

# Portfolio Value at Risk Based On Independent Components Analysis

Ying Chen<sup>1,2</sup>, Wolfgang Härdle<sup>1</sup> and Vladimir Spokoiny<sup>1,2</sup>

<sup>1</sup> CASE - Center for Applied Statistics and Economics

Humboldt-Universität zu Berlin

Wirtschaftswissenschaftliche Fakultät

Spandauerstrasse 1, 10178 Berlin, Germany

<sup>2</sup> Weierstraß - Institute für Angewandte Analysis und Stochastik

Mohrenstrasse 39, 10117 Berlin, Germany

## Abstract

Risk management technology applied to high dimensional portfolios needs simple and fast methods for calculation of Value-at-Risk (VaR). The multivariate normal framework provides a simple off-the-shelf methodology but lacks the heavy tailed distributional properties that are observed in data. A principle component based method (tied closely to the elliptical structure of the distribution) is therefore expected to be unsatisfactory. Here we propose and analyze a technology that is based on Independent Component Analysis (ICA). We study the proposed ICVaR methodology in an extensive simulation study and apply it to a high dimensional portfolio situation. Our analysis yields very accurate VaRs.

**Keywords:** independent component analysis, Value-at-Risk

**MSC code:** 62G05 62H12 62H10

**Acknowledgement:** This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

# 1 Introduction

The Value-at-Risk (VaR) calculation of large portfolios is a challenging task - both numerically and statistically. VaR indicates the possible loss over a given time horizon at a risk level  $a$ . From a statistical point of view, it is the  $a$ -quantile of the joint distribution of the portfolio's risk factors which are modeled heteroscedastically:

$$x_t = \Sigma_t^{1/2} \varepsilon_t, \quad (1)$$

where  $x_t \in \mathbb{R}^d$  is the risk factor vector, e.g. the (log) returns of  $d$  individual financial instruments. The matrix  $\Sigma_t$  denotes the corresponding time-dependent covariance and  $\varepsilon_t$  is the  $d$ -dimensional standardized residual vector. The portfolio VaR calculation becomes technically difficult for high dimensionality of the portfolio.

In order to solve this and other numerical problems, portfolio variations are typically mapped into a conditional multivariate normal framework such as RiskMetrics launched by J.P. Morgan. Recall that, Gaussian distributed residuals or devolatilized returns  $\varepsilon_t = \Sigma_t^{-1/2} x_t$  are independent after devolatilization; the joint density of the residuals is the product of  $d$  marginals. In this sense, the portfolio VaR calculation is simplified and only covariance based. Many well developed covariance estimation methodologies in a high dimensional space have been easily applied in practice, for example, the Constant Conditional Correlation (CCC) model proposed by ? and the subsequent Dynamic Conditional Correlation (DCC) model proposed by ?, ?. The simplicity of this kind of covariance based methodology nevertheless bears a risk of modeling bias since, among other things, the assumed conditional Gaussian marginals are unable to mimic the heavy tailedness of financial time series observed in markets. This issue has been addressed in a variety of papers. For example, ? have studied the conditional Gaussian distribution fits to VaR that deliver satisfactory estimates at a moderate (e.g. 95%) confidence level but underestimate VaRs at more extreme levels such as 99%.

The tail problems are evident from Figure 1, where we compare the marginal density estimations of devolatilized returns of foreign exchange rates, the German Mark to the US Dollar (DEM/USD), from 1979-12-01 to 1994-04-01. In order to mimic the empirical distributional behavior of the real data, we assume 3 different distributional types. The RiskMetrics (left panel) and  $t(15)$ -deGARCH (right panel) methods fit the devolatilized returns, based on a GARCH(1,1) volatility process:  $\hat{\sigma}_t^2 = 1.65 * 10^{-6} + 0.07x_{t-1}^2 + 0.89\hat{\sigma}_{t-1}^2$ , by the normal and Student- $t$  with degrees of freedom 15 distributions. The GHADA (middle panel) technique assumes that the devolatilized returns follow a time stationary Generalized Hyperbolic (GH) distributional mechanism based on locally constant volatilities, see ? and ? for details. The nonparametric kernel density estimations (solid lines) corresponding to the devolatilized return processes are considered as benchmarks. According to the graph-

ical comparison, the GHADA technique is superior to the other two techniques since the empirical GH density coincides to the benchmark, especially in the tails. The  $t$ -deGARCH technique shows although a better tail fit compared to the RiskMetrics, it still deviates from the benchmark. This small comparison provides evidence that the normally distributional assumption is unreliable and will lead to low accuracy of univariate and portfolio VaR calculations.

The weak performance of the normal assumption motivates us to search for a different approach solving the technical problems of portfolio VaR calculations. An “ideal” situation is, as mentioned before, that the residuals  $\varepsilon_t$  are independent. Since based on the independence, the estimation of the joint distribution can be converted to marginals’ estimations.

In the context of sound engineering, signal detection from unknown filters and sources is treated by a method called Independent Component Analysis (ICA). This engineering method is designed for detection of blind folded signals and retrieves out of a high dimensional time series stochastically independent source components. A tutorial on ICA can be found in ? and a variety of numerical techniques to uncover Independent Components (ICs) are discussed in ?. Promoted by the success in engineering, ICA has been applied in different areas such as brain imaging (?) and telecommunication study (?). An early implementation of ICA in financial time series is given in ?, drawing comparisons of ICs and principal components (PCs) applied to 28 Japanese stocks from 1986 to 1989. Few contributions however exist for the application of ICA in risk management.

The first aim of this paper is to bring together the lines of thought of the engineering signal processing literature and newer statistical insights on the high dimensional VaR calculations. We coin the name ICVaR owing to the ICA technology. Given a trading strategy  $b_t \in \mathbb{R}^d$ , the portfolio return  $R_t \in \mathbb{R}$  is:

$$R_t = b_t^\top x_t. \quad (2)$$

Note that the ICVaR technique does not rely on a direct joint density estimation of the high dimensional returns  $x_t \in \mathbb{R}^d$  such as  $x_t = \Sigma_t^{1/2} \varepsilon_t$ . Instead the ICVaR procedure in a first step applies a linear transformation to  $x_t$ , i.e. a nonsingular matrix  $W$  yields (approximately) ICs  $y_t \in \mathbb{R}^d$ :

$$x_t = W^{-1} y_t \quad (3)$$

The matrix  $W$  is different from the Mahalanobis transformation  $\text{cov}(x)^{-1/2}$  - that creates ICs in the Gaussian regime - only in the case of non-Gaussian marginals as we will see later. The second ICVaR step concerns the fit of each IC univariately:

$$y_t = \text{diag}(\sigma_{1t}, \dots, \sigma_{dt}) \varepsilon_t = D_t^{1/2} \varepsilon_t \quad (4)$$

or equivalently:  $y_{jt} = \sigma_{jt} \varepsilon_{jt}, \quad j = 1, \dots, d,$

