Eur. Phys. J. B **60**, 101–109 (2007)
DOI: 10.1140/epjb/e2007-00322-1

THE EUROPEAN PHYSICAL JOURNAL B

Correlation and volatility in an Indian stock market: A random matrix approach

Varsha Kulkarni[1,2] and Nivedita Deo[1,a]

[1] Department of Physics and Astrophysics, University of Delhi, Delhi-110007, India
[2] Department of Statistics, University of Wisconsin-Madison, Medical Science Center, 1300 University Avenue, Madison, WI 53706, USA

Received 12 December 2005 / Received in final form 5 October 2007
Published online 29 November 2007 –

Abstract. We examine the volatility of an Indian stock market in terms of correlation of stocks and quantify the volatility using the random matrix approach. First we discuss trends observed in the pattern of stock prices in the Bombay Stock Exchange for the three-year period 2000–2002. Random matrix analysis is then applied to study the relationship between the coupling of stocks and volatility. The study uses daily returns of 70 stocks for successive time windows of length 85 days for the year 2001. We compare the properties of matrix C of correlations between price fluctuations in time regimes characterized by different volatilities. Our analyses reveal that (i) the largest (deviating) eigenvalue of C correlates highly with the volatility of the index, (ii) there is a shift in the distribution of the components of the eigenvector corresponding to the largest eigenvalue across regimes of different volatilities, (iii) the inverse participation ratio for this eigenvector anti-correlates significantly with the market fluctuations and finally, (iv) this eigenvector of C can be used to set up a Correlation Index, CI whose temporal evolution is significantly correlated with the volatility of the overall market index.

PACS. 89.65.Gh Economics; econophysics, financial markets, business and management – 89.65.-s Social and economic systems – 89.75.-k Complex systems

1 Introduction

Physical phenomena such as Brownian motion, turbulence, chaos, have recently found application in the study of dynamics of financial markets [1–3]. An important question of interest is whether and how volatility (a measure of the market fluctuations) affects the response of market dynamics and vice versa. It has been observed [4–8] that there is an accentuated rise in the overall stock index correlations during a financial crisis. Empirical relations between volatility and information exchange (news) have also been established in recent studies [9]. In this paper, we explore the relationship between two a priori distinct properties of the market, one, volatility an index of market fluctuations, and two, the coupling of stocks with one another captured by the correlation matrix. While there already exists evidence that both fluctuation and correlation are enhanced at the same time, we attempt here to explore this relationship systematically and quantitatively for the Indian Stock market.

A number of researchers [10–20] have applied the methods of Random Matrix Theory (RMT) to financial data and found interesting clues about the underlying interactions. However, while the regression methodology [21,22] has focussed keenly on prediction and finding potential indicators of volatility, RMT has been applied little in that area. The purpose of this paper is two-fold. First, it attempts to understand empirically the closely related aspects of volatility and correlation in the market using RMT. Second, it aims to show that this technique may be used to set up a quantitative indicator which can potentially serve as a measure of correlations of the market.

The paper is organized as follows. Section 2 provides a brief empirical analysis of the Bombay Stock Exchange (BSE) index and shows the volatility pattern. Trends are identified on the basis of some commonly observed features of volatile versus non-volatile situations. Section 3 deals with the random matrix approach and the financial correlation matrix. Section 4 concludes with a discussion of the key observations.

2 Empirical analysis of the Indian stock market BSE index

2.1 Data analyzed

This section uses the daily indices of the Bombay Stock Exchange for a period of 3 years between 2000–2002. Each

[a] e-mail: ndeo@physics.du.ac.in

year corresponds to approximately 250 days of elapsed time and the total number of data points in this set is 750. We consider the opening values of indices to be continuous by removing the holidays.

BSE SENSEX: The BSE-Sensitive Index, (or SENSEX) is a "market capitalization-weighted" index of 30 stocks representing a sample of large, liquid and representative companies. In particular, the selection of these stocks is based on certain technical specifications [23] of criteria such as liquidity, continuity, industry representation, market capitalization, track record, etc. The value of the index at any point reflects the effective market value of 30 component stocks relative to a base period. The total market value of the stocks in the index is set equal to 100. BSE-SENSEX is the oldest index in India and is in effect the proxy for the Indian stock markets.

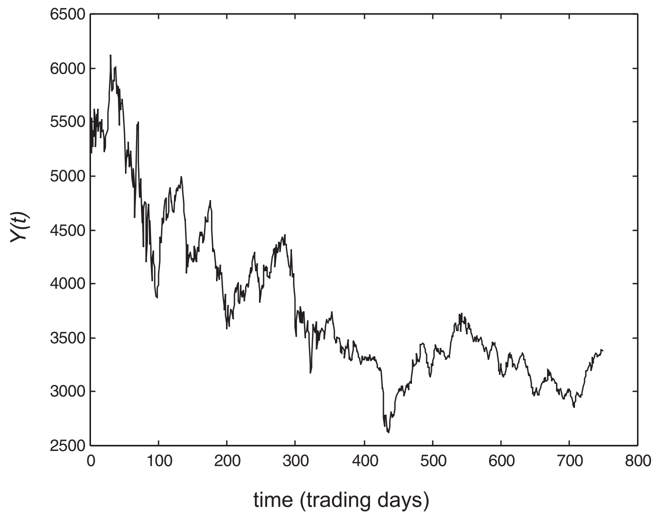

Fig. 1. BSE index plotted for all the days reported in the period 2000–2002. Total number of points plotted is 750. A sharp dip can be seen around September 11th, 2001 (425th day in figure) when the index drops to the lowest.

2.2 Volatility

Our present focus is on linking volatility to correlations in BSE. Hence, it would be worthwhile to understand the volatility pattern of BSE based on the above assertions, before moving to the random matrix treatment.

2.2.1 Computing volatility:

We consider $Y(t-\Delta t), Y(t), Y(t+\Delta t), \ldots$ to be a stochastic process. $Y(t)$ may represent time series of prices, indices, exchange rates etc. The logarithmic returns $G(t)$ over time scale Δt are

$$G(t) = \log(Y(t + \Delta t)) - \log(Y(t)), \qquad (2.1)$$

Δt refers to the time interval. Here $\Delta t = 1$ day.

