results clearly show the benefit of the diversification induced by our specific modelling of the dependence structure between the assets regarding to extreme risk. This phenomenon is also illustrated by Table 3. It shows lower tail statistics of the optimal portfolio obtained by applying our modelling as a function of the number D of market indexes involved in the ICA model (financial indexes being progressively added, by decreasing order of their tail index estimates) and plainly indicates that the lower tail becomes thinner as D grows.

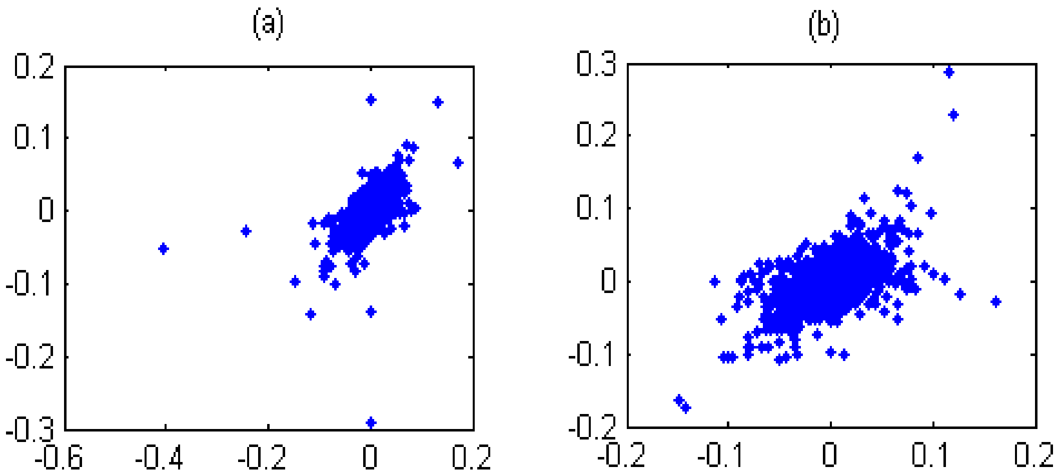

Fig. 3: Lower Tail Dependence of Returns Pairs (a) Singapore (vertical axis) and Hong Kong (horizontal axis) (b) Argentina (vertical axis) and Brasil (horizontal axis).

Table 2. Weights of the elementary portfolios are given under columns in percentages (estimate of the de-mixing matrix), together with their tail index estimates, the minimum return values over the time period considered, the standard deviation, and the EPE and EME computed at levels 1%, 2% and 3%.

	PF1	PF2	PF3	PF4	PF5	PF6	PF7	PF8	PF9	PF10	PF11
						Weights					
Markets											
Canada	7.57	20.64	0.00	4.66	21.58	19.35	23.92	12.88	0.00	6.34	0.00
Chile	24.44	0.00	13.61	8.61	9.20	0.00	20.47	0.00	7.45	7.42	12.52
Germany	17.12	22.96	3.89	28.28	3.53	0.00	0.00	14.43	13.20	16.12	0.00
HongKong	0.00	18.09	16.22	0.00	0.00	0.00	3.78	6.37	0.00	0.00	19.40
Korea	1.92	0.12	12.46	17.43	0.00	21.78	0.00	0.00	0.00	6.54	3.64
Japan	13.43	16.82	9.01	0.00	12.27	8.21	2.76	0.00	16.97	33.01	0.94
Malaisia	5.87	6.89	5.48	8.57	0.00	11.68	20.42	22.83	4.78	19.49	18.68
Singapore	0.00	3.26	0.00	18.77	22.36	9.46	7.38	0.00	8.14	11.08	24.21
Taiwan	11.54	0.00	21.67	0.00	18.28	3.81	0.00	23.35	2.79	0.00	0.78
U.S.	18.11	9.56	4.68	5.54	0.00	23.17	0.00	17.63	8.18	0.00	19.83
U.K.	0.00	1.66	12.98	8.14	12.78	2.54	21.27	2.51	38.49	0.00	0.00
Tail Index	**3.93**	**3.82**	**3.64**	**3.46**	**3.39**	**3.35**	**3.27**	**3.19**	**3.03**	**2.92**	**2.84**
Min (%)	-3.86	-5.89	-8.34	-6.13	-6.35	-9.80	-9.38	-8.47	-7.93	12.31	-9.72
Std (%)	0.58	0.67	0.80	0.75	0.72	0.68	0.66	0.74	0.65	0.78	0.80
						Mean Excess at threshold level u					
u=1%	0.44	0.58	0.58	0.54	0.62	0.58	0.58	0.57	0.59	0.60	0.71
u=2%	0.73	0.75	0.77	0.70	0.79	1.06	0.88	0.82	1.12	0.87	1.08
u=3%	0.46	1.13	1.38	1.05	0.86	1.51	2.07	1.30	1.91	2.03	1.49
						Probability of Excess at threshold level u					
u=1%	0.083	0.105	0.177	0.157	0.139	0.112	0.101	0.138	0.100	0.157	0.154
u=2%	0.008	0.017	0.026	0.021	0.022	0.015	0.015	0.021	0.015	0.025	0.031
u=3%	0.034	0.004	0.005	0.048	0.005	0.053	0.003	0.005	0.004	0.005	0.010

Table 3. Lower tail characteristics of the optimal portfolio obtained by using the Heavy-tailed ICA model with Dmarket indexes, as Dgrows.