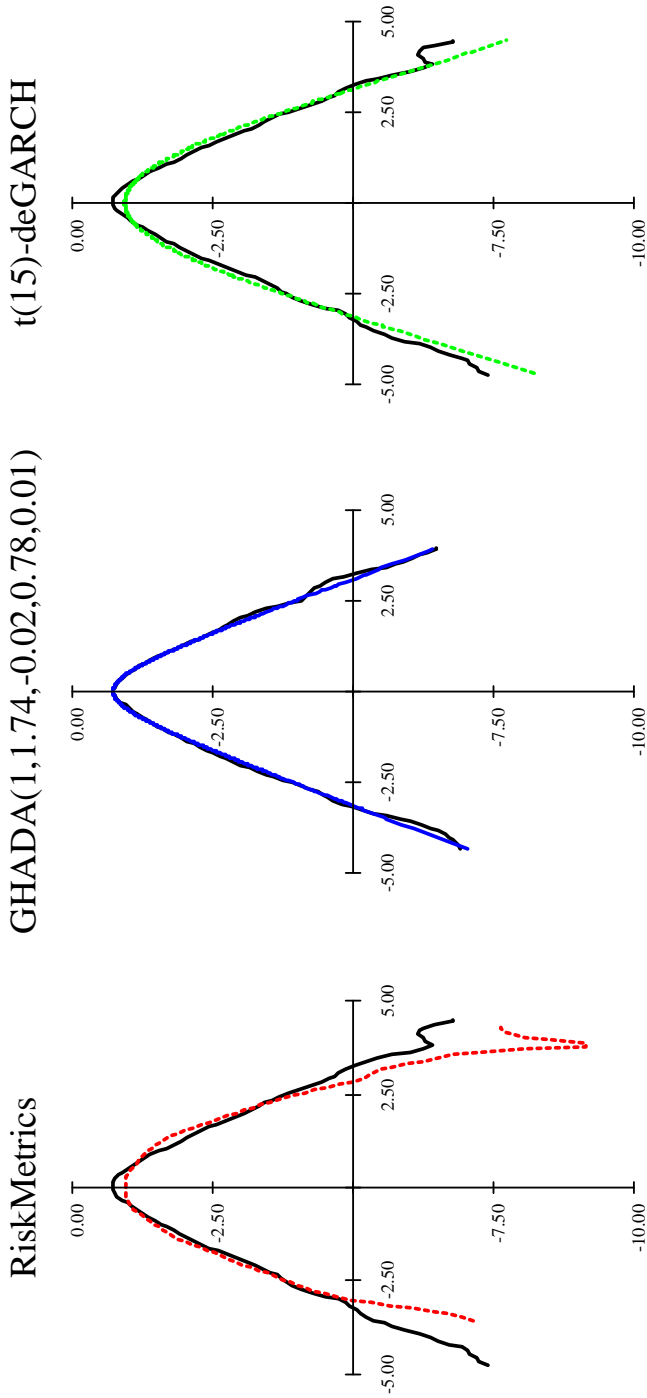


Fig. 1: Graphical comparison of density estimations based on the devolatilized DEM/USD returns from 1979-12-01 to 1994-04-01 (3719 observations). The nonparametric kernel density estimations are considered as benchmarks. The RiskMetrics based density estimation is the dotted curve on the left panel whereas the  $t(15)$ -deGARCH based estimation is displayed on the right. The corresponding GARCH(1,1) process is:  $\hat{\sigma}_t^2 = 1.65 * 10^{-6} + 0.07x_{t-1}^2 + 0.89\hat{\sigma}_{t-1}^2$ . The GHADA technique with the local constant volatility process and the GH(1, 1.74, -0.02, 0.78, 0.01) density estimation is displayed on the middle panel. Data source: FEDC([sfb649.wiwi.hu-berlin.de](http://sfb649.wiwi.hu-berlin.de)).

where the covariance  $D_t$  of the ICs is diagonal due to independence and  $\varepsilon_t$  are cross independent shocks. Based on the ICA, the high-dimensional VaR problem is now converted to simpler univariate VaR calculations. Given the discussion above on fitting VaR this opens a wide avenue of alternative VaR determination.

The second purpose of this paper is to compare the proposed ICVaR technique with the industry standard RiskMetrics and the most-often used  $t$ -deGARCH method. As discussed before, the RiskMetrics often gives underestimated VaRs, an inevitable cost of the Gaussian assumption. On the contrary, the Students- $t$  distribution can better mimic the heavy tailed distributional behavior of financial risk factor than the RiskMetrics, but it is hard to reflect the leptokurtic scenario exactly, as illustrated in Figure 1. In this paper, we show how high accuracy can be reached by the ICVaR compared to the RiskMetrics and  $t$ -deGARCH methods in real data analysis.

The methodological contribution of the study unfolds in Section 2 where the ICA method is discussed. The simulation study is presented in Section 3. Further we apply the ICVaR to exchange rate portfolios with different artificial trading strategies. The ICVaR predicts risk levels precisely and outperforms the RiskMetrics and  $t$ -deGARCH methods. Finally we conclude our study in Section 5. All algorithms and figures can be recalculated by following the net linked indicator below the figures.

## 2 ICVaR Methodology

### 2.1 Basic model

The proposed ICVaR methodology consists of 2 main steps: searching for Independent Components (ICs) based on a linear transformation and modeling the univariate volatility and stochastic distribution.

$$R_t = b_t^\top x_t \tag{5}$$

$$= b_t^\top W^{-1} y_t \tag{6}$$

$$= b_t^\top W^{-1} D_t^{1/2} \varepsilon_t. \tag{7}$$

The idea of the ICA is that risk factors  $x_t \in \mathbb{R}^d$  can be represented by a linear combination of  $d$ -dimensional ICs. The linear transformation matrix  $W$  is assumed to be nonsingular. Due to the independence property of ICs, the covariance  $D_t$  must be a diagonal matrix and the elements of the stochastic vector  $\varepsilon_t$  are cross independent. Furthermore, it fulfills that  $E[\varepsilon_t | F_{t-1}] = 0$  and  $Var[\varepsilon_t | F_{t-1}] = I_d$ .

From a statistical viewpoint, this projection technique is desirable since the  $d$ -dimensional portfolio is decomposed to univariate and independent risk factors through a simple linear

transformation. Recall that, the joint density ( $f$ ) and the covariance of any linear transformed ICs such as  $x_t = W^{-1}y_t$  are analytically computable:

$$\begin{aligned} f_y &= \prod_{j=1}^d f_{y_j}, & \Sigma_{t,y} &= D_t \\ f_x &= \text{abs}(|W|)f_y(Wx), & \Sigma_{t,x} &= W^{-1}D_tW^{-1\top}. \end{aligned}$$

In the second step of the proposed ICVaR, the diagonal elements of  $D_t$  and each component of  $\varepsilon_t$  are estimated univariately since the matrix manipulation is equivalent to:

$$y_{jt} = \sigma_{jt}\varepsilon_{jt}, \quad j = 1, \dots, d, \quad (8)$$

where  $\sigma_{jt}$  is the square root of the  $j$ -th diagonal element of  $D_t$  and  $\varepsilon_{jt}$  is the univariate stochastic term with  $E[\varepsilon_{jt}|F_{t-1}] = 0$  and  $\text{Var}[\varepsilon_{jt}|F_{t-1}] = 1$ . There are various univariate models that estimate the volatility and approximate the distribution of the stochastic term. Here we focus on the IC identification step. For the fit of the marginal IC factors we refer to the GHADA technique as in ?, where one estimates the local constant volatility adaptively in a homogeneous interval and fits the devolatilized returns in the GH distributional framework. An alternative approach is given by the  $t$ -deGARCH setup. It could be used to estimate the heteroscedastic volatility process and the Student- $t$  could be applied to pick up the heavy-tailedness of  $\varepsilon_j$ .