We quantify volatility, as the local average of the absolute value of daily returns of indices in an appropriate time window of T days,

$$v = \frac{\sum_{t=1}^{T-1} |G(t)|}{T-1}. \qquad (2.2)$$

We compute volatility for the three year period $2000-2002$ by taking $T = 20$ days, that is typically a month of trading time in BSE. This method comes close to the computation of historical volatility in literature. The results here are for small time periods though the best estimation of volatility involves use of larger time periods. Alternative estimators have been introduced in different contexts [24].

The BSE index for the period 2000–2002, is shown in Figure 1. The figure shows a significant change in the value of index over the period of three years 2000–2002. The rate of change (decrease) appears to be higher for the first 450 days than later. It seems the market was far more active in the year 2000 than 2001 or 2002. There is a sharp dip in Y near 425th point in the data set, after which the index rises and settles without much fluctuation. This point represents the day 9/11/2001. The event that occurred on this day in the USA had a rattling effect on markets worldwide. $\sum |G(t)|$ may be considered a substitute for volatility and we refer to it as the 'scaled volatility'.

Figure 2 shows the volatility of the market in the period 2000–2002. Each year corresponds to 250 days of elapsed time. We may divide the period into three subperiods:
T_1: 1–250 days, year 2000 (scaled volatility = 5.65),
T_2: 251–500 days, year 2001 (scaled volatility = 3.5) and
T_3: 501–750 days, year 2002 (scaled volatility = 2.25).

We observe that T_1 (the year 2000) was an extremely active subperiod marked by very high fluctuations in the market. Also, regimes of high volatility seem to occur in clusters. Fluctuations are consistently high in say the first 200 days and more. Subsequently they decrease in T_2 (year 2001). The sudden rise observed around the 425th day indicates the effect of the 9/11/2001 event. However its impact on the market was not long lasting and a quiescent state T_3, followed soon after.

3 Random matrix approach

Random Matrix Theory, developed originally [25] to study the interactions in complex quantum systems has been useful in the analysis of universal and non-universal properties of cross-correlations between different stocks. Recently various studies [10–20] have quantified correlations between different stocks by applying concepts and methods of RMT. They have shown that deviations of the properties of correlation matrix of stock-price fluctuations, from a random correlation matrix yield information about the actual correlations existing in the market. Here we compare volatile versus less volatile situations from the point of view of correlations, participation of stocks in the

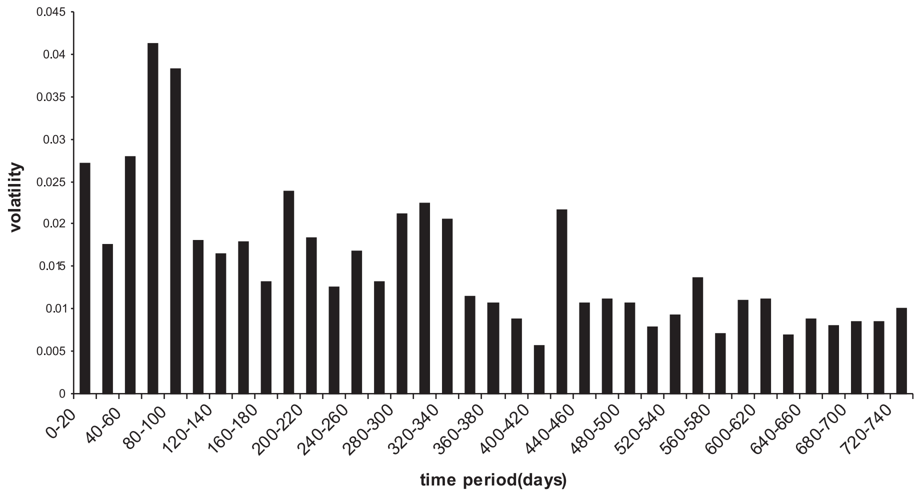

Fig. 2. Volatility, v of BSE index for 20-day periods between 2000–2002. Last ten days of the period have been ignored. The period 420–440 days including the date of September 11th 2001 shows a sudden burst of activity.

market and try to quantify volatility in terms of the deviations.

3.1 Description of data used for the correlation analysis

As mentioned earlier, BSE consists of stocks from various sectors. Many of the stocks are not actively traded and hence not reported regularly in any period of time. Consequently they do not contribute much to the variations in stock price indices. We consider here seventy stocks from largest sectors such as chemical industry, metal and non-metal (diversified including steel, aluminum, cement etc) whose data exists for the full period under consideration[1]. Analysis is done for the period 280–500 days, a subperiod in the year 2001. The data set studied here is that of intra day prices of seventy stocks of BSE which should be distinguished from the previous data set analyzed in Section 2, the index of 30 stocks of BSE-SENSEX.

3.2 Cross correlations

In terms of the price of stock i at time t, $P_i(t)$ ($i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$) we quantify correlations for $T-1$ observations of inter day price changes (returns) as

$$G_i(t) = \log P_i(t+1) - \log P_i(t). \tag{3.3}$$

Since different stocks vary on different scales, we normalize the returns as

$$M_i(t) = \frac{G_i(t) - \langle G_i \rangle}{\sigma_i} \tag{3.4}$$

where $\sigma_i = \sqrt{\langle G_i^2 \rangle - \langle G_i \rangle^2}$ is the standard deviation of G_i and $\langle G_i \rangle = \frac{1}{T-1} \sum_{t=1}^{T-1} G_i(t)$. Then the cross correlation matrix C, measuring the correlations of N stocks is

[1] Questions regarding the data may be addressed to the authors at kulkarni@stat.wisc.edu/ndeo@physics.du.ac.in.

