TVICA – Time Varying Independent Component Analysis*

Ray-Bing Chen[†], Ying Chen[‡] and Wolfgang K. Härdle[§]

March 9, 2011

Abstract

Source extraction and dimensionality reduction are important in analyzing high dimensional and complex financial time series that are neither Gaussian distributed nor stationary. A time varying independent component analysis (TVICA) is proposed to factorize the data into a linear combination of independent components. The key idea is to allow the ICA filter to change over time, and to estimate it in so-called local homogeneous intervals. The question of how to identify these intervals is solved by the LCP (local change point) method. Compared to a static ICA, the TVICA provides more accurate performance both in simulation and real data analysis.

JEL code: C14; C58; G17

*Acknowledgement: This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk" and the Berkeley–NUS Risk Management Institute at the National University of Singapore.

[†]Department of Statistics, National Cheng Kung University, Taiwan

[‡]Department of Statistics and Applied Probability, Risk Management Institute, National University of Singapore, Singapore

[§]Ladislaus von Bortkiewicz Chair of Statistics, Center for Applied Statistics & Economics, Humboldt-Universität zu Berlin, Germany, and Institute of Statistics, National Central University, Taiwan

Keywords: Adaptive Sequential Testing; Independent Component Analysis; Local Homogeneity; Signal processing

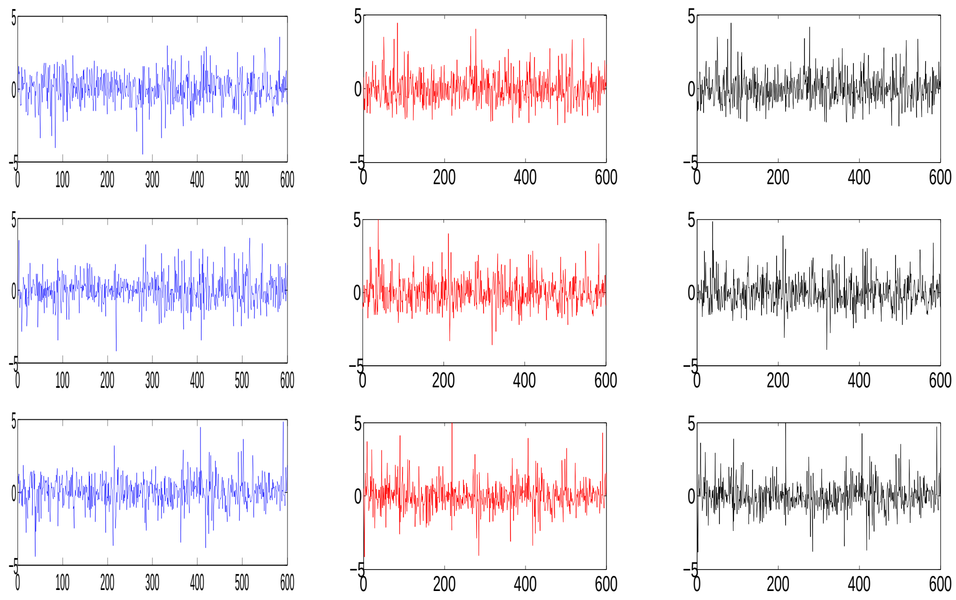
1 Introduction

Source extraction and dimensionality reduction are among the primary goals of multivariate financial time series analysis, which helps to extract features and find latent relations of risk drivers from high dimensional and complex portfolios. With increasing dimension and larger piles of data, attainment of these goals can be challenging.

Conventional statistical methods based on Gaussianity and stationarity do the job of simultaneous dimension reduction and stochastic factor identification. Principal component analysis and factor analysis are the tools here. The assumption of stationarity and Gaussianity is questionable though for the stochastic description of financial data. The Gaussian distribution cannot be used to mark tail dependence of risk factors and it fails in providing the empirical facts like heavy tailedness, volatility clustering and intertemporal dependence of cross moments of order higher than 2. The practical need to retrieve the main driving stochastic factors is accentuated though in risk management and many other fields of applications and must be dealt with even without distributional assumptions. An eigenvalue decomposition of returns' covariance yields uncorrelated factors, see e.g. Jolliffe (2002), Härdle and Simar (2011) and the references therein. With the Gaussian distributional assumption, the factors are independent and this explains why Gaussianity has been widely adopted.