After these two steps, we compute the quantile of portfolio risk by Monte Carlo simulation or analytical methods e.g. saddle point approximation, see ?. For the ease of presentation, we concentrate in this paper on the simulation methodology. In particular, we generate  $d$ -dimensional samples of the fitted distributions with sample size  $M$ , from which we calculate the daily empirical  $a$ -quantile of the portfolio variations. The simulation will repeat  $N$  times and the average value of the empirical quantiles is considered as the portfolio VaR at level  $a$ :

$$\text{VaR}_{a,t} = \frac{1}{N} \sum_{n=1}^N \hat{F}_{a,t}^{-1}(R_t^{(n)}) = \frac{1}{N} \sum_{n=1}^N \hat{F}_{a,t}^{-1}\{b_t^\top \hat{W}^{-1} \hat{D}_t^{1/2} \hat{\varepsilon}_t^{(n)}\},$$

where  $\hat{F}_{a,t}^{-1}$  denotes the empirical quantile function of  $R_t$ .

## 2.2 ICA: Properties and Estimation

Since ICA is a relatively new technique in this context, we present a small pedagogical illustration of its usage.

**Example:** Generate 3 independent GH random variables  $\text{GH}(y; 1, 2, 0, 1, 0)$ ,  $\text{GH}(y; 1, 1.7, 0, 0.5, 0)$  and  $\text{GH}(y; 1, 1.5, 0, 1, 0)$  as sources. The first distributional param-

eter specifies the subclass of the random variables. With the value of 1, they are hyperbolic (HYP) distributed. The other four parameters control the location, scale, asymmetry and likeliness of extreme events, see ?. These source components have mean  $(-0.02, 0.05, -0.00)^\top$  and standard deviation (SD)  $(0.83, 0.92, 0.99)^\top$  respectively. The linear transformation matrix is the estimate based on 3 real German stocks' returns: ALLIANZ, BASF and BAYER from 1974-01-02 to 1996-12-30 (Data source: FEDC at [sfb649.wiwi.hu-berlin.de](http://sfb649.wiwi.hu-berlin.de)).

$$W^{-1} = \begin{pmatrix} 1.31 & 0.14 & 0.18 \\ -0.42 & -1.26 & -1.25 \\ -0.03 & 0.41 & -0.49 \end{pmatrix} 10^{-2} \quad (9)$$

The time series  $x_t = W^{-1}y_t$  are analyzed by the ICA. The covariance of  $x_t$  is

$$\text{cov}(x) = \begin{pmatrix} 13.66 & 6.04 & 6.49 \\ 6.04 & 15.54 & 11.64 \\ 6.49 & 11.64 & 16.07 \end{pmatrix} 10^{-5}, \quad (10)$$

which is comparable to  $W^{-1}D_tW^{-1\top}$ . Recall that in the Gaussian framework, the Mahalanobis transformation delivers independent variables:

$$\widehat{\text{cov}}(x)^{-1/2} = \begin{pmatrix} 0.91 & -0.09 & -0.12 \\ -0.09 & 1.03 & -0.41 \\ -0.12 & -0.41 & 1.04 \end{pmatrix} 10^2 \quad (11)$$

which is clearly distinct from

$$W = \begin{pmatrix} 0.79 & 0.10 & 0.03 \\ -0.11 & -0.44 & 1.08 \\ -0.15 & -0.38 & -1.10 \end{pmatrix} 10^2 \quad (12)$$

indicated by (9). Figure 2 provides an illustration of this procedure. The top row contains the 3 independent source signals  $y_t$ . The middle row displays the time series  $x_t = W^{-1}y_t$ . One sees the scale changes and different random patterns display. The last row of Figure 2 shows the ICs estimated by the ICA method. The time series on the top and the bottom look familiar but the sign and the ordering may change as displayed in the figure; The first estimated IC displays the similar movement as the third generated IC. Furthermore, the third estimated IC has a mirror pattern of the second true IC.

#### *Scale identification*

In fact, the scales of  $y_t$  and  $W$  are not identifiable. Given a matrix  $C = \text{diag}(c_1, \dots, c_d)$

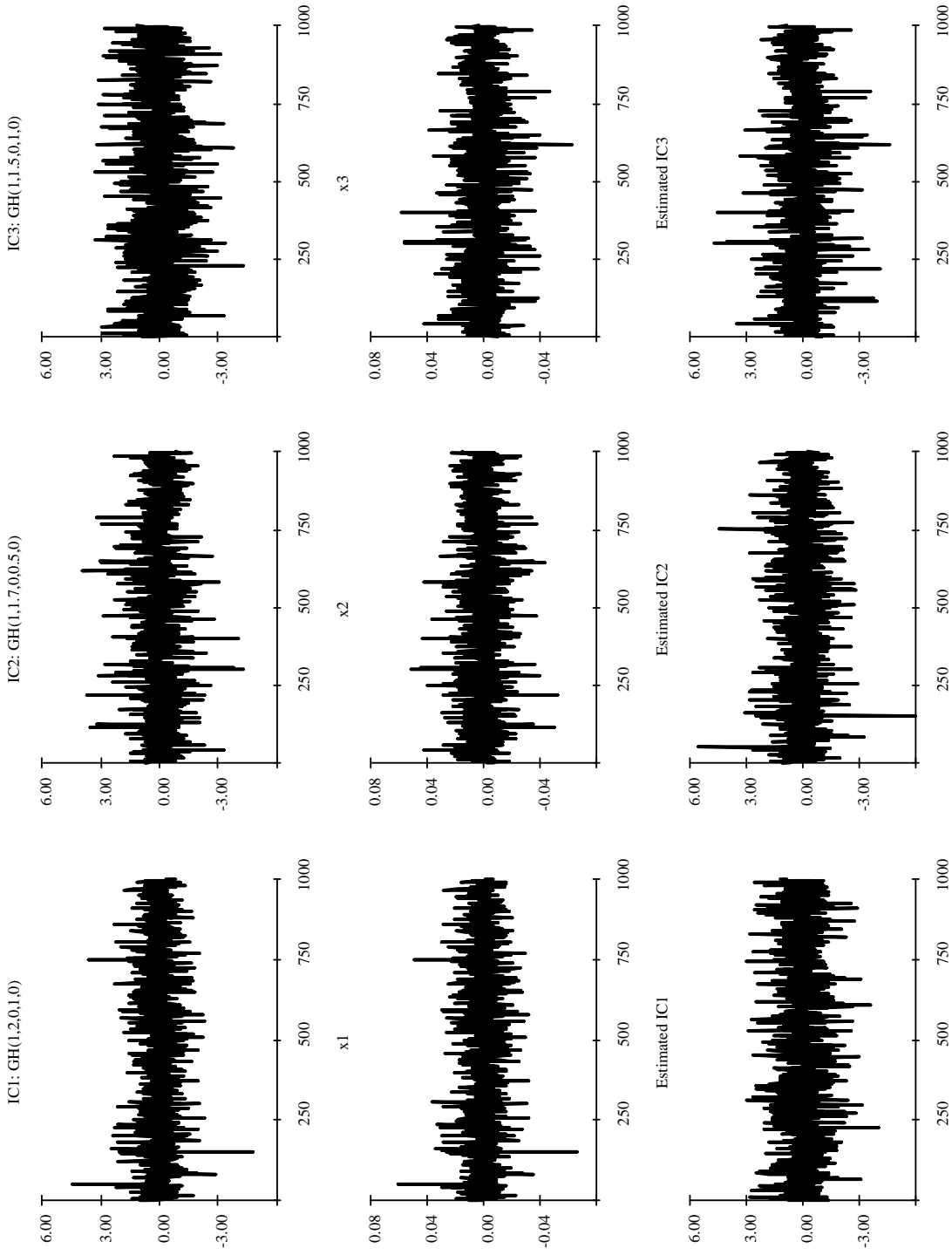


Fig. 2: Simulation: on the top are time plots of 3 independent GH variables. The mixed variables  $x_t = W^{-1}y_t$  are displayed on the middle row where the linear transformation matrix  $W$  is estimated based on 3 German stocks' returns: ALLIANZ, BASF and BAYER from 1974-01-02 to 1996-12-30. The time series of the estimated ICs are displayed on the bottom.



with  $c_j \neq 0$ , the new ICs ( $Cy_t$ ) with the transformation ( $W^{-1}C^{-1}$ ) also fulfill (3). In order to avoid the identification problem, it is suggested to prewhiten  $x_t$  and assume  $y_t$  to be standardized. The Mahalanobis transformation  $\widehat{\text{cov}}(x)^{-1/2}$  does the prewhitening job. It is not hard to see that  $W$  becomes then an orthogonal matrix. Denote by  $\tilde{x}_t$  - the prewhitened  $x_t$  and  $y_t = \tilde{W}\tilde{x}_t$  the corresponding ICs.  $W = \tilde{W}\widehat{\text{cov}}(x)^{-1/2}$  is then the linear transformation for the original observations. For notational convenience, we assume from now on that  $x_t$  has been prewhitened.