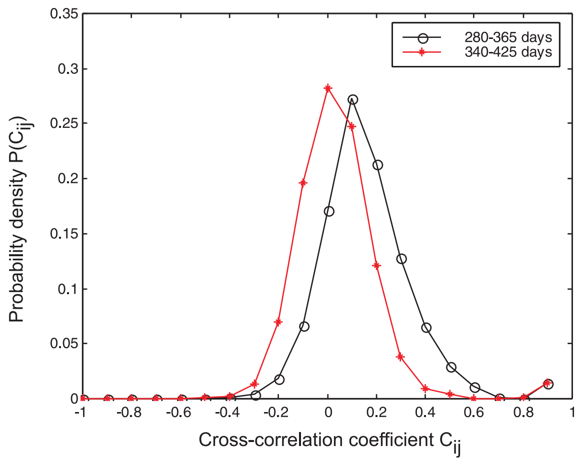

Fig. 3. Plot of the probability density of elements of correlation matrix C calculated using daily returns of 70 stocks for two 85 day analysis periods (i) 280–365 days and (ii) 340–425 days with scaled volatilities (computed as $\sum |G(t)|$) of 1.6 and 0.82 respectively. We find a large value of average magnitude of correlation $\langle |C| \rangle = 0.22$ for (i) compared to $\langle |C| \rangle = 0.14$ for (ii).

constructed with elements

$$C_{ij} = \langle M_i(t) M_j(t) \rangle. \tag{3.5}$$

The elements of C lie between $-1 \leq C_{ij} \leq 1$ where $C_{ij} = 1$ corresponds to complete correlation, $C_{ij} = 0$ corresponds to no correlation and $C_{ij} = -1$ corresponds to complete anti-correlation.

We construct the cross correlation matrix C from daily returns of $N = 70$ stocks for two analysis periods of 85 days each, T_a: 280–365 days with scaled volatility $\sum |G(t)| = 1.6$ and T_b: 340–425 days with scaled volatility $\sum |G(t)| = 0.82$) (see Fig. 2). The probability densities of elements of C, $P(C_{ij})$ for both periods are compared in Figure 3. We see that the distribution for period T_b is more symmetric, implying that both positive and negative

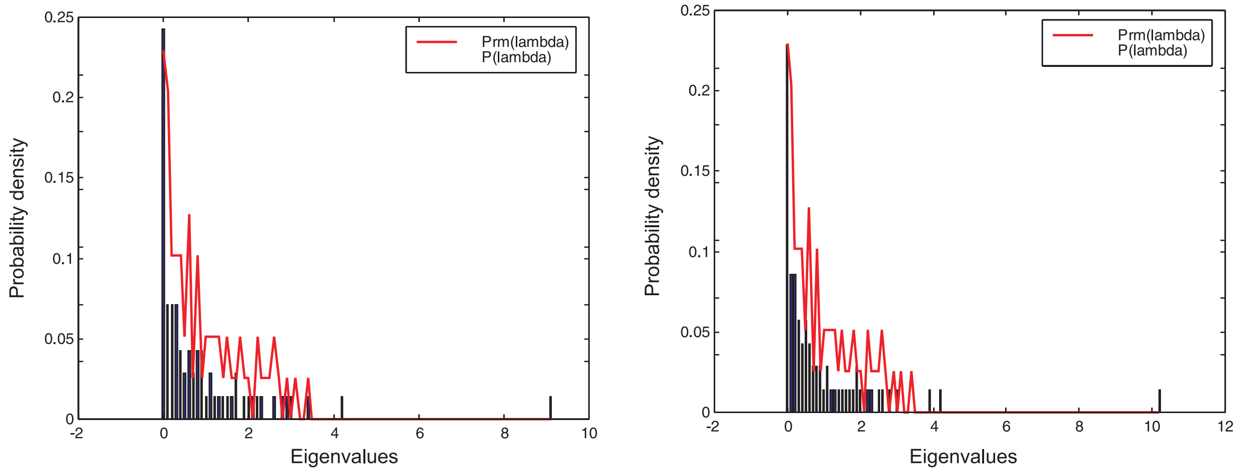

Fig. 4. Probability density of eigenvalues is shown by bars for a period considered (i) 334–419 before 9/11/2001 and having a volatility (scaled) of 0.8 (Left) and (ii) 346–431 including 9/11/2001 and having a volatility (scaled) of 0.9 (Right). A comparison is made with the probability density of eigenvalues of a random matrix R of the same size as C, shown by the solid line. The number of deviating eigenvalues is 4 in (i) and 6 in (ii). Largest eigenvalue for (i) is 9.17 and for (ii) is 10.28.

correlations are more or less equal in extent. T_a, however is characterized by a larger, positive mean. The figure also suggests that there is a lower value of $P(C_{ij})$ in higher levels of correlation magnitudes in the less volatile period T_b as compared to the more volatile period T_a. The existence of more pronounced correlations in periods of high volatility is indicated in Figure 5. The simple Pearson's correlation coefficient [26] between the $\langle |C| \rangle$ and volatility is found to be 0.94 which is highly significant.

3.3 Statistics of eigenvalues of C

The eigenvalues of C have special implications in identifying the nature of the correlations. In the past, studies using RMT methods have analyzed C and shown that 98% of eigenvalues of C lie within the RMT limits whereas 2% of them lie outside [10]. It is known that the largest eigenvalue deviating from RMT prediction reflects that some influence of the full market is common to all stocks, and that it yields information about the actual correlations in the market. The next few sub-leading eigenvalues carry information regarding the market sectors. The range of eigenvalues within the RMT bounds corresponds to noise and does not yield any system specific information.

3.3.1 Eigenvalue distribution of the correlation matrix

In order to extract information about the cross correlations from the matrix C, we compare the properties of C with those of a random correlation matrix. We now define a random correlation matrix as

$$R = \frac{AA^{\top}}{T} \tag{3.6}$$

where A is $N \times T$ matrix with random entries (zero mean and unit variance) that are mutually uncorrelated. Statistics of random matrices such as R are known. In the limit of both N and T tending to infinity, such that $Q = T/N(> 1)$ is fixed, it has been shown that the probability density function Prm (λ) of eigenvalues of R is given by [14,15]

$$\mathrm{Prm}(\lambda) = \frac{Q\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\lambda} \tag{3.7}$$

for λ lying in $\lambda_- < \lambda < \lambda_+$ where λ_- and λ_+ are the minimum and maximum eigenvalues of R, respectively given by

$$\lambda_{\pm} = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}. \tag{3.8}$$

We set up two correlation matrices C from the daily returns of $N = 70$ stocks for $T = 85$ days for two time periods in the year 2001. The periods are chosen so that one of them includes the data reported on the day - September 11, 2001 (85th day being September 18, 2001) and the other one does not (85th day being August 31, 2001). Here $Q = 1.21$, and maximum and minimum eigenvalues predicted by RMT from (3.8) are 3.6385 and 0.0086. See Figure 4.