A recently developed multivariate statistical method – Independent Component Analysis (ICA) is different from the conventional approaches. ICA extracts Independent Components (ICs) using a linear filter but does not project onto the eigenvectors of the covariance matrix as PCA does. A rich set of algorithms exists e.g. FastICA proposed by Hyvärinen and Oja (1997) and other methods in Hyvärinen, Karhunen and Oja (2001). The factors are estimated via solving an optimization problem, in which the statistical cross dependence between the extracted ICs is minimized. The dimensionality reduction feature of ICA is that it actually converts a high dimensional problem to a set of univariate ones, and all components are approximately independent. Therefore well-developed univariate methods can be applied to each IC, without considering the dependence among the components anymore. This technique has been implemented in stock returns analysis by Back and Weigend (1998), in risk management by Chen, Härdle and Spokoiny (2010), in high frequency analysis by Kouontchou and Maillet (2007), and in an intertemporal GARCH context by Wu, Yu and Li (2006).

One essential assumption though is common to these papers: the observed series and as well the ICs are stationary and the filter the same for the entire time series. As a consequence, the dynamics of cross dependence is constant over time which in light of the ever occurring turbulences is questionable. In order to demonstrate how the performance of ICA is affected, consider first 3 independent components, each normal-inverse Gaussian (NIG) distributed. The NIG distributions are selected according to the estimated parameters on three ICs obtained for the log returns of Home Depot (HD), Hewlett-Packard (HPQ) and IBM. Two theoretical ICA filters, A_1 and A_2, that are used for generating the mixing series ($X_t = A_t \mathrm{IC}_t$), are learned from the real financial returns over different time periods: 3rd September 2008 to 31th August 2009 (a period with market turbulence), and 30th July 2004 to 29th December 2006 (a relatively quiet period). The first 300 observations are generated by using $A_t = A_1$, and the last 300 time points with $A_t = A_2$. Figure 1 displays, from left to right, (a) the estimated ICs by using the static ICA where the filter is assumed to be constant



(a) (b) (c)

Figure 1: Demonstration: The simulated series are $\mathbf{X}_t = A_t \mathrm{IC}_t$, where $A_t = A_1$ for $t = 1, \cdots, 300$ and changes to A_2 after then. (a) The ICs are estimated based on the simulated series either over the whole sample. (b) The original ICs are shown. (c) The ICs are estimated separately over each stationary sample.

$A_t = A$, (b) the theoretical values of ICs and (c) the estimated ICs by respectively doing ICA based on the first 300 observations and the last 300 observations. We observe that for case (a) the estimated ICs deviate from the theoretical values when the estimation is done over the whole sample in the static ICA. On the other hand, the estimated ICs in case (b) well represent the theoretical independent series when we consider the change of the linear filter.

The above (reality driven) example makes it clear that one not only needs a non-Gaussian low dimensional factor extraction but also a technique that locally (in time) identifies us a “trust interval” over which one can safely do ICA. The importance of identifying such an interval of approximate stationarity is often under-evaluated. The little demonstration above indicates that TVICA is the preferable method when dynamics are changing over time. Improving the quality of IC extraction for varying intervals is the aim here. The question is of course how to identify the intervals in practice! Matteson and Tsay (2009) gave an answer by allowing the mixing matrix to vary over time via a smooth function of other transition variables. This idea is similar to time-varying models proposed in the volatility and co-volatility literature, see e.g. Baillie and Morana (2009), Scharth and Medeiros (2009). Also it resembles time variation models incorporating changes via Markov-Switching or mixture multiplicative error specifications that have been proposed by e.g. Hamilton and Susmel (1994), So, Lam and Li (1998), Lanne (2006). These techniques though take a globally given mechanism for this time variation in contrast to e.g. Mercurio and Spokoiny (2004) who use a local change point approach. This completely data driven approach for filter and homogeneity determination motivates us to develop a local estimation approach for ICA.