Number of assets D	3	5	7	9
Pareto Index	2.57	2.90	3.04	3.30
Minimum (%)	-36.54	-14.60	-7.25	-6.71
EME at $u = 1\%$	1.27	0.92	0.58	0.56
EME at $u = 2\%$	1.53	1.09	0.73	0.88
EME at $u = 3\%$	1.98	1.39	1.39	1.19

Example 2. As a second illustration, we also applied our method to the 11 international equity indexes, of developing markets only, listed in Table 4 over the period running from 16-December-1994 to 22-October-2002 (here the time series length is $N = 2048$), the model being fitted by conditional

MLE with the $k = 150$ lowest values.

Table 4. Extreme value statistics based on daily returns of the 11 international equity markets involved in Example 2 for the period 12/1994 - 10/2002.

Country (Index)	k	$\widehat{\alpha}_g$	Mean Excess (%)			Min (%)
			1%	2%	3%	
Argentina (Merval)	174	2.30	1.85	1.94	2.00	-14.76
Brazil (Bovespa)	139	2.27	1.79	1.84	1.95	-17.23
Chile (Igpa)	53	3.15	0.58	0.71	0.41	-3.86
China (Shangaï)	115	2.21	1.43	1.62	2.08	-17.91
Greece (Athens SE)	96	2.98	1.28	1.51	1.49	-9.69
Hong Kong (Hang Sang)	122	2.41	1.27	1.42	1.53	-14.73
Mexico (Ipc)	130	2.62	1.20	1.25	1.36	-14.31
Russia (RTS)	188	2.21	2.18	2.31	2.25	-19.02
S. Africa (JSE All Share)	99	3.28	0.91	1.23	1.47	-11.86
Tchekia (Prague SE)	95	3.68	0.88	0.81	0.89	-7.08
Taiwan (SE)	108	3.30	1.18	1.10	1.05	-9.94

As shown by Table 5, the maximum relative losses sustained by the elementary portfolios we obtained over the period studied range from 2.61% to 5.64% only, which are globally much lesser than the ones of the single market indexes, as the empirical Mean Excess values. One may notice for instance that the Argentina and Brazil market indexes have zero weights in the resulting optimal portfolio: these financial indexes are very heavy-tailed, as the Hong Kong index which is also zero weighted, on the one hand and the plot in Fig. 3b) strongly suggests that extreme values tend to occur simultaneaously for these two financial markets on the other hand.

5.2 Further comparisons

In order to highlight the solutions that may be obtained by using the specific modelling we propose for guarding investors from big losses, we first compared the results described above with the ones obtained by the standard *Mean-Variance* (MV) approach (*cf* Markowitz (1952)). When fixing the mean return target, the latter method consists then in solving a system of linear equations for finding the portfolio strategy with minimum variance.

Table 5. Weights of the elementary portfolios are given under columns in percentages (estimate of the de-mixing matrix), together with their tail index estimates, the minimum return values over the time period considered, the standard deviation, and the EPE and EME computed at levels 1%, 2% and 3%.

	PF1	PF2	PF3	PF4	PF5	PF6	PF7	PF8	PF9	PF10	PF11
						Weights					
Markets											
Argentina	0.00	0.00	19.74	11.13	7.41	17.00	15.92	0.00	0.00	11.21	8.70
Brazil	0.00	0.00	5.00	23.29	18.46	0.00	11.19	0.00	9.31	6.09	0.00
Chile	9.79	8.14	0.00	13.41	0.00	2.51	6.16	15.55	24.60	17.32	26.93
China	12.43	0.00	9.92	0.00	16.17	11.42	11.15	22.65	9.62	26.81	0.00
Greece	0.00	9.71	15.98	17.33	2.09	20.48	0.00	16.15	14.71	0.00	1.43
H. Kong	14.74	0.00	0.00	2.87	18.93	15.26	11.17	12.14	1.22	0.00	20.17
Mexico	6.74	17.17	14.20	0.00	13.62	10.76	6.07	0.00	29.84	8.70	6.09
Russia	10.12	16.51	0.00	6.36	0.00	3.71	27.31	6.32	5.88	0.00	0.00
S. Africa	22.62	7.42	26.21	9.38	5.36	0.00	0.00	20.60	3.37	0.00	9.71
Tchekia	0.00	25.50	4.61	0.00	14.57	3.25	11.02	6.60	0.00	9.74	21.85
Taiwan	23.56	15.55	4.35	16.23	3.40	15.61	0.00	0.00	1.45	20.13	5.11
Tail Index	**3.54**	**3.38**	**3.34**	**3.13**	**3.05**	**2.95**	**2.92**	**2.80**	**2.78**	**2.73**	**2.74**
Min (%)	-2.61	-3.90	-6.57	-6.59	-6.27	-4.96	-5.75	-4.57	-6.09	-4.58	-5.41
Std (%)	0.73	0.82	1.04	1.05	1.01	0.89	1.17	0.82	0.95	0.82	0.77
						Mean Excess at threshold level u					
u=1%	0.41	0.52	0.70	0.73	0.70	0.58	0.80	0.53	0.67	0.59	0.59
u=2%	0.26	0.54	0.68	0.78	0.82	0.71	0.94	0.57	0.81	0.54	0.90
u=3%	0.00	0.83	0.87	0.72	0.95	0.70	1.07	1.13	0.79	0.90	1.02
						Probability of Excess at threshold level u					
u=1%	0.163	0.198	0.286	0.279	0.273	0.230	0.311	0.192	0.226	0.186	0.168
u=2%	0.018	0.035	0.070	0.077	0.066	0.040	0.079	0.025	0.049	0.035	0.024
u=3%	0.000	0.002	0.014	0.025	0.021	0.010	0.024	0.004	0.015	0.006	0.007