Figure 4 indicates an increased deviation in case of a perturbed correlation matrix (including the shock of 9/11/01), Figure 4 (ii), as compared to a non-perturbed one (excluding the shock of 9/11/01), Figure 4 (i). The deviation is observed in terms of the number of eigenvalues lying outside RMT range and the magnitude of the maximum eigenvalue. In the case of non-perturbed correlation matrix, 4 eigen-values lie outside RMT bounds; 2 larger than λ_+ and 2 smaller than λ_-. The largest eigenvalue is 9.17. Whereas, in the case of a perturbed correlation matrix, we find 6 eigen-values deviating from RMT limits; 3 larger than λ_+ and 3 smaller than λ_-. The largest eigenvalue is 10.28.

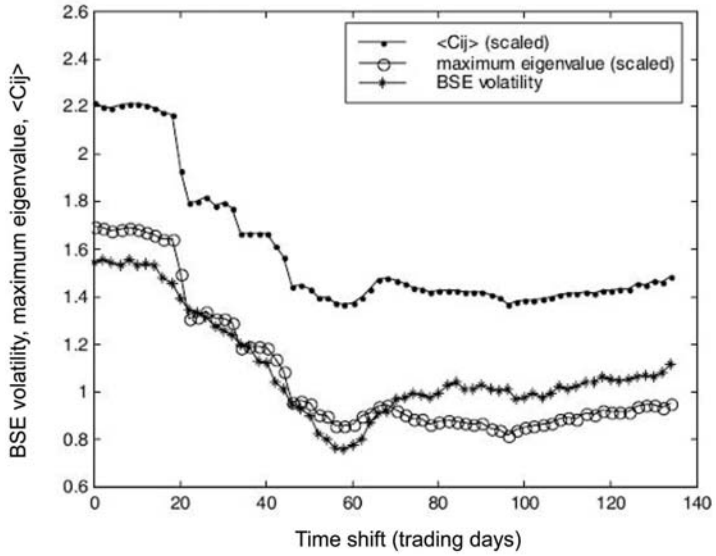

Fig. 5. Variation of largest eigenvalue and $\langle|C|\rangle$, with the time shift, j. Time shift j increases in steps of 2 days each time to span a total time of 280–500 days (see Fig. 2). Volatility has been scaled (as $\sum|G|$) for convenience.

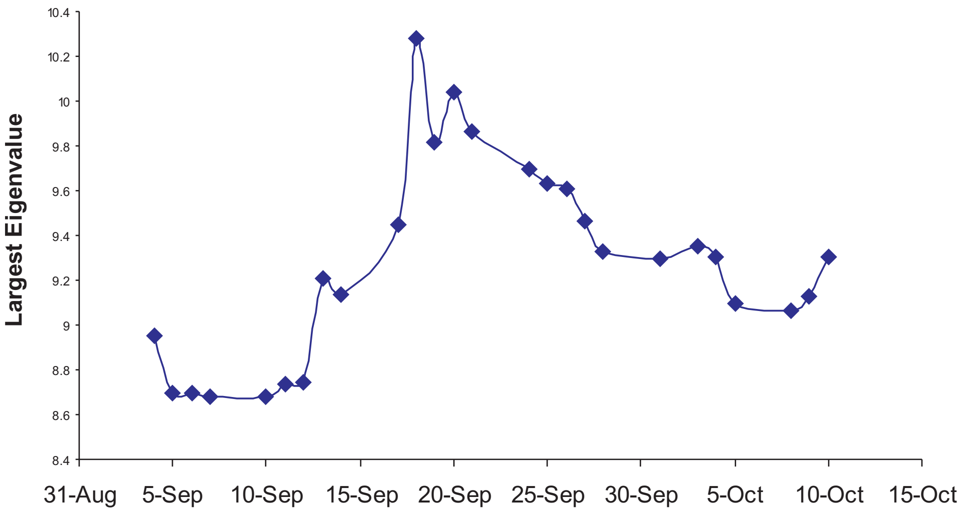

Fig. 6. Variation of largest eigenvalue with the time shift, j. Time shift j increases in steps of 1 day each time to span a total time of 333–444 days, in order to capture the impact of the 9/11 shock (see Fig. 2). The horizontal axis represents the last calendar day of all the time periods.

Of course the difference observed in the magnitude of the maximum eigenvalue in the two cases is consistent with intuition because large changes in correlations occur during market shocks. What may seem surprising is that an additional eigenvalue emerges from the RMT noise band in the second case. However the emergence of this eigenvalue is accompanied by a concentration of eigenvalues around the peak of the distribution, suggesting more systematic synchronization during this time. This observation is in agreement with the results for other markets [27].

3.3.2 Trend of largest eigenvalue

Since the largest eigenvalue represents collective information about the correlations between stocks, we expect its trend to be dependent on the market conditions. To see this we set up C using daily returns of $N = 70$ stocks for fixed time periods of length $T = 85$ days, progressing from quiescent (no-shock) periods to the ones hit by the shock. The trace of the correlation matrix is a constant throughout, $\mathrm{Tr}(C) = N$. The closer the maximum eigenvalue is to the trace, the more information it contains and the more correlated the prices would be. Variation of the largest eigenvalue is seen by advancing the time window each time by two days. Labelling the first and last day of all periods as t_f and t_l respectively, we set up C as

$$C(t_f, t_l) = C(280 + j, 280 + j + 85) \qquad (3.9)$$

where $j = 0, 2, 4, 6, \ldots, 134$ denotes the time shift.

The trend of the largest eigenvalue is shown in Figure 5. We observe a decrease in its magnitude for the time periods between 280–425 days after which it is more or less constant. It is found to be strongly correlated with the volatility of the BSE index (the simple correlation coefficient is found to be 0.94). In order to capture the impact of 9/11/2001 shock, we carry out a similar exercise, taking $j = 0, 1, 2, 3, \ldots, 26$. The aftermath of the event can be seen in Figure 6 by the sudden, impulsive rise in the maximum eigenvalue around September 13th, 18th, 2001. The impact was localized in time. Other deviating eigenvalues are also of significance in this context and this analysis may be applied to study their patterns.