Here a time varying ICA (TVICA) framework is put into action, where the mixing matrix (linear filter) is allowed to change over time without imposement of a global structure. For each time point we determine a "trust interval" by conducting a sequence of tests on a structural change. In this selected trust interval one performs ICA. The selection is controlled by a set of critical values. The approach is different from the existing ones in the sense that it is data-driven and applicable for various kinds of breaks (macroeconomic or political changes) with different magnitudes and abrupt or smooth types. Neither prior information (on say states of the market) nor distributional assumption is required.

The remainder of the paper is structured as follows. The next section presents in detail the time varying (constrained) ICA approach and the estimation procedure. Section 3 investigates the performance of the proposed approach along with a simulation study, and the real data analysis is reported in Section 4.

2 How TVICA works

Suppose that there are p assets with log returns $\mathbf{X}_t = \{x_1(t), \cdots, x_p(t)\}^\top$. The aim is to factorize the financial returns into a linear combination of independent components $\mathbf{Z}_t = \{z_1(t), \cdots, z_p(t)\}^\top$. The TVICA approach is based on:

$$\mathbf{X}_t \quad = \quad A_t \mathbf{Z}_t \tag{1}$$

where A_t is a $p \times p$ time varying matrix. In the static ICA approach, the observed series $\mathbf{X}_t$ in (1) are assumed to be stationary and $A_t = A = \text{const}$ i.e. to be time homogeneous. Here the linear filter A_t is time dependent and the estimation of ICs is customized under *Local Homogeneity* for any time point of interest.

Local homogeneity means that, for any particular time point t there exists a past time interval $I_t = [t - m_t, t]$, over which the linear filter A_t is approximately constant, i.e. $A_s \approx A$, $\forall\ s \in I_t$. Given t and its past information, the challenge is of course to determine I_t (or m_t) – the trust "interval of local homogeneity". In order to rise to this challenge, the Local Change Point (LCP) detection method of Mercurio and Spokoiny (2004) is applied. Note that the LCP method nests the above mentioned "smooth transition" and "regime switching" techniques used in earlier literature. Based on the identified interval, TVICA can provide more accurate performance than a constant ICA filter.

2.1 The LCP method

In this section, we present the LCP detection procedure to identify the interval of local homogeneity at point t. The estimation of the TVICA is carried out via the (quasi) maximum likelihood ICA method by treating the mixing matrix (or the inverse of the mixing matrix) as the unknown parameters. Suppose that at time point τ, an interval of *homogeneity* $I_t = [t - m_t, t]$ is given with m_t indicating the length of the interval. Then with pdf $f_j(z_j)$ of IC z_j, $j = 1, \cdots, p$, the pdf of $\mathbf{X}$, according to Jacobian transformation, is:

$$f_{\mathbf{X}}(x_1, \cdots, x_p) = \{\prod_{i=1}^{p} f_j(z_j)\} \times |\det B_t|,$$

where B_t is the inverse of A_t. With $B_t = (b_{1t}, \ldots, b_{pt})^\top$, this gives:

$$f_{\mathbf{X}}(x_1, \ldots, x_p) = |\det B_t| \prod_{j=1}^{p} f_j(b_{jt}^\top \mathbf{X}).$$

The log-likelihood function on the interval I_t is:

$$L(I_t, B_t) = \sum_{t=1}^{m_t} \sum_{j=1}^{r} \log\{f_j(b_{jt}^\top \mathbf{X}_t)\} + T \log |\det B_t|,$$

and the MLE is denoted as $\widetilde{B}_t$.

Relaxing this to *local homogeneity* on I_t means that B_t does not deviate too much from a constant filter. More precisely, using this almost constant parameter gives roughly a small modeling bias that measures the divergence of a time varying model to a static model, see Spokoiny (2009). Take now a family of nested intervals, $I_0 \subset I_1 \subset \cdots \subset I_{K-1} \subset I_K$ (the subscript t is omitted for notation simplification), the LCP method attempts to find the longest interval of local homogeneity among them. The longer the length of intervals, the smaller the variance of the estimator (under local homogeneity) but the higher the bias.