#### *Order identification*

Furthermore, the order of the ICs is ambiguous. Given a permutation matrix  $P$ , the ICs ( $Py_t$ ) fulfill (3) with a new transformation ( $W^{-1}P^{-1}$ ) as well.

#### *IC is necessarily non-Gaussian*

Given a  $d$ -dimensional standardized Gaussian vector  $x_t$  and an orthogonal matrix  $W$ , the joint pdfs of  $x_t$  and  $y_t$  are in fact identical and unrelated with  $W$ :

$$\begin{aligned} f_x &= |2\pi\mathbf{I}_d|^{-\frac{1}{2}} \exp\left(-\frac{x^\top x}{2}\right) \\ &= f_x(W^{-1}y)|\det W^{-1}| \\ &= |2\pi\mathbf{I}_d|^{-\frac{1}{2}} \exp\left(-\frac{y^\top y}{2}\right) = f_y. \end{aligned}$$

The third condition is although strict, it is naturally fulfilled in financial applications, where financial time series, even after devolatilization, display heavy tailed behavior.

How can we estimate the linear transformation matrix  $W$  after we have prewhitened the non-Gaussian variables? Independence of the components of a random vector  $y \in \mathbb{R}^d$  can be measured by the mutual information:

$$\begin{aligned} I(W, y) &= I(W, f_y) = \sum_{j=1}^d H(y_j) - H(y) \\ &= \sum_{j=1}^d H(y_j) - H(x) - \log|\det(W)| \end{aligned} \quad (13)$$

where  $H(y) = H(f_y) = -\int f_y(u) \log f_y(u) du$  is the entropy of the vector  $y$  with a joint pdf  $f_y$ . If the components of  $y$  are independent, then the mutual information will reach its minimum with a value of 0. Therefore, the IC searching is identical to minimizing (13) w.r.t.  $W$ . Since  $H(x)$  is fixed given the data, and the matrix  $W$  is orthogonal after prewhitening, this problem is further equivalent to minimizing the term  $\sum_{j=1}^d H(y_j)$ . Now we replace the objective function of the optimization problem by  $\min H(y_j)$  w.r.t.  $w_j$ , the  $j$ -th row of  $W$ .

Note that:

$$\sum_{j=1}^d \min H(y_j) \leq \min \sum_{j=1}^d H(y_j) = \min \sum_{j=1}^d H(w_j x) \quad (14)$$

This replacement leads to some loss in the  $W$  estimation but extensively speeds up the estimation procedure. Moreover, the entropy and negentropy  $J(w_j, y_j) = H\{N(0, 1)\} - H(y_j)$  are in one-to-one correspondence, we can also formulate the optimization problem as:

$$\hat{w}_j = \operatorname{argmin} H(y_j) = \operatorname{argmax} J(w_j, y_j).$$

The negentropy is always nonnegative, since with a fixed variance, the Gaussian random variable has the largest entropy among all distributional types, ?. Therefore, the negentropy is considered as a non-Gaussian measure and widely used in Projection Pursuit (PP), see ?. Furthermore, the PP methods of searching non-Gaussian direction can be applied in the IC identification as well. Nevertheless, compared to the cumulant based PP method, the entropy or negentropy is less sensitive to outliers and therefore preferable.

A question however raises at this stage, i.e. the marginal pdfs of the ICs in the entropy or the negentropy are unknown. A distributional free approximation of the univariate negentropy has been proposed by ?:

$$J(w_j, y_j) \approx C \{E\{G(y_j)\} - E\{G\{N(0, 1)\}\}\}^2$$

where  $C$  is a constant and  $G$  is an even function, e.g.  $G(y_j) = \frac{1}{\kappa} \log \cosh(\kappa y_j)$ ,  $1 \leq \kappa \leq 2$ .

The estimation of the linear transformation matrix in the ICA is estimated by:

$$\hat{w}_j = \operatorname{argmax} \{E\{G(w_j X)\} - E\{G\{N(0, 1)\}\}\}^2 \quad (15)$$

$$s.t. \quad W^\top W = I_d \quad (16)$$

Based on the Kuhn-Tucker condition and the Newton's method, ? have proposed the FastICA algorithm. ? has shown that the estimates based on the FastICA algorithm are consistent.

**FastICA algorithm: set  $j = 1$**

1. Choose an initial vector  $w_j$  of unit norm,  $W = (w_1, \dots, w_d)^\top$ .
2. Let  $w_j^{(n)} = E\{g(w_j^{(n-1)} x)x\} - E\{g'(w_j^{(n-1)} x)\}w_j^{(n-1)}$ , where  $g$  denotes the first derivative of  $G(y_j)$  and  $g'$  the second derivative. In practice, the sample mean is applied for  $E[\cdot]$ .
3. Orthogonalization:  $w_j^{(n)} = w_j^{(n)} - \sum_{k \neq j} (w_j^{(n)\top} w_k) w_k$ .

4. Normalization:  $w_j^{(n)} = w_j^{(n)} / \|w_j^{(n)}\|$ ,  $\|\cdot\|$  denotes the norm.
5. If not converged, i.e.  $\|w_j^{(n)} - w_j^{(n-1)}\| \neq 0$ , go back to 2.
6. Set  $j = j + 1$ . For  $j \leq d$ , go back to step 1.

### 3 Simulation Study

The reliability of the proposed ICVaR depends on the linear transformation matrix estimation and the univariate modeling on the ICs estimated. In order to fit the local distributions of ICs, we apply the GHADA technique due to its good performance in the simulation and empirical studies in ?. The target of the simulation study here is to search for ICs and compare the marginal densities of the estimated and generated ICs.

In particular, we pursue an experiment to check the validation of the FastICA approach with Normal Inverse Gaussian (NIG) distributed ICs, where the NIG distribution is a subclass of the GH distribution with the fixed parameter  $\lambda = -0.5$ , see ?. We generate  $d = 50$  NIG samples with  $T = 1000$  observations, i.e.  $y_j \sim \text{NIG}(\alpha_j, \beta_j, \delta_j, \mu_j)$  for  $j = 1, \dots, 50$ . Without loss of generality, we set  $\mu_j = 0$  and  $\delta_j = 1$ . The parameter  $\alpha_j$  is uniformly distributed in  $[1, 2]$  and  $\beta$  fulfills the condition:

$$\text{Var}[y_j] = \frac{1}{\sqrt{\alpha_j^2 - \beta_j^2}} \frac{\alpha_j^2}{\alpha_j^2 - \beta_j^2} = 1,$$

such that the generated ICs have unit variances. Furthermore, the sign of  $\beta$  is chosen arbitrarily. Table 1 shows the distributional parameters of the generated ICs. The linear transformation matrix  $W^{-1}$  is obtained via the Jordan decomposition of a square matrix, whose elements are standard normally distributed. The mixed time series  $x_t = W^{-1}y_t$  are analyzed by the ICA.