3.4 The eigenvector corresponding to the largest eigenvalue and the correlation index

The general conclusion of some of the earlier work [10–13,20] is that the eigenvectors of C corresponding to eigenvalues deviating from RMT predictions bring out the collective response of the market to perturbations. In this section we see how the collective motion of all the stocks can be interpreted in terms of volatile versus non-volatile market conditions. In the following two subsections (3.4.1 and 3.4.2), we see that the degree of such synchronization is indicated by the eigenvector corresponding to the largest eigenvalue (the last eigenvector) through the evolution of its structure and components. Finally in (3.4.3), we try to quantify volatility in terms of the eigenvector corresponding to the largest eigenvalue to yield a strong indicator of correlation, the Correlation Index.

3.4.1 Distribution of eigenvector components

We study and compare the distributions of the components of the eigenvectors corresponding to largest eigenvalues for three time periods characterized by different volatilities (i) T_a: 280–365 days, volatility = 1.6, (ii) T_b: 340–425 days, volatility = 0.82 and (iii) T_c: 380–465 days, volatility = 0.99.

Figure 7 shows that the distributions of components of U^{70}, the eigenvector corresponding to the largest eigenvalue, are smaller and broader in the less volatile regimes (T_b and T_c) than in the more volatile regime T_a. Although the maximum contribution is found to be higher in distributions for T_b and T_c, the number of significant contributions (identified as components differing significantly from zero) is far lower than that in the period T_a; in fact it is half as much. (This is dealt with in the next subsection.) In addition, all the components of U^{70} in the period T_a have a positive sign, which confines the distribution to one side. This finding has been interpreted previously [10] to imply that there is a common component of the significant contributions of U^{70} that affects all of them with a similar bias. This means that the market forces at the time drive all the stocks in the same direction. For instance, in the event of a newsbreak, all the stocks would correlate and hence the collective contribution to the market would be high. However the periods T_b, T_c and all those following in succession (whose distributions have not been shown here) are relatively quiescent. The components of U^{70} in these time periods are found to be both positive and negative. It shows that the market conditions are not strong enough to drive the stocks in the ensemble together. This suggests an interesting link between the strength of the common influence and volatility. A collective or ensemble-like behavior seems to be more pertinent to volatile situations rather than non-volatile ones.

3.4.2 Inverse Participation Ratio

We analyze the evolution of the structure of the last eigenstate, U^{70} by evaluating the Inverse Participation Ratio.

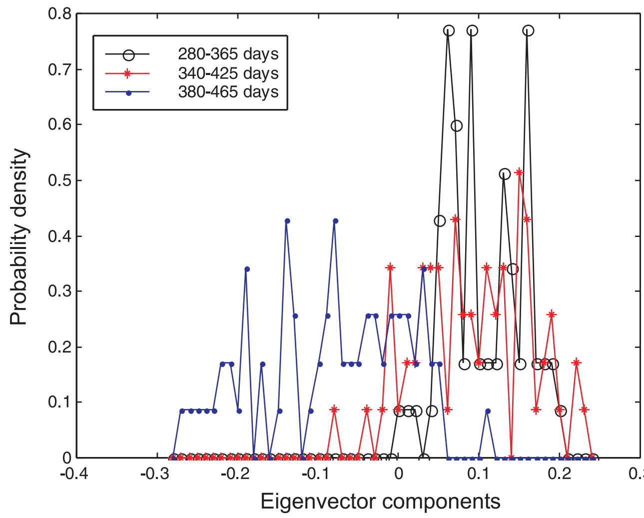

Fig. 7. Probability density of the eigenvector components for the largest eigenvalue for three periods (i) 280–365 days (ii) 340–425 and (iii) 380–465 days marked by volatilities 1.6, 0.82, 0.99 respectively. The plots are for C constructed from daily returns of 70 stocks for $T = 85$ days.

The IPR quantifies the contribution of different components of eigenvector to the magnitude of an eigenvector. We define ν_{ik}, $i = 1, 2, \ldots, N$ to be the components of eigenvector U^k. The IPR is given by

$$I_k = \sum_{i=1}^{N} \nu_{ik}^4. \tag{3.10}$$

Since IPR is the reciprocal of the number of eigenvector components that contribute significantly, if all components contribute identically, $\nu_{ik} = 1/\sqrt{N}$ then $I_k = 1/N$. As before we set up a correlation matrix C with $N = 70$ stocks for $T = 85$ days, each time shifting the time window forward in steps of 2 i.e. $j = 0, 2, 4, \ldots, 134$ spanning a period of 280–500 days as before. The pattern of IPR (Fig. 8) indicates that the number of significant participants in U^{70} decreases as we advance to less volatile periods. The IPR is closest to 0.0143 (=1/70), the value we would expect when all components contribute equally, in the most volatile periods of the time span. The values of IPR deviate more and more from 0.0143 as we move to the less volatile periods. In fact the correlation between IPR and volatility was found to be equal to −0.63.

3.4.3 Correlation Index

Another interesting feature brought out in the analysis of eigenvectors is the large-scale correlated movements associated with the eigenvector corresponding to the largest eigenvalue. The average magnitude of correlations of prices of every stock m with all stocks $n = 1, 2, \ldots, N$ is $\langle |C| \rangle_m = \frac{1}{(N-1)} \sum_{k=1}^{N} |C_{mk}|$, when $m \neq k$. Variation of $\langle |C| \rangle_m$ for $m = 1, 2, \ldots, N$ with the corresponding components of U^{70} shows (Fig. 9) a strong linear