The identification of the trust interval is done via a sequential algorithm. At the first step at time t, the interval I_0 is accepted. Next for an interval I_k, $k = 1, \cdots, K$, the procedure is to sequentially screen $J_k = I_k \setminus I_{k-1} = [t - m_k, t - m_{k-1}]$ and check it for a possible change point. The interval I_k is accepted if every point in J_k is tested to be insignificant as a location of change point. One continues this way until a change point is detected or the longest interval I_K is reached. Otherwise, the algorithm is terminated and the last accepted interval is selected.

More specifically in the k-th step, given J_k as the testing interval we choose $I = [t', t'']$ to be a larger interval such that $J_k \subset I$. Then for each point $t \in J_k$, we separate the interval I into two sub-intervals, $I' = [t', t)$ and $I'' = [t, t'']$. That is $I = I' \bigcup I''$ and $I' \bigcap I'' = \emptyset$. Figure 2 demonstrates the relation of the intervals used in the testing procedure. Let $L_I(B)$ be the log-likelihood function for the observations in I. The

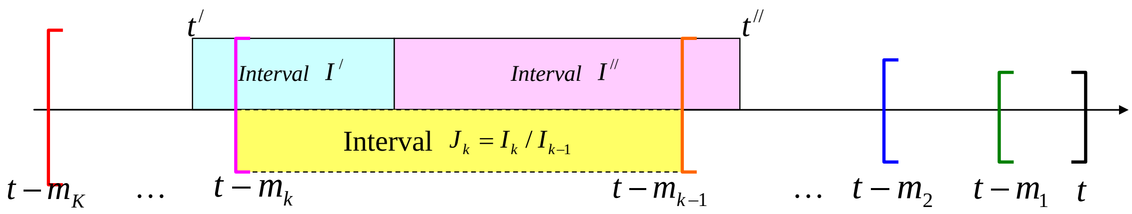


Figure 2: Local change point detection procedure.

LCP method employs the likelihood ratio:

$$T_{I,t} \quad = \quad \max_{B'', B'} \{L_{I''}(B'') + L_{I'}(B')\} - \max_{B} L_I(B). \tag{2}$$

This test statistic (2) is only calculated at one $t \in J_k$. It is therefore not indicating "a possible change point" over the whole interval J_k. To do so, the maximum (over t) of (2) is the finally used test statistics:

$$T_k = \max_{t \in J_k} T_{I,t} \tag{3}$$

If T_k is greater than a critical value η_k, the null hypothesis of local homogeneity on I_k

is rejected. The critical values $\{\eta_k\}$ are computed in Monte Carlo simulation, since the distributional properties of (3) are even asymptotically unknown. The details are described in Section 2.2.

The formal definition of the LCP algorithm is as follows:

1. Initialization: The null is not rejected on I_0. Denote the initial homogeneous estimate by $\widehat{B}_t^{(0)} = \widetilde{B}_t^{(0)}$.

2. Set $k = 1$. While $T_k \leq \eta_k$ and $k \leq K$,
   update the present homogeneous estimate by $\widehat{B}_t^{(k)} = \widetilde{B}_t^{(k)}$ and set $k = k + 1$.

3. Final Estimate: $\widehat{B}_t = \widehat{B}_t^k$, which is actually the maximum likelihood estimate from the longest interval of local homogeneity.

It is worth mentioning that the numerical complexity of the LCP algorithm is not high.

2.2 Selection of Hyperparameters

The LCP method is driven by a small set of "adjustable screws" or hyperparameters that we present here.

Set of interval: The family of intervals $\{I_k\}_{k=0}^{K}$ is either given or selected as:

$$I_k = [t - m_k, t],$$

where $m_k = m_0 a^k$ with a pre-specified initial length m_0 and a multiplier $a > 1$. The coefficient a controls the increasing speed of the candidate intervals. The starting value m_0 should be sufficiently small to provide a reasonable local homogeneity.

Critical values: The critical values $\{\eta_k\}$ are calculated under the null, i.e. homogeneity.