We apply the FastICA algorithm to the transformed time series  $x_t$  and estimate the NIG parameters of each estimated IC. We order the 50 estimated independent series to minimize the mean absolute error (MAE) of the marginal pdfs between the estimated and generated ICs. Overlapping is avoided in the ordering:

$$\text{MAE}_j = \frac{1}{T} \sum_{t=1}^{T=1000} |f(\hat{y}_{jt}) - f(y_{jt})|.$$

The largest two MAEs are 0.09 (ICest 20) and 0.04 (ICest 17), indicating the worst cases of IC searching. In this case, it is expected to get accurate VaR estimations based on these fits.

ICest	$\hat{\alpha}$	$\hat{\beta}$	IC	$\alpha$	$\beta$	MAE	ICest	$\hat{\alpha}$	$\hat{\beta}$	IC	$\alpha$	$\beta$	MAE
1	1.89	0.63	23	1.39	0.61	0.63	26	1.84	0.60	2	1.92	1.14	2.56
2	1.76	0.60	25	1.97	1.19	1.57	27	1.82	-0.18	16	1.57	-0.80	2.56
3	1.45	-0.51	30	1.06	-0.21	0.32	28	1.30	-0.37	8	1.17	-0.37	0.32
4	1.64	0.42	36	1.22	0.43	0.52	29	1.34	0.13	17	1.06	0.21	0.68
5	2.11	-0.67	13	1.31	-0.53	0.70	30	2.02	-0.39	44	1.54	0.77	2.73
6	2.05	-0.62	28	1.39	-0.61	0.93	31	3.05	0.80	35	1.78	1.00	3.27
7	1.73	-0.41	50	1.60	-0.83	0.76	32	1.79	-0.45	32	1.34	-0.57	0.77
8	1.62	0.45	7	1.15	0.35	0.45	33	1.58	0.08	4	1.24	0.46	0.88
9	1.27	0.16	6	1.96	1.17	1.58	34	1.59	0.35	47	1.45	0.67	0.74
10	1.51	0.43	27	1.51	-0.74	2.33	35	2.00	-0.33	34	1.43	-0.66	1.18
11	2.44	1.33	38	1.27	0.49	0.98	36	1.63	-0.06	31	1.85	1.07	1.37
12	1.65	-0.25	26	1.35	-0.57	0.79	37	1.81	0.29	29	1.93	1.15	1.29
13	2.58	-1.35	10	1.89	-1.11	1.69	38	1.49	-0.01	22	1.66	-0.89	2.97
14	1.60	0.15	24	1.54	0.77	0.98	39	2.39	-0.30	42	1.53	-0.76	1.27
15	1.52	-0.03	20	1.04	0.18	0.67	40	1.52	-0.22	21	1.39	0.62	1.48
16	2.14	-1.00	1	1.72	-0.94	3.09	41	2.32	-0.001	41	1.70	0.93	3.37
17	2.20	0.64	3	1.75	0.97	4.07	42	2.87	0.30	9	1.50	0.73	1.37
18	1.44	0.48	15	1.40	0.63	2.28	43	2.27	0.24	14	1.28	0.50	1.13
19	1.44	-0.40	33	1.70	0.93	3.22	44	2.44	0.46	43	1.87	1.09	1.39
20	1.90	-0.54	39	1.72	0.95	9.73	45	1.88	-0.07	18	1.44	-0.67	1.23
21	1.57	0.39	11	1.63	0.86	0.72	46	2.22	0.13	37	1.39	0.62	1.24
22	1.80	-0.56	12	1.78	-1.00	0.84	47	1.79	0.26	49	1.45	-0.68	1.41
23	1.69	0.20	46	1.69	0.92	1.09	48	3.03	0.75	40	1.95	1.17	2.37
24	1.76	-0.27	5	1.39	0.62	1.53	49	3.31	-0.18	48	1.50	-0.73	1.49
25	1.50	0.26	19	1.69	-0.92	1.38	50	3.77	0.29	45	1.53	0.76	1.52

Tab. 1: ML estimators  $\hat{\alpha}_j$  and  $\hat{\beta}_j$  of the estimated ICs, the parameters of the true ICs and the MAE (unit:  $10^{-2}$ ).

ICVaRsimplpdfcompICVaRsimgchoice

Time Series	Mean	SD	Skewness	Kurtosis	$\rho_1$	$\rho_2$
DEM/USD	0.00	$0.71 \times 10^{-2}$	-0.13	4.94	0.02	0.01
GBP/USD	0.00	$0.69 \times 10^{-2}$	-0.01	5.64	0.08	0.01
IC1	-0.02	1.00	-0.62	8.71	0.07	0.02
IC2	0.01	1.00	-0.08	5.19	0.05	0.01

Tab. 2: Descriptive statistics of the log returns and the two estimated independent processes of the DEM/USD and GBP/USD rates.

ICVaRfxdescriptive

## 4 Empirical Study

In this section, we analyze foreign exchange rate portfolios with artificial and time-invariant trading strategies. Since the position of these individual risk factors are constant in time, one can simply use the historical simulation approach where the portfolio returns are considered as the single risk factor. It speeds up VaR computations without losing the portfolio's dependence information. However in practice, it may slow down the calculation as the trading strategies changes, see ?. Three VaR models are evaluated: the proposed ICVaR approach based on the multivariate risk factors, the RiskMetrics and  $t$ -deGARCH methodologies based on the univariate portfolio returns. In the RiskMetrics and  $t$ -deGARCH frameworks, we apply the GARCH(1,1) setup to estimate the dependence structure of real data and assume that the devolatilized returns are Gaussian or Student- $t$  distributed. The degrees of freedom (df) of the Student- $t$  distribution are selected by the maximum likelihood method.

The foreign exchange market, or "Forex" market, is by far the largest financial market in the world with trading volumes surpassing USD 1.5 trillion on some days. The very active buying and selling of traders make it further the most liquid financial market. Here we consider portfolios including two exchange rates: DEM/USD and the British Pound to the US Dollar (GBP/USD) from 1979-12-01 to 1994-04-01. We forecast VaRs one-day-ahead w.r.t. four artificial trading strategies, i.e.  $b_1 = (1, 1)^\top$ ,  $b_2 = (1, 2)^\top$ ,  $b_3 = (-1, 2)^\top$  and  $b_4 = (-2, 1)^\top$ . For example,  $b_1$  means holding one unit DEM/USD and one unit GBP/USD forward contracts.