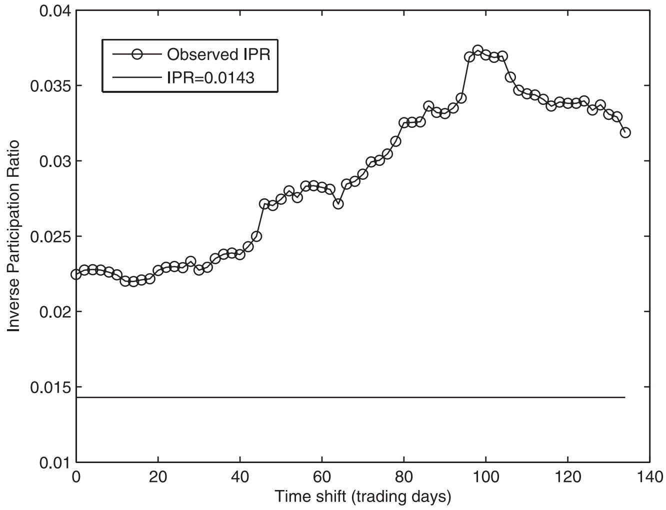

Fig. 8. Inverse participation ratio (IPR) for the eigenvector U^{70} as a function of time. Results have been obtained from correlation matrix C constructed from daily returns of 70 stocks for 68 time windows of 85 days each, progressed each time by 2 days spanning a time of 280–500 days. The solid line marks the value 0.0143 of IPR when all components contribute equally.

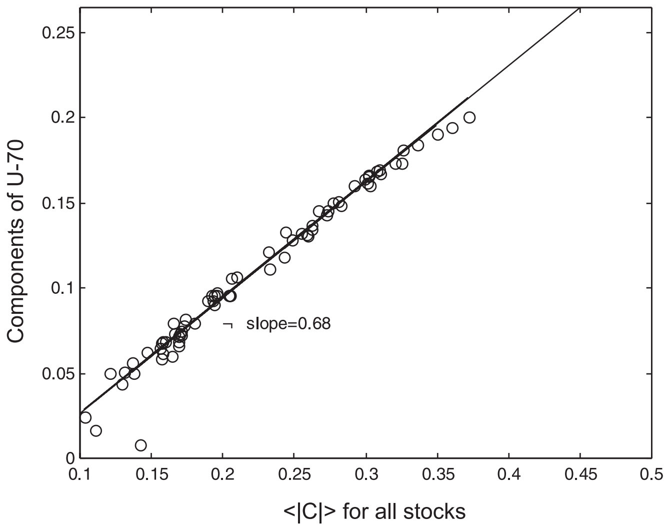

Fig. 9. Plot of the components of the eigenvector U^{70} corresponding to the largest eigenvalue with the extent to which every individual stock is correlated in the market, denoted by $\langle |C| \rangle_m$. In this case, the correlation matrix C was constructed using daily returns of 70 stocks for the period 280–365 days. The line obtained by least square fitting has a slope $= 0.68 \pm 0.01$.

positive relationship between the two at all times. However its variation with U^2, an eigenvector lying within the RMT range, shows almost zero dependence (Fig. 10). In this final sub-section we make use of this dependence to set up a Correlation Index, which is strongly related to the correlation of the BSE index.

We define a projection vector S with elements $S_m = \langle |C| \rangle_m$ where $m = 1, 2, \ldots, 70$, as calculated before. We

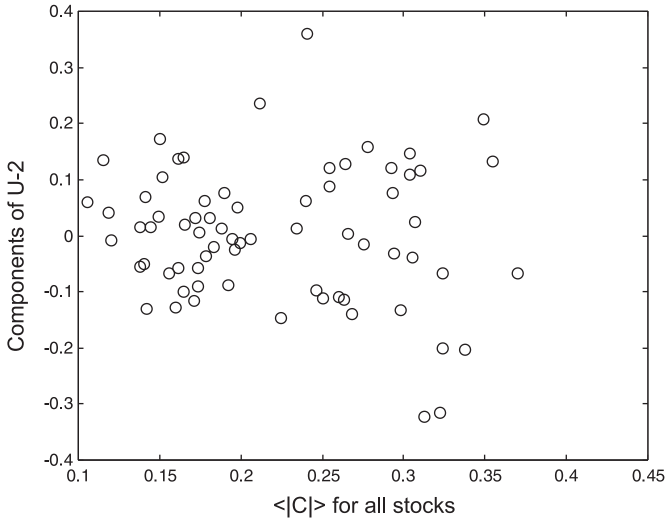

Fig. 10. Plot of the components of eigenvector U^2 associated with an eigenvalue from the bulk of RMT, λ_2. The variation shows no significant dependence between the two. The picture is quite the same for successive time periods considered.

obtain a quantity $X_m(t)$ by multiplying each element S_m by the square of the corresponding component of U^{70} (in analogy with the probability of the component of the eigenvector in quantum mechanics) for each time window t,

$$X_m(t) = (U_m^{70})^2 S_m, \qquad m = 1, 2, \ldots, 70. \tag{3.11}$$

The idea is to weight the average correlation possessed by every stock m in the market according to the contribution of the corresponding component to the last eigenvector U^{70}, thereby neglecting the contribution of less significant participants (the ones negligible in magnitude) in U^{70}. The quantity X in some sense represents the effective magnitude of correlations of stocks. Inclusion of other deviating components could also be useful in explaining the variation.

The sum of the correlation magnitudes is obtained as

$$CI(t) = \sum_{m=1}^{70} X_m(t), \qquad \text{at time } t \tag{3.12}$$

and may be expected to reflect the correlation of the market at that time. We call it the Correlation Index (CI). We note from Figure 11 that the Correlation Index behaves remarkably similar to the volatility of BSE index as the time window is slid forward. A highly statistically significant coefficient of correlation of 0.95 is obtained and, a positive linear relationship between the two can be seen in the plot of CI and BSE index volatility set out in Figure 12. A coefficient of correlation of 0.98 is found between the largest eigenvalue and CI. From its definition the correlation index is a more direct measure of correlation between stocks than the largest eigenvalue. We thus find the relevance of the last eigen-vector in quantifying the volatility of the overall market. Similar procedures have