More precisely, for both examples we calculated here the allocations corresponding to the portfolio with minimum variance among all portfolios with the same mean as the (estimated) one of the optimal portfolio resulting from the Heavy-tailed ICA model (which portfolio is to be refered to as the *MV portfolio* in what follows, while the optimal elementary portfolio is called the *ICA portfolio*): in the first example, the empirical mean return of the optimal elementary portfolio is 0.04%, while it is 0.02% in the second example. MV portfolios are displayed in Table 6. The performance of these portfolios regarding to huge losses may be compared by representing the lower tail of their empirical distributions (see the time-plots and the histograms in Fig.4) in the one hand and plotting their empirical mean excess (EME) functions in the other hand (see Fig. 5 and Table 7). In these experiments the Heavy-tailed ICA model clearly outperforms the classical MV approach: in both examples, the graph of the EME function of the elementary portfolio is always much below the one of the EME function of the MV portfolio and this phenomenon is more and more pronounced as the threshold level grows. This suggests that the Heavy-tailed ICA model describes much more pertinently the dependence structure of the lower fluctuations of the asset returns considered in these examples than the simple covariance structure, which essentially describes the spread of the distribution of the flow of returns about its mean value and does not take sufficiently into account the tail behaviour, given its nongaussian character. One can see for instance, that although the time plot in Fig. 6 clearly shows that in most cases extreme lower variations of the Merval index and the ones of RTS index do not occur simultaneously, the (linear) correlation between these return series is almost zero (namely, -0.86%), extreme lower values being not enhanced by the latter statistical indicator due to their rarity.

We also emphasize that many other models have been suggested for solving the portfolio selection problem, involving different measures of risk. Providing a systematic list of their respective advantages and limitations is beyond the scope of the present paper, but we nevertheless point out that a possible approach, could consist in fixing a (large) threshold loss u and implementing a greedy search algorithm for finding a portfolio strategy ω that minimizes the mean excess function at the point u, which objective function is $-\mathbb{E}(\omega r + u \mid \omega r < -u) = (\sigma + \xi u)/(1 - \xi)$, when modelling the lower tail of the portfolio return ωr by a *Generalized Pareto Distribution* $\mathbb{P}(\omega r < -z) = F_{\xi,\sigma,u}(z) = 1 - \mathbb{P}(\omega r < -u)(1 + \xi(z-u)/\sigma)_+^{-1/\xi}$, with $\xi \in (0,1)$ and $\sigma > 0$, in the same spirit as the approach developped by Bradley & Taqqu (2004) (except that the *Value at Risk* $Var_\omega(\alpha) = u + \sigma(((1-\alpha)/\mathbb{P}(\omega r < -u))^{-\xi} - 1)/\xi$ at a fixed level α is chosen as loss threshold), the shape and scale parameters ξ and σ being estimated by maximum likelihood for each possible allocation vector ω. We call such a portfolio the ES_u portfolio in the following. Beyond the practical difficulties encountered when performing such an optimization procedure, the resulting portfolio strategy is highly dependent of the loss threshold u chosen, as confirmed by our experiments. As a matter of fact, the EME plots in Fig. (a) and (b) and the results displayed in Table 7 show that, for a given threshold u_0, though the ES_{u_0} portfolio may have smaller empirical mean excess and probability of excess at level $u = u_0$ (in some cases only, namely for low threshold levels), its performance is rapidly becoming worse, compared to the ICA portfolio, when the risk level u increases. These findings suggest that, for guarding from catastrophic losses, optimizing globally the tail behaviour of the portfolio, or equivalently the asymptotic behaviour of the Mean Excess function (see Remark 2.1) as permitted by the Heavy-tailed ICA model may be preferable to performing a local optimization of the Mean Excess function at an arbitrary fixed level and lead to stabler results.

Table 6. MV and ES_u portfolio allocations (%). ES_u is minimized at threshold 1%, 2% and 3%.