Under the null, one generates M sets of independent samples and mixes them with a constant filter matrix A^* ($B^* = A^{*-1}$). Conventionally, the maximum likelihood estimate over the whole sample period is used as the constant. For any $r > 0$ and $\rho > 0$, the fitted log likelihood with $B_t = B^*$ for all $t \in I_K$ satisfies:

$$\mathsf{E}_{B^*}|L_{I_k}(\widetilde{B}_k, B^*)|^r \leq \rho R_r(B^*), \tag{4}$$

where $L_{I_k}(\widetilde{B}_k, B^*) = L_{I_k}(\widetilde{B}_k) - L_{I_k}(B^*)$ and $R_r(B^*) = \max_{k \leq K} \mathsf{E}_{B^*}|L_{I_k}(\widetilde{B}_k, B^*)|^r$. Note that given the values of r and ρ, $R_r(B^*)$ can be computed straightforwardly.

The parameter r specifies the loss function (4) under the null. The parameter ρ is similar to the test level parameter. Small values of ρ indicates that one expects a small divergence of the estimate to a constant filter (the null), which leads to relatively large critical values and a rather conservative procedure for possible time variation. Increasing ρ would result in a decrease of the critical values and an increase of the sensitivity of the method to the changes of filter in the underlying process.

In the homogeneous situation, the longest interval I_K is the optimal choice. The selected final estimate $\widehat{B} = \widehat{B}_K$ that depends on the critical values $\eta_1, \cdots, \eta_K$ is expected to perform as good as B^* in the sense that a "feasible" version of (4) is satisfied:

$$\mathsf{E}_{B^*}|L_{I_K}(\widetilde{B}_K, \widehat{B})|^r \leq \rho R_r(B^*), \tag{5}$$

where we mimic the practical situation, i.e. the constant B^* is unknown. Notice that the sequential tests accumulate uncertainty in estimation due to the increase in the degrees of freedom. To take this into account, we require that at each step $k = 1, \cdots, K$, the present "final" estimate $\widehat{B}_k$ provides the prescribed performance on the interval I_k in the sense:

$$\mathsf{E}_{B^*}|L_{I_k}(\widetilde{B}_k, \widehat{B}_k)|^r \leq \frac{k}{K}\rho R_r(B^*). \tag{6}$$

where $\widehat{B}_k$ depends on all the critical values $\eta_1, \cdots, \eta_k$.

Now we select the critical values sequentially. At the initial step $k = 0$, we set $\eta_0 = \infty$ in agreement with the local homogeneity in the shortest interval. To specify the next critical value η_1, we set the values of $\eta_2, \ldots, \eta_K$ to be infinity. Then η_1 is selected as the minimum value to provide the prescribed performance:

$$\mathsf{E}_{B^*}\left|L_{I_k}\left\{\widetilde{B}_k, \widehat{B}_k(\eta_1, \eta_2)\right\}\right|^r \leq \frac{1}{K}\rho R_r(B^*), k = 1, \ldots, K.$$

We then continue to select η_k given $\eta_1, \ldots, \eta_{k-1}$ and set $\eta_{k+1} = \cdots = \eta_K = \infty$, $k = 2, \cdots, K$. The value of η_k is determined in the sense:

$$\mathsf{E}_{B^*}\left|L_{I_l}\left\{\widetilde{B}_l, \widehat{B}_l(\eta_1, \ldots, \eta_k)\right\}\right|^r \leq \frac{k\rho R_r(B^*)}{K}, \quad l = k, \ldots, K.$$

It is worth mentioning here that the LCP procedure is robust w.r.t. these hyper-parameters. In a later simulation study, we will show that the performance of the LCP detection approach is stable to the selection of these parameters.

2.3 Finding ICs in a selected interval

Given an identified interval of local homogeneity, (quasi) maximum likelihood estimation can be used to obtain ICs. For leptokurtic original sources, one considers the log density:

$$\log f_j(x_j) = \alpha_1 - 2\log\cosh(x_j) = \alpha_1 - 2\log\{\frac{1}{2}(e^{x_j} + e^{-x_j})\}, \quad (7)$$

where α_1 is a normalizing constant to make this function a pdf. Let the derivative of $\log f_j$ be $g_j(x_j) = \frac{\partial}{\partial x_j}\log f_j(x_j) = \frac{f_j'(x_j)}{f_j(x_j)}$. We have:

$$g_j(x_j) = -2\tanh(x_j) = -\frac{2\{\exp(2x_j) - 1\}}{\exp(2x_j) + 1}, \quad \forall\, j = 1, \ldots, p, \quad (8)$$

The motivation of this selection is that the log density is close to the absolute value that would give Laplace density, see Hyvärinen and Oja (1999). Moreover small misidentification in the density doesn't affect the local consistency of the ML estimator, see Hyvärinen et al. (2001).