The data is available at FEDC ([sfb649.wiwi.hu-berlin.de](http://sfb649.wiwi.hu-berlin.de)). Each time series consists of 3721 observations. Table 2 summarizes the statistical properties of the original data and the estimated ICs based on the linear transformation. These four time series are all centered around 0, approximately symmetric but distinctly non-Gaussian indicated by their large kurtoses. The temporal correlations are very small; The ACF plots show that the serial correlations decay at the very beginning lags, indicating a weak stationarity of the four series, see Figure 3. On the other hand, the cross correlation of the two exchange returns

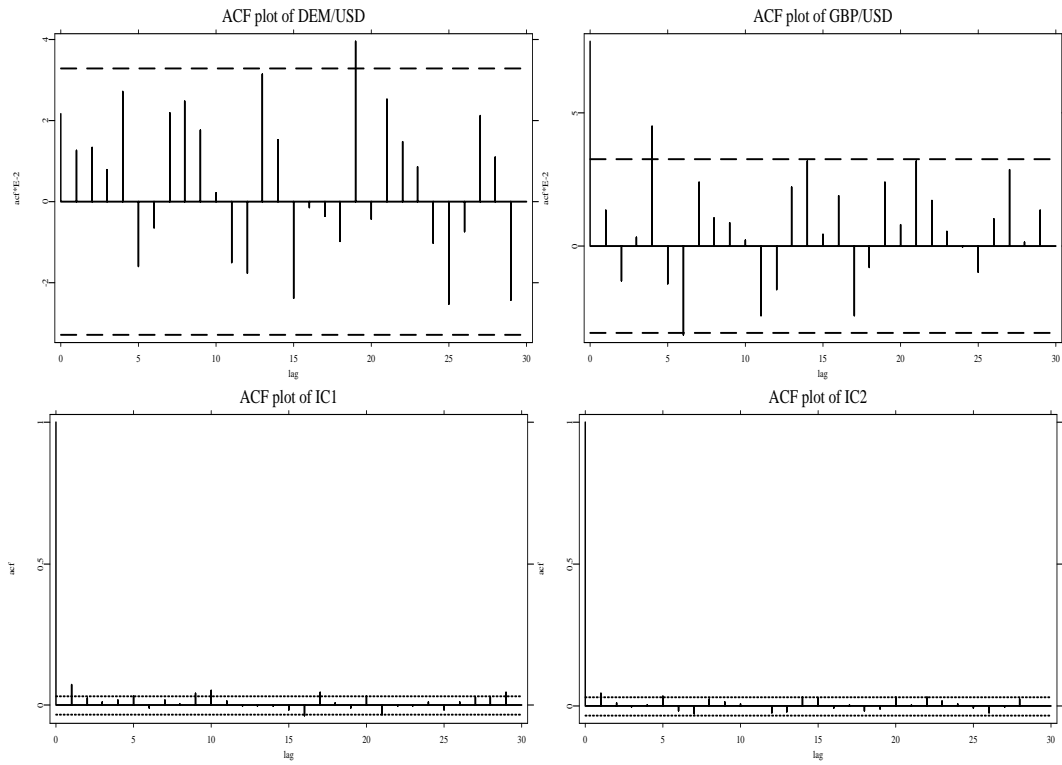


Fig. 3: ACF plots of the log returns of the DEM/USD (left) and the GBP/USD (right) are displays on the top. Below are the ACF plots of the estimated IC series: IC1 (left) and IC2 (right).

ICVaRfxdescriptive

is over 0.77, referring a strongly linear dependence between them.

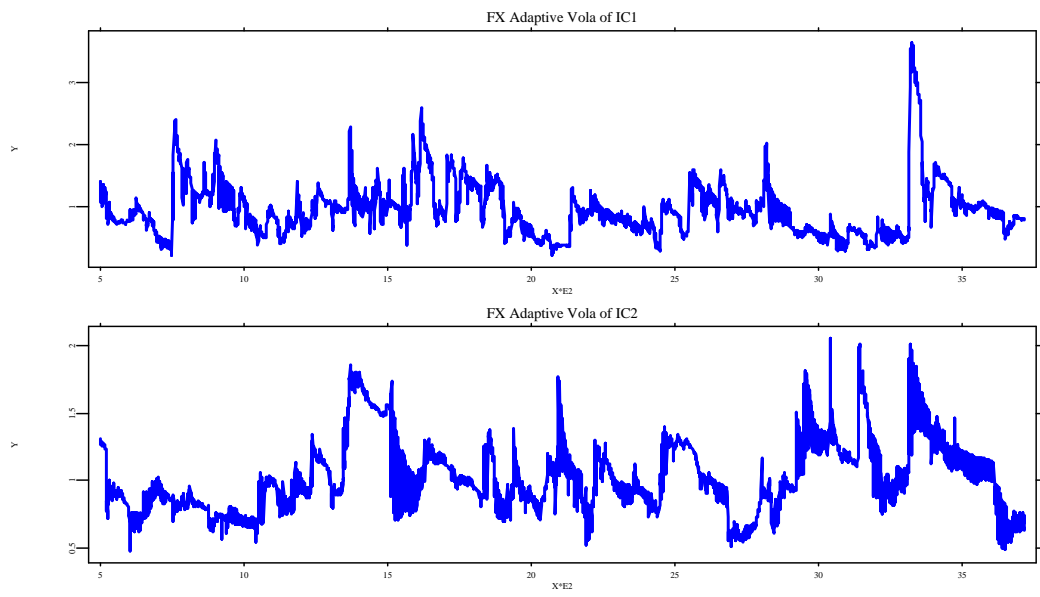


Fig. 4: Adaptive volatility processes of the FX ICs.

ICVaRfxdescriptive

GH Type	Time Series	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$	$\hat{\mu}$
HYP	IC1	1.71	-0.17	0.55	0.12
HYP	IC2	1.77	0.02	0.71	-0.01
NIG	IC1	1.22	-0.18	1.10	0.13
NIG	IC2	1.37	0.03	1.28	-0.02

Tab. 3: GH parameters of the ICs.

ICVaRfxdescriptive

Applying the FastICA algorithm, we estimate the linear transformation matrix and its inverse:

$$\hat{W} = \begin{pmatrix} 207.93 & -213.63 \\ 77.72 & 73.29 \end{pmatrix}, \quad \hat{W}^{-1} = \begin{pmatrix} 2.30 & 6.71 \\ -2.44 & 6.53 \end{pmatrix} 10^{-3} \quad (17)$$

We then implement the GHADA approach to fit the distributional feature of each estimated IC. Figure 4 shows the adaptive volatility series based on the two ICs. The well-known volatility clustering is reflected and volatility jumps appear. These jumps happen in most cases at different time and have individual influences on the return processes.

The estimated HYP and NIG parameters of the devolatilized ICs are reported in Table 3. In the Monte Carlo simulation to find the empirical quantiles of the portfolios, we generate  $d = 2$  samples with  $M = 10,000$  observations. The daily empirical quantiles at 3 risk levels  $a = 5\%$ ,  $1\%$  and  $0.5\%$  are the average values of  $N = 100$  repetitions. We redo the simulation for the last  $T = 1000$  days. The daily means and standard deviations of the 3 empirical quantiles given different trading strategies and two GH subclasses (HYP and NIG) are reported in Table 4 to Table 7. The largest standard deviation of daily empirical quantiles is underlined in each category. The values are small due to the large sample size, indicating efficient estimation of the daily quantiles.

The backtesting results based on the ICVaR, the RiskMetrics and  $t$ -deGARCH methodologies are reported in Table 8. The dfs of the Student- $t$  fits are 16, 16, 13 and 14 w.r.t. different trading strategies. Based on two likelihood ratio tests: LR1 for risk level and LR2 for exceedances clustering (?), the proposed ICVaR model is superior to the other two candidates. In the risk level test (LR1), the NIG fit performs even better than the HYP fit. In the RiskMetrics framework, the exceedances happen minimal 2.6 times and maximal 23 times more than the expected risk level, e.g. for the trading strategy  $b = (1, 1)^\top$ . In some cases, the underestimation is even over 25 times. Compared to the RiskMetrics, the  $t$ -deGARCH method improves the VaR forecasting as the extreme risk levels such as  $0.5\%$  are considered. However both models are rejected everywhere in the two tests at 99% level, meaning unreliable predictions. Exemplary graphical illustrations of the VaR time plot are displayed in Figure 5 to Figure 6.