Example 1					Example 2				
Country	MV	ES_u			Country	MV	ES_u		
		1%	2%	3%			1%	2%	3%
Canada	0.00	19.08	12.34	0.13	Argentina	5.76	0.00	0.00	0.00
Chile	48.62	23.45	31.62	31.13	Brazil	0.00	0.00	0.00	0.00
Germany	6.92	15.99	18.40	11.88	Chile	27.05	45.80	46.78	45.83
H. Kong	0.00	0.00	4.74	4.73	China	40.10	8.79	7.87	7.71
Korea	0.00	3.67	12.08	9.54	Geece	0.00	3.29	2.32	3.28
Japon	0.00	5.53	0.20	0.04	H. Kong	2.56	0.00	0.00	1.00
Malaisia	12.62	9.09	1.49	2.01	Mexico	1.98	0.00	0.00	0.00
Singapore	17.07	3.88	0.28	0.68	Russia	9.48	3.59	3.67	3.59
Taiwan	0.00	0.59	18.04	38.80	S. Africa	7.62	19.65	20.08	19.66
U.K.	2.97	7.87	0.76	0.02	Tchekia	5.45	12.38	12.64	12.42
U.S.	11.81	10.85	0.05	1.04	Taiwan	0.00	6.50	6.64	6.51

Table 7. Empirical Mean Excess and Probability of Excess over a threshold u_0 for portfolios obtained by direct ES minimization, Heavy-tailed ICA and MV optimization.

	Example 1			Example 2		
	ES_{u_0}	ICA	MV	ES_u	ICA	MV
EME at $u_0 = 1\%$	0.57	0.46	0.61	0.37	0.41	0.58
EPE at $u_0 = 1\%$	0.1910	0.0604	0.1121	0.0747	0.1636	0.1717
EME at $u_0 = 2\%$	0.57	0.54	1.19	0.39	0.26	0.63
EPE at $u_0 = 2\%$	0.0290	0.0067	0.0166	0.0070	0.0185	0.0384
EME at $u_0 = 3\%$	0.56	0.20	2.25	0.41	0.00	0.72
EPE at $u_0 = 3\%$	0.0041	0.0008	0.0043	0.0341	0.0000	0.0114

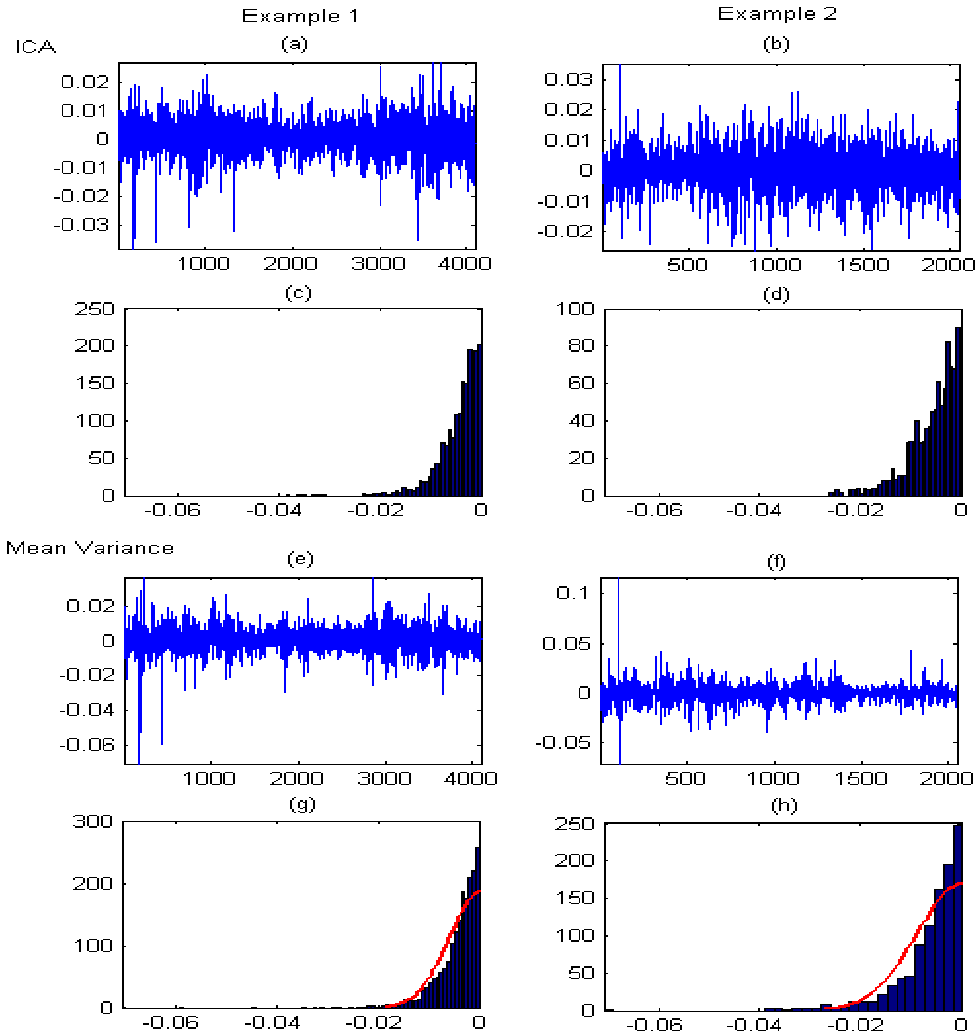

Fig. 4. Figures on the left (resp. on the right) are related to Example 1 (resp. to Example 2) (a), (b): Time plots of the returns of the ICA optimal portfolio. (c), (d): Histograms of the returns of the ICA optimal portfolio. (e), (f): Time plots of the returns of the MV portfolio. (g), (h): Histograms of the returns of the MV portfolio fitted to the gaussian density.