3 Simulation

This section investigates the performance of the TVICA method in different scenarios. In particular, we assess its detection power under homogeneity and a situation with a change point. Under homogeneity, LCP should select the longest interval in the estimation of ICs. In a scenario with a change point, LCP is expected to detect a change point, and further to locate the position of the change point properly. The ICs are estimated from the identified interval of local homogeneity. We also analyze the impact of the hyperparameters (r, ρ) on the LCP algorithm. It turns out that they have little influence on the performance of TVICA.

The setup of the simulation scenarios are practical. The real data (10 highly traded stocks at NYSE) are used to generate the simulation processes. The 10 stocks are The Home Depot (HD), Hewlett-Packard (HPQ), IBM, Intel (INTC), Johnson & Johnson (JNJ), JPMorgan Chase (JPM), Coca-Cola (KO), McDonald's (MCD), 3M (MMM) and Merck (MRK). First we choose the historical log returns of these stocks over an approximately stationary time period from 7th December 2009 to 28th October 2010. Then we estimate ICs by selecting a quasi log likelihood g as described in (8). Due to a universally good description of NIG distribution (one heavy-tailed distribution, see Barndorff-Nielsen (1997) for more details), we assume that the ICs are NIG distributed and estimate the distribution. Accordingly, we generate 10 independent univariate series, with 1210 sample points for each series and with 1000 replications. The simulated observations are obtained by mixing these independent sources with the mixing matrix A_t. Two kinds of scenarios are discussed here: a scenario under homogeneity where the matrix A_t is an identity matrix and time independent and two scenarios with a change point occurred close to the boundaries of an interval respectively.

The set of the time intervals is defined with $m_0 = 225$, $a = 1.4$ and $K = 5$, which corresponds to the investment horizon from one year to 5 years. The parameters (r, ρ) in the LCP procedure are assigned to be $(0.5, 0.5)$ and $(0.1, 0.1)$, where the first selection is a conventional selection and the second is close to extremal. For computation of the critical values, we fix the distribution of ICs over the whole sample period and generate 5000 independent series, with 1210 sample points for each. When we apply ICA for the interval $J_k = I_k/I_{k-1}$, we set I to be a superset of J_k that also includes the neighboring 25 observations of J_k. The critical values for different scenarios are displayed in Figure 3. The critical values are decreasing corresponding to the fact for longer interval, probability of rejecting the null increases.

The detection results are reported in Table 1. The ratio of rejecting the null over the 1000 replications is reported. Under scenario of homogeneity, only little (6.2%

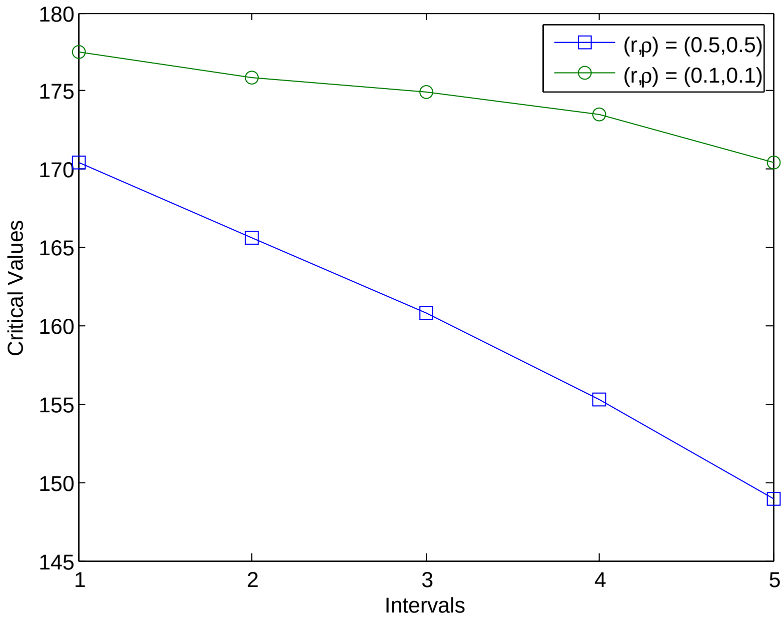


Figure 3: Critical values for two sets of parameters: $(r, \rho) = (0.5, 0.5)$ and $(r, \rho) = (0.1, 0.1)$. The computations are based on the generate oracle independent series, with 1210 sample points for each series and with 5000 replications.

