

# On Measuring Extreme Synchrony with Network Entropy of Bipartite Graphs



Aki-Hiro Sato

## 1 Introduction

Recently a comprehensive measurement of agent behavior can be conducted based on high-resolution data on socioeconomic systems due to the development of information and communication technology. In particular, we want to consider a problem that we need to deal with the collective behavior of millions of nodes contribute to the observable dynamical features of such a complex system (de Menezes and Barabási 2004).

To do so, it is necessary to develop an adequate model both which considers states of agent behavior from the comprehensive point of view and which is as simple as one can estimate its parameters from actual data under assumptions. Specifically in socioeconomic activities, if we regard the situation where people exchange things, money, and information with each other as a network, then it further seems to be fruitful to study such social or economic systems from a network point of view.

To make advances in this direction, we need to treat structure on the basis of information transmission among heterogeneous agents in a given socioeconomic system from a limited amount of available data without precise knowledge on communication networks.

Several networks relating to human activity can be described as bipartite graphs with two kinds of nodes as shown in Fig. 1. For example, it is known that financial markets (financial commodities and participants), blog systems (blogs and

---

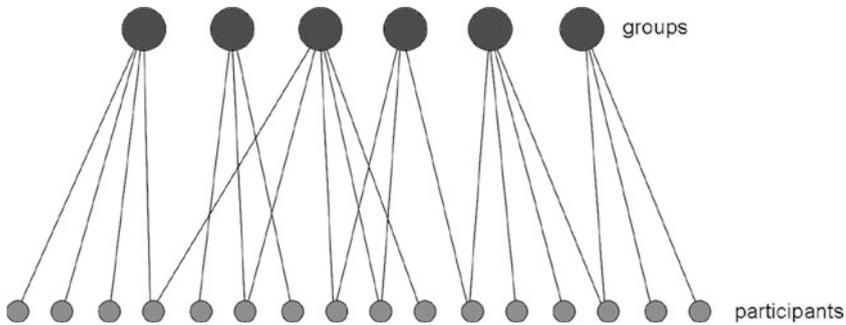
A.-H. Sato (✉)

Yokohama City University, Kanazawa-ku, Yokohama-shi, Kanagawa, Japan

Japan Science and Technology Agency PRESTO, Kawaguchi-shi, Saitama, Japan

Statistical Research and Training Institute, Ministry of Internal Affairs and Communications, Shinjuku-ku, Tokyo, Japan

e-mail: [ahsato@yokohama-cu.ac.jp](mailto:ahsato@yokohama-cu.ac.jp)



**Fig. 1** A conceptual illustration of social systems where  $M$  agents ( $B$  node) communicate with one another in  $K$  groups ( $A$  node)

bloggers), and economic systems (firms and goods/consumers) can be described as a directed bipartite graph (Lambiotte et al. 2007; Chmiel et al. 2007; Sato 2007; Sato and Hołyst 2008).

Such a bipartite network is recognized by observers when constituents are transmitted between two kinds of nodes. Links are recognized by observers when a constituent moves from one node to another node. Therefore bipartite network representation also seems to be one of our cognitive categories such as causality, time and space, intensity, quantity, and so on.

This chapter considers a model-based comparative measurement of collective behavior of groups based on their activities (Sato 2017). Specifically, a descriptive multivariate model of a financial market is proposed from a comprehensive point of view. Using comprehensive high-resolution data on the behavior of market participants, correlations of log returns and quotations (transactions) are analyzed on the basis of theoretical insights from the proposed model.

As a subject, we focus on financial markets which are attracting numerous researchers in various fields since they are of complex systems consisting of various types of heterogeneous agents. In the context of finance, the normal mixture hypothesis (or more generally “mixture of distribution hypothesis”) has been proposed as an alternative explanation for the description of return distribution of financial assets by several studies (Mandelbrot and Taylor 1967; Clark 1973; Tauchen and Pitts 1983; Richardson and Smith 1994). They have considered trading volume or the number of transactions (or quotations) as a proxy of the latent number of information arrivals. The proposed model may lead to better understanding of dynamics for return-volume relationship including volatility persistence (Lamoureux and Lastrapes 1990; Andersen 1996; Liesenfeld 1998; Watanabe 2000).