5.3 On-line implementation - Dynamic asset allocation

As mentioned in § 4.2.3, the dynamic of financial markets may evolve through time (economic structural changes, ruptures due to shocks, business cycles...). Hence it may appear as legitimate to make the parameters of the ICA Heavy-tailed possibly evolve for tracking eventual changes. Such a problem may be solved by taking a *sliding window approach.* As new values of financial returns keep on being observed each (opening) day, this method simply amounts to

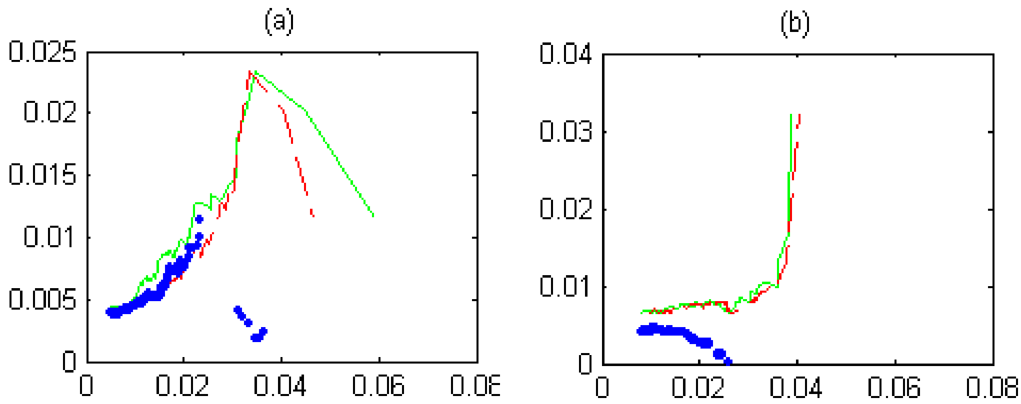

Fig. 5: (a) Example 1: Mean Excess of the ICA optimal strategy (black dotted line) compared to the ES_u portfolio relative to the threshold $u = 2\%$ (grey slashed line) and the MV portfolio (grey line). (b) Example 2: ICA versus ES_u ($u = 3\%$) and MV.

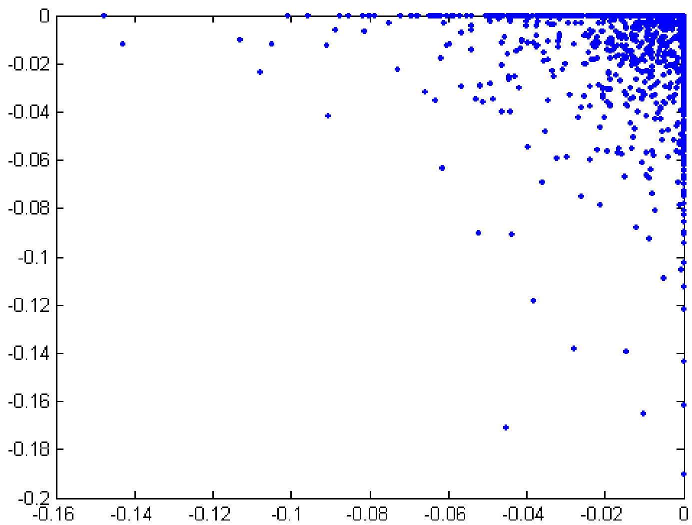

Fig. 6: Lowertail dependence: Merval Index (horizontal axis) versus RTS Index (vertical axis).

implement the estimation algorithm we proposed from the latest observed return values only, at time intervals of fixed length L (it would be naturally vain to attempt to track changes of the lower tail behaviour of the flow of returns each day). In the following illustration, we consider the problem of constructing a *safety-first* portfolio based on eight stock indexes (among which two indexes of developed markets and six indexes of emerging markets) and describe the performance of the investment strategy consisting in rebalancing the portfolio every $L = 120$ days (that is to say every six months roughly speaking) according to the optimal elementary portfolio computed from the least observed returns values (see the time-plot of the returns of the resulting portfolio in Fig. 7). By examining the results displayed in Table 8, one can see that some of the portfolio weights exhibit high degrees of variability, which does not necessarily affect the lower tail index estimate (for instance the weight of the Chile index ranges from 19.61% to 0% over the period 18/11/96-20/10/97, whereas the Pareto index is almost constant). Some economic considerations may provide helpful interpretation for these results. The significant decrease of the weights of the two developed market (mainly in favor of the Chile and Taiwan indexes) from the year 2001 may be attributed to the effects of the burst of the speculative bubble and the growing integration of the Chile market into the global financial market. It seems thus that this market has not been much affected by the "tequila crisis" in 1994, contrary to most emerging markets of South America (the effect of this crisis on the weight of the mexican market is obvious and has apparently lasted until new informations could restore the confidence). In a similar fashion, the Taiwan index seems to have escaped the effects of the asian crisis in 1997, which certainly explains why the two other asian market indexes correspond to very small weights since then.