HYP	a 5%		a 1%		a 0.5%		NIG	a 5%		a 1%		a 0.5%	
	mean	SD%	mean	SD%	mean	SD%		Day	mean	SD%	mean	SD%	mean
Day													
1	-0.01	0.01	-0.02	0.04	-0.02	0.06	1	-0.01	0.02	-0.02	0.05	-0.02	0.06
2	-0.01	0.01	-0.01	0.04	-0.02	0.05	2	-0.01	0.01	-0.02	0.04	-0.02	0.06
.	.	.	.	.	.	.	.	.	.	.	.	.	.
424	-0.03	<u>0.07</u>	-0.06	0.15	-0.07	0.21	423	-0.04	0.06	-0.07	0.15	-0.07	0.23
425	-0.03	0.06	-0.06	0.14	-0.07	0.22	425	-0.04	0.06	-0.07	0.14	-0.07	<u>0.24</u>
426	-0.03	0.04	-0.05	0.11	-0.06	0.18	426	-0.03	0.05	-0.06	0.12	-0.06	0.18
597	-0.03	0.06	-0.06	<u>0.16</u>	-0.06	<u>0.22</u>	607	-0.04	0.06	-0.06	<u>0.15</u>	-0.07	0.21
.	.	.	.	.	.	.	.	.	.	.	.	.	.
999	-0.01	0.02	-0.02	0.04	-0.02	0.06	622	-0.03	<u>0.07</u>	-0.06	0.12	-0.07	0.18
1000	-0.01	0.02	-0.02	0.04	-0.02	0.07	1000	-0.01	0.02	-0.02	0.05	-0.02	0.07

Tab. 4: Descriptive statistics of daily empirical quantile estimates:  $b = (1, 1)^\top$ , MC simulation with  $M = 10000$ ,  $N = 100$ ,  $T = 1000$ .  
ICVaRmcqstat

HYP	a 5%		a 1%		a 0.5%		NIG	a 5%		a 1%		a 0.5%	
	mean	SD%	mean	SD%	Day	mean		SD%	mean	SD%	mean	SD%	mean
Day													
1	-0.01	0.02	-0.03	0.06	-0.03	0.09	1	-0.01	0.02	-0.03	0.06	-0.03	0.09
2	-0.01	0.02	-0.02	0.05	-0.03	0.08	2	-0.01	0.02	-0.03	0.06	-0.03	0.09
.	.	.	.	.	.	.	.	.	.	.	.	.	.
424	-0.05	<u>0.10</u>	-0.09	0.23	-0.10	0.31	325	-0.06	<u>0.11</u>	-0.10	0.22	-0.11	0.30
425	-0.05	0.10	-0.10	0.21	-0.10	<u>0.33</u>	423	-0.06	0.11	-0.10	<u>0.25</u>	-0.11	0.32
597	-0.05	0.09	-0.09	<u>0.24</u>	-0.10	0.33	424	-0.06	0.10	-0.10	0.22	-0.11	0.32
.	.	.	.	.	.	.	.	.	.	.	.	.	.
999	-0.01	0.03	-0.03	0.07	-0.03	0.08	597	-0.05	0.09	-0.10	0.22	-0.10	<u>0.34</u>
1000	-0.02	0.03	-0.03	0.06	-0.03	0.10	1000	-0.02	0.03	-0.03	0.08	-0.04	0.11

Tab. 5: Descriptive statistics of daily empirical quantile estimates:  $b = (1, 2)^\top$ , MC simulation with  $M = 10000$ ,  $N = 100$ ,  $T = 1000$ .  
ICVaRmcqstat

HYP	a 5%		a 1%		a 0.5%		NIG	a 5%		a 1%		a 0.5%	
	mean	SD%	mean	SD%	mean	SD%		Day	mean	SD%	mean	SD%	mean
Day													
1	-0.01	0.01	-0.01	0.03	-0.01	0.04	1	-0.01	0.01	-0.01	0.03	-0.02	0.05
2	-0.01	0.01	-0.01	0.03	-0.01	0.05	2	-0.01	0.01	-0.01	0.03	-0.02	0.05
.	.	.	.	.	.	.	.	.	.	.	.	.	.
601	-0.03	0.05	-0.06	0.15	-0.06	<u>0.22</u>	601	-0.03	0.05	-0.06	0.15	-0.07	0.19
607	-0.04	0.06	-0.06	0.13	-0.07	0.18	606	-0.03	0.06	-0.06	<u>0.15</u>	-0.07	<u>0.23</u>
611	-0.03	<u>0.06</u>	-0.06	0.13	-0.06	0.20	607	-0.04	<u>0.07</u>	-0.07	0.13	-0.07	0.19
613	-0.03	0.05	-0.06	<u>0.15</u>	-0.07	0.20	613	-0.04	0.06	-0.06	0.12	-0.07	0.18
.	.	.	.	.	.	.	.	.	.	.	.	.	.
999	-0.01	0.01	-0.01	0.03	-0.01	0.04	999	-0.01	0.01	-0.01	0.03	-0.01	0.04
1000	-0.01	0.01	-0.01	0.03	-0.01	0.04	1000	-0.01	0.01	-0.01	0.03	-0.01	0.05

Tab. 6: Descriptive statistics of daily empirical quantile estimates:  $b = (-1, 2)^\top$ , MC simulation with  $M = 10000$ ,  $N = 100$ ,  $T = 1000$ .  
ICVaRmcqstat

HYP	a 5%		a 1%		a 0.5%		NIG	a 5%		a 1%		a 0.5%	
	mean	SD%	mean	SD%	mean	SD%		Day	mean	SD%	mean	SD%	mean
Day													
1	-0.01	0.01	-0.01	0.03	-0.01	0.04	1	-0.01	0.01	-0.01	0.03	-0.02	0.04
2	-0.01	0.01	-0.01	0.03	-0.01	0.05	2	-0.01	0.01	-0.01	0.03	-0.02	0.05
.	.	.	.	.	.	.	.	.	.	.	.	.	.
607	-0.04	<u>0.06</u>	-0.06	0.13	-0.07	0.18	601	-0.03	0.06	-0.06	<u>0.15</u>	-0.07	<u>0.22</u>
608	-0.03	0.05	-0.06	0.13	-0.06	0.19	605	-0.04	0.06	-0.06	0.14	-0.07	0.19
612	-0.03	0.05	-0.06	<u>0.14</u>	-0.07	<u>0.21</u>	612	-0.04	<u>0.07</u>	-0.07	0.14	-0.07	0.19
.	.	.	.	.	.	.	.	.	.	.	.	.	.
999	-0.01	0.01	-0.01	0.03	-0.01	0.04	999	-0.01	0.01	-0.01	0.02	-0.01	0.04
1000	-0.01	0.01	-0.01	0.03	-0.01	0.04	1000	-0.01	0.01	-0.01	0.03	-0.02	0.04

Tab. 7: Descriptive statistics of daily empirical quantile estimates:  $b = (-2, 1)^\top$ , MC simulation with  $M = 10000$ ,  $N = 100$ ,  $T = 1000$ .  
ICVaRmcqstat