Cloud-based services can be used to provide various kinds of real-world applications. In fact, rich data on both mobility and activity of users have been accumulated by the cloud service providers. If secondary usage of the data collected in the cloud-based servers is permitted both legally and adequately and data analytics procedures for anonymized data can be developed, then we will be able to quantify risk and

chance of our society based on rich data on human mobility and activity with high resolution.

The chapter aims to propose a method to detect change points of temporally evolutionary bipartite network regarding anonymity of personal information. Moreover, in order to confirm an ability of the proposed method empirically, the proposed method is applied to data of trading activity collected in a cloud-based service of the foreign exchange market and quantified trading activity of the foreign exchange market from a comprehensive point of view.

The network structure of various kinds of physical and social systems has attracted considerable research attention. A many-body system can be described as a network, and the nature of growing networks has been examined well (Albert and Barabási 2002; Miura et al. 2012). Power-law properties can be found in the growing networks, which are called complex networks. These properties are related to the growth of elements and preferential attachment (Albert and Barabási 2002).

A network consists of several nodes and links that connect nodes. In the literature on the physics of socioeconomic systems (Carbone et al. 2007), nodes are assumed to represent agents, goods, and computers, while links express the relationships between nodes (Milaković et al. 2010; Lämmer et al. 2006). The network structure is perceived in many cases through the conveyance of information, knowledge, and energy, among others.

In statistical physics, the number of combinations of possible configurations under given energy constraints is related to “entropy.” Entropy is a measure that quantifies the states of thermodynamic systems. In physical systems, entropy naturally increases because of the thermal fluctuations on elements. Boltzmann proposed that entropy  $S$  is computed from the possible number of ensembles  $g$  by  $S = \log g$ . For a system that consists of two sub-systems whose respective entropies are  $S_1$  and  $S_2$ , the total entropy  $S$  is calculated as the sum of one of two subsystems  $S_1 + S_2$ . This case is attributed to the possible number of ensembles  $g_1 g_2$ . Entropy in statistical physics is also related to the degree of complexity of a physical system. If the entropy is low (high), then the physical configuration is rarely (often) realized. Energy injection or work in an observed system may be assumed to represent rare situations. Shannon entropy is also used to measure the uncertainty of time series (Carbone and Stanley 2007).

The concept of statistical–physical entropy was applied by Bianconi (2009) to measure network structure. She considered that the complexity of a network is related to the number of possible configurations of nodes and links under some constraints determined by observations. She calculated the network entropy of an arbitrary network in several cases of constraints.

Researchers have used a methodology to characterize network structure with information-theoretic entropy (Dehmer and Mowshowitz 2011; Wilhelm and Hollunder 2007; Rashevsky 1955; Trucco 1956; Mowshowitz 1968; Sato 2009). Several graph invariants such as the number of vertices, vertex degree sequence, and extended degree sequences have been used in the construction of entropy-based measures (Wilhelm and Hollunder 2007; Sato 2009).

## 2 An Entropy Measure on a Bipartite Network

The number of elements in socioeconomic systems is usually very large, and several restrictions or finiteness of observations can be found. Therefore, we need to develop a method to infer or quantify the affairs of the entire network structure from partial observations. Specifically, many affiliation relationships of socioeconomic systems can be expressed as a bipartite network. Describing the network structure of complex systems that consist of two types of nodes by using the bipartite network is important. A bipartite graph model also can be used as a general model for complex networks (Guillaume and Latapy 2006; Chmiel et al. 2007; Tumminello et al. 2011). Tumminello et al. proposed a statistical method to validate the heterogeneity of bipartite networks (Tumminello et al. 2011).

Suppose a symmetric binary two-mode network can be constructed by linking  $K$  groups ( $A$  node) and  $M$  participants ( $B$  node) if the participants belong to groups (Fig. 1). Assume that we can count the number of participants in each group within the time window  $[t\delta, (t+1)\delta]$  ( $t = 1, 2, 3, \dots$ ), which is defined as  $m_i(t)$  ( $i = 1, 2, \dots, K$ ).