or 0.6%) is rejected. Under scenario with change point, the location of the change point has been meaningfully detected. When the change point $t = 351$ is close to the left end-point of the interval $I_4/I_3 = [\mathbf{346}, 593]$, about 99% of the interval is actually homogeneous. In this case, one may not want to discard so many useful observations in interval I_4. The ratio of rejecting the null is 0.6% (or 0.1%) in interval I_4 and 99.4% (or 99.9%) in interval I_5. Alternatively, when the change point $t = 501$ is around the right boundary $I_4/I_3 = [346, \mathbf{593}]$, the ratio of rejecting the null becomes 99.7% (or 95.9%) in interval I_4. Moreover, the ratio of rejecting the null under scenario with change point, no matter where the change point is detected, is 100%. In general, the ratio of rejecting the null is reasonable for all scenarios. We also notice that the performance is very stable for different set of (r, ρ), since the TVICA method is data-driven.

Parameter (r, ρ)	Homogeneity	Change Point at $t = 351$ $I_4/I_3 = [\mathbf{346}, 593]$	Change Point at $t = 501$ $I_4/I_3 = [346, \mathbf{593}]$
$(0.5, 0.5)$	0.062	$@I_4 = 0.006$, $@I_5 = 0.994$	$@I_4 = 0.997$, $@I_5 = 0.003$
$(0.1, 0.1)$	0.006	$@I_4 = 0.001$, $@I_5 = 0.999$	$@I_4 = 0.959$, $@I_5 = 0.041$

Table 1: The rejection ratio of the LCP detection tests for 1000 replications. In the scenario of homogeneity, A_t is an identity matrix, while in the scenarios with change point, the $(2, 1)$-component of A_t is changed from 0 to 2. The results show the TVICA method can detect the local change point precisely, if existed. In addition, the method also works well under homogeneity.

4 Real Data Analysis

In this section, we implement TVICA to 6 highly traded stocks at NYSE: The Home Depot (HD), Hewlett-Packard (HPQ), IBM, Intel (INTC), Johnson & Johnson (JNJ) and JPMorgan Chase (JPM). Does the proposed method detect intervals of local homogeneity? Do we identify the intervals in a post-financial crisis world and in a relatively stationary situation?

The first experiment considers the time interval from 30th March 2007 to 31st August 2009, during which the stock market crash occurred in 2008. The set of intervals for testing is defined as $m_0 = 200$, $a = 1.25$ and $K = 5$. The initial interval is $I_0 = [2008/11/12, 2009/08/31]$, over which we use the quasi likelihood approach to estimate ICs. The results are then used to generate independent series with 5000 replications for leaning critical values. The parameters $(r, \rho) = (0.5, 0.5)$ and $(r, \rho) = (0.1, 0.1)$ are considered respectively. Again we set I to be a superset of $J_k = I_k/I_{k-1}$ that also includes the neighboring 25 observations of J_k.

Given the sample period, we expect a detection of local change point around year 2008. The influence of the financial crisis 2008 will definitely remain for a long while, however it would be interesting to ask whether the financial markets have reached a new (approximately) stationary situation, though with a different perspective on structure. If this is the case, when does the stationary world start? The test statistic and the critical values are reported in Table 2. According to the table, we shall reject the hypothesis of local homogeneity at the interval I_2/I_1, see Figure 4. In other words, the interval $[2008/09/03, 2009/08/31]$ is stationary.

Figure 4: Identified interval of local homogeneity. The hypothesis is rejected at the interval I_2/I_1. In other words, the interval $[2008/09/03, 2009/08/31]$ is stationary. The results are same for $(r, \rho) = (0.5, 0.5)$ and $(r, \rho) = (0.1, 0.1)$.