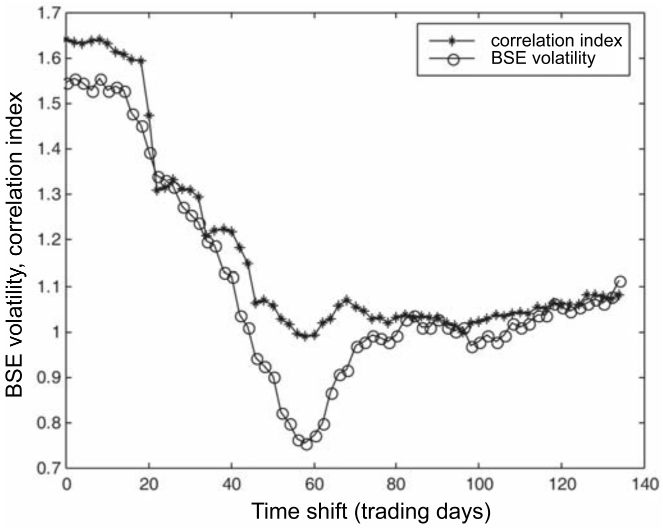

Fig. 11. Temporal evolution of the correlation index, CI (scaled by a factor of ∼6) and the scaled volatility (computed as $\sum |G(t)|$) of the BSE index. The results are obtained from the correlation matrix C constructed from daily returns of 70 stocks for 68 progressing time windows of 85 days each. The time shift increases in steps of 2 days each time and is represented by the horizontal axis.

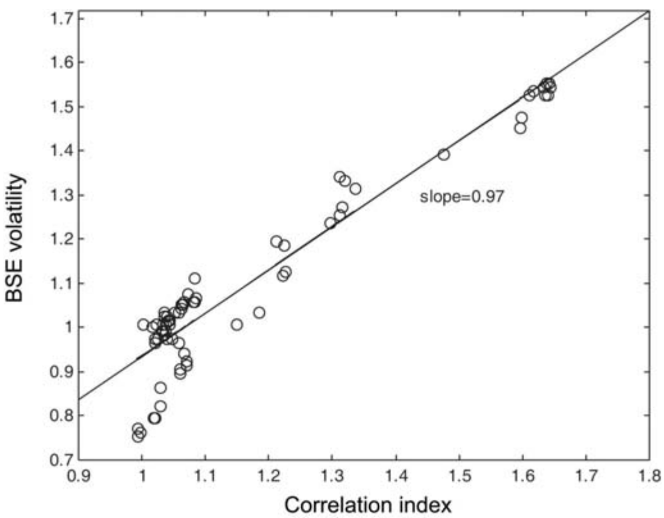

Fig. 12. The correlation index, CI with the volatility of BSE index approximates a linear fit with slope = 0.97 ± 0.04

been carried out in other studies [11] in different contexts to verify the relevance of this last eigenvector.

A similar study may also be done based on the Factor analysis. The one-factor model [28] postulates that X is linearly dependent on a factor F_1 as $X_m = l_{1m}F_1$ where l_{1m} is the factor loading of the m^{th} component on F_1. In this set up, $(U_m^{70})^2$ may be viewed as the loading on S_m, the participation factor. However further analysis is required to be able to establish the relevance of the one factor model to the procedure outlined in the present context. The correlation index, proposed here, is related to the regression coefficients β_m in the standard one-factor model [16], roughly as $CI(t) \approx \sum_m(\beta_m^3 + \ldots)$.

The conditions involved in the constructions of the correlation index and the BSE index are quite different. The correlation index uses only the prices, whereas the BSE index is a composite index defined by the set of variables – price and number of shares [23]. It is based on the idea that the total of the magnitudes of stock-price correlations (weighted according to their participation in U^{70}) is a good measure of the correlation between stocks. The fact that the correlation index indicates the level of volatility in the market at any time is remarkable.

4 Conclusion

In this paper we study the volatility of the Bombay Stock Exchange in India using the RMT approach. We find that the deviations from RMT bounds are more pronounced in volatile time periods as compared to the not so volatile ones for the Bombay Stock Exchange. The largest eigenvalue, which is in some sense an index of information contained in the entire market, is seen to be highly sensitive to the trends of market activity. A comparison of eigenvalue distributions for two analysis periods before and after the event on 9/11/2001, show that not only the number of eigen-values deviating from RMT bounds but also the magnitude of the maximum eigenvalue increase after the event. The simple correlation coefficient between $\lambda_{\max}$ and *BSE volatility* is 0.94. Analysis of the correlation matrix C as a function of time reveals a strong dependence between the average of magnitude of elements of C and volatility, indicating highly synchronous movements of stocks in highly fluctuating times or vice versa. A highly significant correlation coefficient of 0.94 is observed here as well.

The eigenvector associated with the largest eigenvalue, the last eigenvector of C has been enunciated in previous studies as the collective response of the whole market to certain bursts of activity. We have tried to explore its role in quantifying the fluctuations. It has been shown in Plerou et al. [10] that if all the components of the eigenvector have the same sign then there is some common component of the significant participants that affects all of them with similar bias. The probability density patterns of the components of U^{70}, found here, shows that while the distribution in period T_a is confined to the positive values of participation, the distributions in the other two periods (T_b and T_c), have spread to the negative side as well, indicating a gradual decrease of a common influence on the components as we move from a more volatile period T_a to less volatile periods T_b, T_c. Hence our finding here may suggest that ensemble-like behavior is more prominent in volatile situations than non-volatile ones. Further, the number of significant participants in T_b,T_c drops to almost half that in T_a, a finding better demonstrated by the time evolution of the inverse participation ratio for components of U^{70}. A strong anti-correlation between IPR and volatility ($= -0.63$) confirms the existence of a positive association between the number of significant participants in U^{70} with the volatility.

It is verified that the eigenvector U^{70} indicates the extent to which the stock movements are synchronized. We find a positive, linear relationship between the extent to which all individual stocks correlate or anti-correlate in the market ($\langle|C|\rangle_m, m = 1, 2, \ldots, N$) and the corresponding elements of U^{70}. Finally we investigate how this may lead to a quantification of the correlation of the market by taking the product of $\langle|C|\rangle_m$ with squares of corresponding elements of U^{70}. The products for all components may be put together as a sum to obtain a Correlation Index, CI. It is quantified as the sum of correlations of individual stocks, each weighted according to its participation in U^{70}. Temporal evolution of CI and BSE index volatility, have identical trends and there exists a highly statistically significant correlation of 0.95 between the two. In addition we find a close positive linear relationship between the two. It is noted that the ensemble of 70 stocks employed in the construction of CI may have no bearing on the Sensex basket of 30 stocks, the latter being chosen according to certain technical specifications. Further, the Correlation Index is based on stock-price correlations and not the market value of (or number of shares of) these stocks. However what is interesting is that even though a priori conditions underlying the constructions of CI and BSE index are different, the overall patterns of CI and BSE volatility match very well. We may conclude that the last eigenvector of the cross correlation matrix can be set up usefully to obtain a statistically significant indicator for the correlation of the market at any time. This establishes that there exists a close relationship between two distinct properties of the market, correlation and volatility.