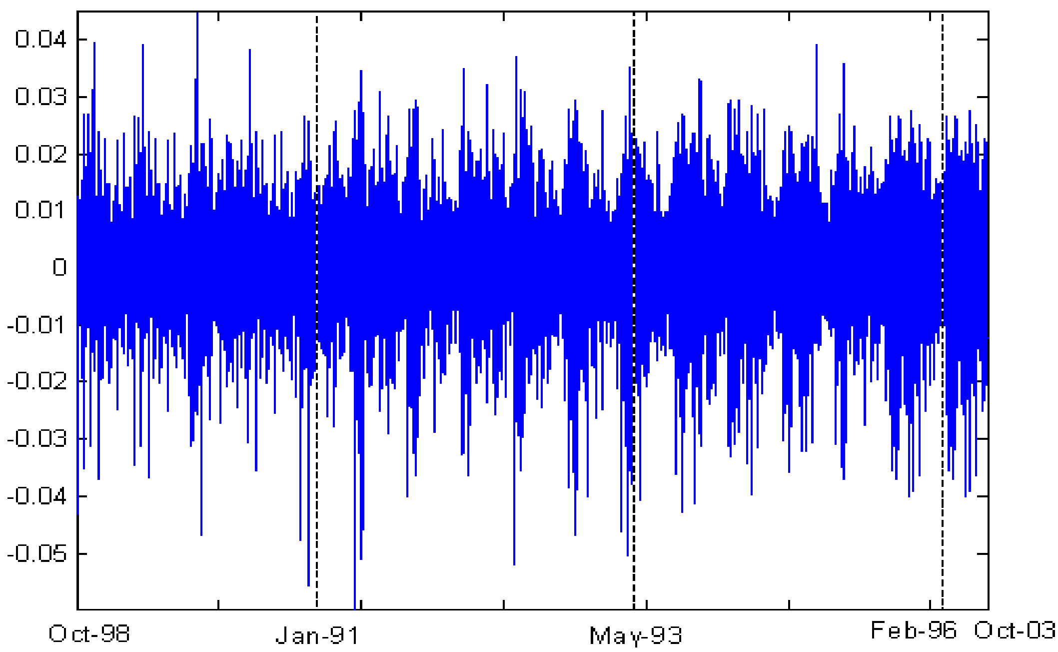

Fig. 7: Time plot of the returns of the dynamic ICA portfolio.

Table 8. Temporal evolution of the constitution of the safety-first portfolio constructed by rebalancing the capital every 120 days using the Heavy-tailed ICA model and of its lower tail index estimate.

Rebalancing date	Chile	Fra.	Gre.	H.K.	Korea	Mexi.	Tai.	U.S.	$\widehat{\alpha}$
04/10/1988	15.20	41.14	8.70	0.00	1.00	8.00	1.44	24.52	2.62
20/03/1989	19.61	29.40	2.00	12.88	11.97	5.00	3.00	16.14	3.11
04/09/1989	4.48	15.18	5.56	18.45	23.35	2.27	0.03	30.68	3.15
19/02/1990	0.00	13.96	20.99	11.05	9.66	5.46	8.00	30.88	3.13
21/01/1991	14.76	20.00	13.03	6.00	5.00	2.00	7.81	31.40	2.80
08/07/1991	21.64	17.75	14.48	4.88	2.00	5.41	0.32	33.52	2.60
23/12/1991	25.29	18.31	7.95	4.13	8.00	4.49	11.81	20.02	2.78
08/06/1992	22.54	8.36	6.21	11.29	3.00	10.00	19.55	19.05	2.83
23/11/1992	12.79	13.43	10.81	9.51	0.65	12.27	20.46	20.08	2.53
10/05/1993	16.56	18.04	4.86	5.88	2.12	14.32	28.18	10.04	3.19
25/01/1993	8.00	27.04	19.13	0.00	0.00	10.70	24.03	11.11	2.54
11/04/1994	28.70	14.84	7.41	0.00	0.00	12.42	17.58	19.05	2.67
26/09/1994	35.79	0.33	15.66	4.81	4.00	5.02	30.11	4.28	2.98
13/03/1995	46.07	10.07	11.07	0.47	0.00	0.00	21.50	10.82	3.12
28/08/1995	32.64	4.93	14.63	0.00	0.00	6.20	24.18	17.42	2.79
12/02/1996	37.37	0.00	26.30	1.01	0.00	0.00	31.32	4.00	3.04