The other experiment considers the same stocks and uses the same parameter selection as the first experiment. However, the selected sample period is different, from 30th July 2004 to 29th December 2006. During this period, no influential economic or financial events occurred so that the sample interval can be estimated as stationary. We expect that the TVICA method select the longest interval as an interval of local homogeneity. The test statistic and the critical values are reported in Table 2. The results indicate that the stationarity assumptions are verified for the whole time interval.

References

Back, A. and Weigend, A. (1998). A first application of independent component analysis to extracting structure from stock returns, *International Journal of Neural Systems* **8**: 473–484.

	2007/03/30-2009/08/31			2004/07/30-2006/12/29		
	CV		T_I	CV		T_I
(r, ρ)	$(0.5, 0.5)$	$(0.1, 0.1)$		$(0.5, 0.5)$	$(0.1, 0.1)$	
I_1	66.50	78.68	43.66	59.32	68.98	21.77
I_2	61.46	74.90	**87.04**	54.89	65.81	32.05
I_3	58.35	73.24	**76.55**	51.75	64.02	42.88
I_4	54.37	70.79	**136.95**	49.21	62.35	34.92
I_5	51.97	69.25	**186.78**	46.68	61.59	39.10

Table 2: The critical values and the test statistic for two experiments. The set of intervals is $\{200 \times 1.25\}_{k=0}^{5}$. The parameters are set to be $(r, \rho) = (0.5, 0.5)$ and $(r, \rho) = (0.1, 0.1)$. The critical value computations are based on the generate 6 independent series, with 610 sample points for each series and with 5000 replications.

Baillie, R. T. and Morana, C. (2009). Modelling long memory and structural breaks in conditional variances: An adaptive FIGARCH approach, *Journal of Economics Dynamics and Control* **33**: 1577–1592.

Barndorff-Nielsen, O. (1997). Normal inverse gaussian distributions and stochastic volatility modelling, *Scandinavian Journal of Statistics* **24**: 1–13.

Chen, Y., Härdle, W. and Spokoiny, V. (2010). GHICA risk analysis with GH distributions and independent components, *Journal of Empirical Finance* **17**: 255–269.

Hamilton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics* **64**: 307–333.

Härdle, W. and Simar, L. (2011). *Applied Multivariate Statistical Analysis*, 3rd edn, Springer-Verlag Berlin Heidelberg New York.

Hyvärinen, A., Karhunen, J. and Oja, E. (2001). *Independent Component Analysis*, John Wiley & Sons, Inc.

Hyvärinen, A. and Oja, E. (1997). A fast fixed-point algorithm for independent component analysis, *Neural Computation* **9**: 1483–1492.

Hyvärinen, A. and Oja, E. (1999). Independent component analysis: Algorithms and applications, *Neural Networks* **13**: 411–430.

Jolliffe, I. T. (2002). *Principal component analysis*, Springer-Verlag Berlin Heidelberg New York.

Kouontchou, P. and Maillet, B. (2007). ICA-based high frequency VaR for risk management, *ESANN'2007 proceedings - European Symposium on Artificial Neural Networks*, Bruges, Belgium.

Lanne, M. (2006). A mixture multiplicative error model for realized volatility, *Journal of Financial Econometrics* **4**: 594–616.

Matteson, D. S. and Tsay, R. S. (2009). Independent component analysis for multivariate financial time series. working paper, submitted.

Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time-inhomogeneous volatility models, *Ann. Statist.* **12**: 577–602.

Scharth, M. and Medeiros, M. C. (2009). Asymmetric effects and long memory in the volatility of dow jones stocks, *International Journal of Forecasting* **25**: 304–327.

So, M. K. P., Lam, K. and Li, W. K. (1998). A stochastic volatility model with markov switching, *Journal of Business & Economic Statistics* **16**: 244–253.

Spokoiny, V. (2009). Parameter estimation in time series analysis. Preprint no. 1404, WIAS, Berlin.

Wu, E., Yu, P. and Li, W. (2006). Value at risk estimation using independent component analysis-generalized autoregressive conditional heteroscedasticity (ICA-GARCH) models, *International Journal of Neural Systems* **16**: 371–382.