We thank Dheeraj Bharadwaj and Sachin Maheshwari for making available their useful resources to carry out this work. Stanislaw Drozdz is thanked for valuable comments. We also thank P. Kaur (Delhi School of Economics, Delhi) for the data. VK thanks Raghav Gaiha, Suresh Kulkarni and Vidyadhar Mudkavi for the stimulating discussions and useful clarifications. ND thanks Sanjay Jain, Fabrizio Lillo, Rosario Mantegna, Matteo Marsili and Seeteez Mohapatra for discussions. ND gratefully acknowledges the CSIR project research grant No-03(1019)/05/EMR-II for support and the Abdus Salam International Center for Theoretical Physics for the Associateships and Visitorships during the summers.

References

1. L. Bachelier, *Theorie de la speculation* [Ph.D. thesis in mathematics], Annales Scientifiques de l'École Normale Superieure III-17, 21-86 (1900)
2. S.N. Kulkarni, Sankhya-The Indian Journal of Statistics **40**, Series D (1978)
3. R.N. Mantegna, H.E. Stanley, *An Introduction to Econophysics* (Cambridge University Press, 2000)
4. Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.K. Peng, H.E. Stanley, Phys. Rev. E **60**, 1390 (1999)
5. C. Shalen, *International Stock Index Spread Opportunities in Volatile Markets*, The CBOT Dow Jones Industrial Average Futures and Future Options, 02613 (1999)
6. S.B. Lee, K.J. Kim, Rev. Financial Econom. **3**, 89 (1993)
7. G.A. Karolygi, R.M. Stulz, J. Finance **51**, 951 (1996)
8. L. Ramchand, R. Susmel, J. Empirical Finance **5**, 397 (1998)
9. R. Engle, V.K. Ng, J. Finance, **XLVIII**, No. 5 (1993)
10. V. Plerou, P. Gopikrishnan, B. Rosenow, L.A. Amaral, T. Guhr, H.E. Stanley, Phys. Rev. E **65**, 066126 (2002)
11. P. Gopikrishnan, B. Rosenow, V. Plerou, H.E. Stanley, Phys. Rev. E **64**, 035106 (2001)
12. D. Kim, H. Jeong, Phys. Rev. E **72**, 046133 (2005)
13. H.E. Stanley, L.A.N Amaral, X. Gabaix, P. Gopikrishnan, V. Plerou, Physica A 126 (2001)
14. L. Laloux, P. Cizeau, J.P. Bouchaud, M. Potters, Phys. Rev. Lett. **83**, 1467 (1999); Risk **12**, No. 3, 69 (1999); V. Plerou, P. Gopikrishnan, B. Rosenow, L.N. Amaral, H.E. Stanley, Phys. Rev. Lett. **83**, 1471 (1999); M. Potters, J.P. Bouchaud, L. Laloux, `arXiv: Physics/0507111`
15. A.M. Sengupta, P.P. Mitra, Phys. Rev. E **60**, 3389 (1999)
16. F. Lillo, R. Mantegna, Phys. Rev. E **62**, 6126 (2000); Eur. Phys. J. B **15**, 603 (2000); Eur, Phys. J.B **20**, 503 (2001); F. Lillo, G. Bonanno, R.N. Mantegna, *Empirical Science of Financial Fluctuations*, Econophysics on the Horizon, edited by H. Takayasu (Springer-Verlag Tokio, 2002); S. Micciche, G. Bonanno, F. Lillo, R. Mantegna, Physica A **314**, 756 (2002); F. Lillo, R.N. Mantegna, Phys. Rev. E **72**, 016219 (2005)
17. S. Drozdz, F. Grummer, A.Z. Gorski, F. Ruf, J. Speth, Physica A **287**, 440 (2000); S. Drozdz, F. Grummer, F. Ruf, J. Speth, Physica A **294**, 226 (2001); *Empirical Science of Financial Fluctuations*, edited by H. Takayasu (Springer-Verlag Tokio 2002), p. 41
18. Z. Burda, J. Jurkiewiez, M.A. Nowak, G. Papp, I. Zahed, e-print `cond-mat/0103108`; Z. Burda, A. Goerlich, A. Jarosz, J. Jurkiewicz, e-print `cond-mat/0305627`
19. D. Wilcox, T. Gebbie, `arXiv:cond-mat/0402389`; `arXiv:cond-mat/0404416`
20. C. Mounfield, P. Ormerod, *Market Correlation and Market Volatility in US Blue Chip Stocks*, Crowell Prize Submission, 2001
21. S.S.S. Kumar, *Forecasting volatility - Evidence from Indian Stock and Forex Markets*, IIMK/WPS/08/FIN/2006/06
22. S. Bianco, R. Reno, *Unexpected Volatility and intraday serial correlation*, `arXiv:physics/0610023`, v1 3 October, 2006
23. *Technical Analysis of Indian stock market BSE Sensex Index*, Trend trading Newsletter, Traders Edge India
24. *An Introduction to High Frequency Finance*, G. Ramazan, M. Dacorogna, U. Muller, R. Olsen, O. Pictet (Academic Press Inc., London Ltd)
25. M. Mehta, *Random Matrices* (Academic Press, 1991)
26. Standard formula's for Karl Pearson's correlation coefficient can be found at `http://en.wikipedia.org/wiki/Correlation`
27. R. Rak, S. Drozdz, J. Kwapien, `arXiv:physics/0603071`; R. Rak, S. Drozdz, J. Kwapien, P. Oswiecimka, `arXiv:physics/0606041`
28. R. Johnson, D.W. Wichern (*Applied Multivariate Statistical Analysis*) (Prentice Hall International Inc., London, 1982)