6 Concluding remarks

Although we are far from having covered the application of the heavy-tailed ICA model to financial data in this paper, we endeavoured to present here enough material to illustrate the interest of the method. Now we conclude by pointing out several issues and sketching some lines of further research. First, in ICA applications, a common issue consists in determining the minimum number of independent components for explaining the data well enough (refer to Chap. 13 in Hyvärinen *et al.* (2001) for instance). In the application of the ICA methodology presented here, there is as much IC's as assets considered. And it is straightforward to extend our specific ICA model to the case when the number d of IC's is *a priori* known and smaller than the number D of assets. It would be thus interesting to develop a valid practical procedure for selecting an adequate value for d in the same fashion as methods based on Akaike, Bayesian or other information-theoretic criteria. This defines an ambitious direction for further investigation. Secondly, the parametrization of our ICA model concerns the lower tail behaviour of the IC's only, since we focussed here on the downside risk. Hence, another problem would consist in considering other parametrizations of the distributions of the IC's, so as to deal with different measures of risk, involving different features of the portfolio distribution and permitting pertinent trade-offs between potential profits and losses. Finally, we emphasize that the ICA model described in this paper is based on the observation of i.i.d. samples of linearly combined independent random variables. However, numerous statistical studies in the economic literature have exhibited a (possibly linear) dependence structure of financial returns both in time and across stocks and motivated intense research for modelling the latter in a time series framework and going past the i.i.d. assumption. Therefore several ICA methods have been recently developed for exploiting the time structure of data series, among which procedures based on local autocovariance estimates (see Belouchrani & Amin (1998) and Chap. 18 in Hyvärinen *et al.* (2001)). Hence a possible approach for analyzing returns data could consist in combining ICA to a method for local autocovariance estimation (as in Clémençon & Slim (2004) under the assumption of local stationarity), we leave this question for further research.

References

[1] Acerbi, C. Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking and Finance,* **26**, 1487-1503.

[2] Ané, T., Labidi, C. (2001). Implied Volatility Surfaces and Market Activity over Time. *Journal of Economics and Finance,* ***25,*** 3, 259-275.

[3] Arzac, E.R., Bawa, V.S. (1977). Portfolio choice and equilibrium in capital markets with safety-first investors. *Journal of Financial Economics,* **4**, 277-288.

[4] Back, A. D., Weigend, A. S. (1997). A First Application of Independent Component Analysis to Extracting Structure from Stock Returns. *Int. J. Neur. Syst.,* **8**, 473-484.

[5] Belouchrani, A., Amin, M. (1998). Blind source separation based on Time-frequency signal representations. *IEEE Trans. on Signal Processing, **46**, 11, 2888-2897.*

[6] Bertsekas, D.P. (1996). *Constrained Optimization and Lagrange Multiplier Methods.* Athena Scientific, Belmont, Massachussets.

[7] Capobianco, E. (2003). Independent Multiresolution Component Analysis and Matching Pursuit. *Comp. Statist. Data Analysis*, **42**, 385-402.

[8] Chan, L.W., Cha, S.M. (2001). Selection of independent factor model in finance. *Proceedings of the Third International Conference on Independent Component Analysis and Signal Separation.*

[9] Chavez-Demoulin, V., Embrechts, P. (2004). Smooth Extremal Models in Finance and Insurance. *J. Risk Insurance*, **71**, No 2, 183-199.

[10] Chin, E., Weigend, A. S., Zimmermann, H. (1999). Computing Portfolio Risk Using Gaussian Mixtures and Independent Component Analysis. *Proceedings of the IEEE/IAFE Conference on Computational Intelligence for Financial Engineering (CIFE)* , New York, March, 74-117.

[11] Clémençon, S., Slim, S. (2004). Statistical Analysis of Financial Time Series under the Assumption of Local Stationarity. *Quantitative Finance,* **4**, 208-220.

[12] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance*, **1**, 223-236.

[13] Crouhy, M., Galai, D., Mark, R. (2000). *Risk Management.* McGraw-Hill.

[14] Dacorogna, M.M., Müller, U.A., Pictet, O.V., de Vries, C.G. (2001). Extremal foreign exchange returns in extremely large data sets. *Extremes,* **4**, No. 2, 105-127.

[15] Danielsson, J., de Haan, L., Peng, L., de Vries, C.G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *J. Multiv. Analysis*, **76**, 226-248.

[16] Deheuvels, P., Haüsler, E., Mason, D.M. (1988). Almost sure convergence of the Hill estimator. *Math. Proc. Cambridge Philos. Soc.,* **104**, 371-380.

[17] Drees, H., Haan, L. de, Resnick, S. (2000). How to make a Hill-plot. *Ann. Stat.,* **28**, No. 1, 254-274.

[18] Drees H., Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stoch. Proc. App.*, Vol. **75**, No. 2, 149-172.

[19] Embrechts, P (2000). *Extremes and Integrated Risk Management.* London: Risk Books, Risk Waters Group.

[20] Embrechts, P., Goldie, C.M., Veraverbeke (1979). Subexponentiality and infinite divisibility. *Z. Warsch. Verw. Geb.,* **49**, 335-347.

[21] Embrechts, P., Klüppelberg, C., Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Springer Verlag.

[22] Embrechts, P, Resnick, D., Samorodnitsky, G. (1999). Extreme Value Theory as a Risk Management Tool. *North Amer. Act. Journ.,* **3**, 30-41.