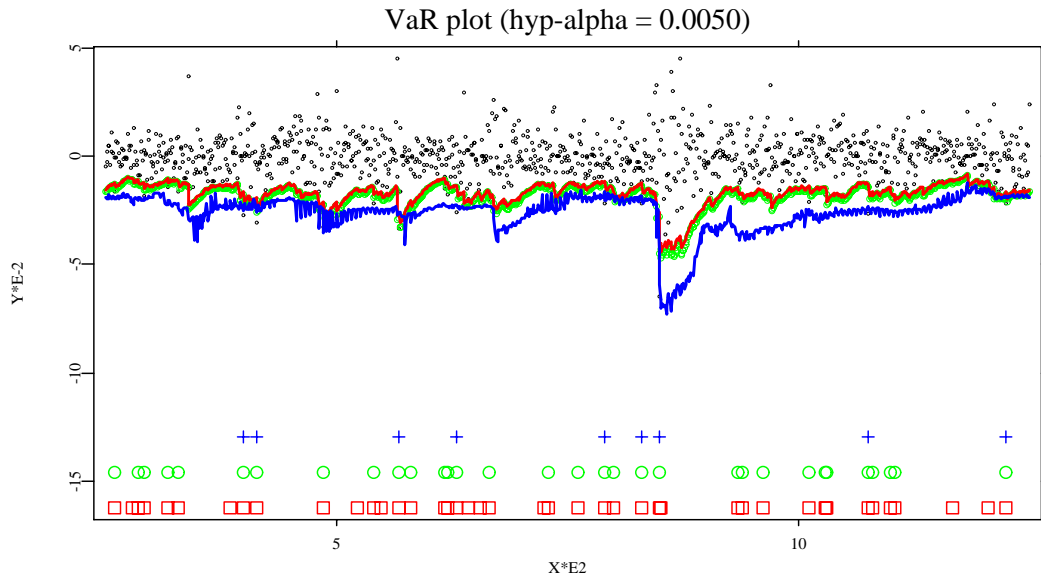


Fig. 5: VaR time plots of the exchange rate portfolio with trading strategy  $b_4 = (-2, 1)^\top$  and risk level  $a = 0.5\%$ . The one-day-ahead predicted VaRs are based on the HYP distribution. The dots are the real portfolio returns, the RiskMetrics (red) and the  $t$ -deGARCH (green) forecasts are displayed as solid line and circles while the ICVaR is displayed as a straight line. The exceedances w.r.t. the risk management models are displayed as cross on the bottom (from the top are the exceedances w.r.t. the ICVaR, the  $t$ -deGARCH and the RiskMetrics).

ICVaRfxVaRplot

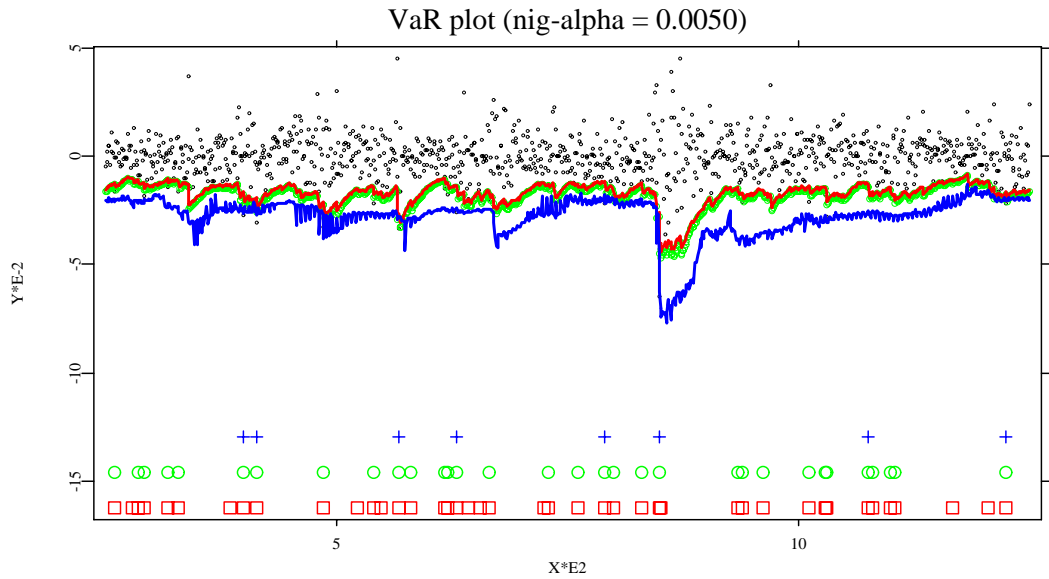


Fig. 6: VaR time plots of the exchange rate portfolio with trading strategy  $b_4 = (-2, 1)^\top$  and risk level  $a = 0.5\%$ . The one-day-ahead predicted VaRs are based on the NIG distribution. The dots are the real portfolio returns, the RiskMetrics (red) and the  $t$ -deGARCH (green) forecasts are displayed as solid line and circles while the ICVaR is displayed as a straight line. The exceedances w.r.t. the risk management models are displayed as cross on the bottom (from the top are the exceedances w.r.t. the ICVaR, the  $t$ -deGARCH and the RiskMetrics).

ICVaRfxVaRplot

$b^T$	$a(10^{-2})$	ICVaR (HYP fit)			ICVaR (NIG fit)			RiskMetrics Method			$t$ -deGARCH Method		
		$n/T\%$	LR1	LR2	$n/T\%$	LR1	LR2	$n/T\%$	LR1	LR2	$n/T\%$	LR1	LR2
(1,1)	5	6.7	5.52	*27.60	5.8	1.28	*28.82	10.8	*53.95	*19.84	11.0	*57.33	*26.05
(1,1)	1	1.9	6.47	*16.29	1.4	1.43	*11.14	5.5	*99.59	*21.70	5.3	*92.67	*22.37
(1,1)	0.5	0.9	2.59	5.58	0.7	0.71	0.00	5.0	*142.33	*23.44	4.2	*106.16	*18.96
(1,2)	5	6.8	6.16	*27.23	5.4	0.32	*30.50	10.9	*55.63	*24.78	11.0	*57.33	*24.48
(1,2)	1	1.6	3.07	*10.51	1.4	1.43	*11.14	5.3	*92.67	*22.37	4.9	*79.30	*23.81
(1,2)	0.5	0.9	2.59	0.00	0.8	1.52	0.00	4.7	*128.43	*21.77	4.4	*114.93	*22.83
(-1,2)	5	6.4	3.80	*38.56	5.5	0.51	*28.17	12.0	*75.40	*35.66	12.0	*75.40	*35.66
(-1,2)	1	1.4	1.43	*11.14	1.2	0.37	*11.86	6.0	*117.58	*25.86	5.5	*99.59	*23.40
(-1,2)	0.5	1.2	*7.06	*11.86	1.1	5.38	*15.84	4.4	*114.93	*18.33	3.2	*65.54	*20.90
(-2,1)	5	6.5	4.34	*22.35	5.0	0.00	*20.34	12.7	*89.18	*18.41	12.7	*89.18	*18.41
(-2,1)	1	1.6	3.07	0.00	1.1	0.09	0.00	5.6	*103.12	*14.79	4.7	*72.87	4.49
(-2,1)	0.5	0.9	2.59	0.00	0.8	1.52	0.00	4.2	*106.16	2.40	3.2	*65.54	5.54

Tab. 8: Backtesting of the VaR forecast of the exchange portfolios: MC simulation with  $M = 10000$ ,  $N = 100$ ,  $T = 1000$ . \* indicates the model is rejected at 99% level.

ICVaRfxVaRplot

## 5 Conclusion

In this paper, we proposed an easy and fast multivariate risk management model. The study is mainly based on the ICA. Instead of estimating the joint density and covariance of high dimensional returns, the searching of ICs transfers the calculation to unidimensional studies. In the empirical study, the proposed ICVaR is superior to the RiskMetrics and *t*-deGARCH methods, above all in the risk level controlling. In addition, in the ICVaR methodology, the joint distribution of portfolio does not rely on trading strategy and therefore can be further applied to calculate VaRs as the investing positions change. Moreover, the ICVaR approach can be easily applied to calculate and forecast other risk measures such as expected shortfall.

## References