[23] Everson, R., Roberts, S. (2001). *Independent Component Analysis: Principles and Practice.* Cambridge University Press.

[24] Gabaix, X., Gopikrishnan, P., Plerou, V., Stanley, E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, **423**, 267-270.

[25] Guillaume, D.M., Dacarogna, M.M., Davé, R.R., Müller, U.A., Olsen, R.B., Pictet, O.V. (1997). From the bird's eye to the microscope. A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics,* **1**, 95-129.

[26] Haan, L. de (1984). Slow variation and the characterization of domains of attraction. In *Statistical Extremes and Applications,* Ed. Tiago de Oliveira, Reidel, Dordrecht, 31-48.

[27] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.,* **3**, 1163-1174.

[28] Hyung, N., de Vries, C. (2004). Portfolio Selection with Heavy Tails. *Tinbergen Institute Discussion Papers* 05-009/2.

[29] Hyvärinen, A., Karhunen, J., Oja, E. (2001). *Independent Component Analysis.* Wiley-Interscience.

[30] Jansen, D., Koedijk, K., de Vries, C. (2000). Portfolio selection with limited downside risk. *Journal of Empirical Finance,* **7**, 247-269.

[31] Kahneman, D., Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, **47** , No 2, 263-291.

[32] Kiviluoto, K., Oja, E., (1998). Independent component analysis for parallel financial time series. In *Proc. Proc. Int. Conf. on Neural Information Processing (ICONIP'98),* Vol. 2, 895-989, Tokyo, Japan.

[33] Koedijk, K.G., Kool, C.J. (1992). Tail estimates of East European Exchange Rates. *J. Busin. Econ. Stat.,* **10**, 83-96.

[34] Longin, F. (1996). The Asymptotic Distribution of Extreme Stock Market Returns. *J. Busin.,* **63**, 383-408.

[35] Loretan, M., Phillips, C.B. (1994). Testing the covariance stationarity of heavy-tailed time-series: an overview of the theory with applications to several financial instruments. *J. Emp. Finance,* **1**, 211-248.

[36] Malaroiu, S., Kiviluoto, K., Oja, E. (2000). ICA preprocessing for time series prediction. In *Proc. Second Int. Workshop on Independent Component Analysis and Blind Source Separation*, June 19-22, Helsinki, 453-457.

[37] Malevergne, Y., Sornette D. (2001). General framework for a portfolio theory with non-Gaussian risks and non-linear correlations. In *Proceedings of the 18th INTERNATIONAL CONFERENCE IN FINANCE*, 26, 27 & 28 JUNE 2001 NAMUR - Belgium.

[38] Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, **7**, 77-91.

[39] Mason, D.M. (1982). Laws of large numbers for sums of extreme values. *Ann. Prob.,* **10**, 756-764.

[40] Moody, J. E., Wu, L. (1997). What is the "true price"? - State space models for high frequency FX data. In *Decision Technologies for Financial Engineering (Proceedings of the Fourth International Conference on Neural Networks in the Capital Markets, NNCM-96)*, Eds Weigend, A.S., Abu-Mostafa, Y.S. and Refenes, A., 346-358, Singapore, World Scientific.

[41] Moody, J. E., Wu, L. (1997). Multi-effect decompositions for financial data modelling. In *Advances in Neural Information Processing Systems 9 (NIPS-96),* eds Mozer, M.C., Jordan, I. and Petsche, T. 995-1001. MIT Press, Cambridge, MA.

[42] Pisarenko, V., Sornette, D. (2004). New statistic for financial return distributions: power-law or exponential? *Physica A.*

[43] Rabin, M. (2000). Risk Aversion and Expected Utility Theory: A Calibration Theorem. *Econometrica*, **68**, 1281-1292.

[44] Rabin, M., Thaler, R. H. (2001). Anomalies: Risk Aversion. *Journal of Economic Perspectives,* **15**, 219-232.

[45] Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer-Verlag, New York.

[46] Roy, A.D. (1952). Safety first and the holding of assets. *Econometrica,* **20**, 431-449.

[47] Slim, S. (2005). PhD Thesis, Université Paris X Nanterre.

[48] Straetmans, S. (1998). *Extreme Financial Returns and their Comovements.* PhD Thesis, No. 181, Erasmus University, Rotterdam.

[49] Susmel, R. (2001). Extreme observations and diversification in Latin American emerging equity markets. *Journal of International Money Finance,* ***20***, 971-986.

[50] Szegö, G. (2004). *Risk Measures for the 21-th Century.* Edited by G. Szegö, Wiley.

[51] Vessereau, T. (2000). Factor Analysis and Independent Component Analysis in Presence of High Idiosyncratic Risks. *Scientific series (CIRANO),* ***46****.*

[52] Zajdenweber, D. (1996). Extreme Values in Business Interruption Insurance. *J. Risk Insurance,* **63**, 95-110.

[53] Zajdenweber, D. (2000). *Economie des Extrêmes.* Flammarion, Collection Nouvelle Bibliothèque Scientifique, Paris (in French).