Let us assume a bipartite graph consisting of  $A$  nodes and  $B$  nodes, of which the structure at time  $t$  is described as an adjacency matrix  $C_{ij}(t)$ . We also assume that  $A$  nodes are observable and  $B$  nodes are unobservable. That is, we only know the number of participants ( $B$  node) belonging to  $A$  nodes  $m_i(t)$ . We do not know the correct number of  $B$  nodes, but we assume that it is  $M$ . In this setting, how do we measure the complexity of the bipartite graph from  $m_i(t)$  at each observation time  $t$ ?

The network entropy is defined as a logarithmic form of the number of possible configurations of a network under a constraint (Bianconi 2009). We can introduce the network entropy at time  $t$  as a measure to quantify the complexity of a bipartite network structure. The number of possible configurations under the constraint  $m_i(t) = \sum_{j=1}^M C_{ij}(t)$  may be counted as

$$N(t) = \prod_{i=1}^K \binom{M}{m_i(t)} = \prod_{i=1}^K \frac{M!}{m_i(t)!(M - m_i(t))!}. \quad (1)$$

Then, the network entropy is defined as  $\Sigma(t) = \ln N(t)$ . Inserting Eq. (1) into this definition, we have

$$\Sigma(t) = K \sum_{n=1}^M \ln n - \sum_{i=1}^K \sum_{n=1}^{M-m_i(t)} \ln n - \sum_{i=1}^K \sum_{n=1}^{m_i(t)} \ln n. \quad (2)$$

Note that because  $0! = 1$ ,  $\sum_{n=1}^0 \ln n = 0$ . Obviously, if  $m_i(t) = M$  for any  $i$ , then  $\Sigma(t) = 0$ . If  $m_i(t) = 0$  for any  $i$ , then  $\Sigma(t) = 0$ . The lower number of combinations

gives a lower value of  $\Sigma(t)$ . To eliminate a difference in the number of links, we consider the network entropy per link defined as

$$\sigma(t) = \frac{\Sigma(t)}{\sum_{i=1}^K m_i(t)}. \quad (3)$$

This quantity shows the degree of complexity of the bipartite network structure. We may capture the temporal development of the network structure from the value of  $\sigma(t)$ . The network entropy per link  $\sigma(t)$  is also an approximation of the ratio of the entropy rate for  $m_i(t)$  to its mean so that

$$\sigma[m_1(t), \dots, m_K(t)] = \frac{\frac{1}{K} \Sigma[m_1(t), \dots, m_K(t)]}{\frac{1}{K} \sum_{i=1}^K m_i(t)} \approx \frac{\Sigma[\mathbf{m}(t)]}{\langle m(t) \rangle}, \quad (4)$$

where the entropy rate and the mean are, respectively, defined as

$$\Sigma[\mathbf{m}(t)] = \lim_{K \rightarrow \infty} \frac{1}{K} \Sigma[m_1(t), \dots, m_K(t)], \quad (5)$$

$$\langle m(t) \rangle = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K m_i(t). \quad (6)$$

The ratio of the entropy rate to the mean tells us the uncertainty of the mean from a different point of view from the coefficient of variation ( $C.V. = \text{standard deviation}/\text{mean}$ ).

To understand the fundamental properties of Eq. (3), we compute  $\sigma(t)$  in simple cases. Consider values of entropy for several cases at  $K = 100$  with different  $M$ . We assume that the total number of links is fixed at 100, which is the same as the number of  $A$  nodes, and we confirm the dependence of  $\sigma(t)$  on the degree of monopolization. We assign the same number of links at each  $A$  node. That is, we set

$$m_i(t) = \begin{cases} 100/k & (i = 1, \dots, k) \\ 0 & (i = k + 1, \dots, K) \end{cases}, \quad (7)$$

where  $k$  can be set as 1, 2, 4, 5, 10, 20, 50, or 100. In this case, we can calculate  $\sigma(t)$  as follows:

$$\begin{aligned} \sigma(t) &= \frac{\sum_{i=1}^k \ln \left( \frac{M}{100/k} \right)}{\sum_{i=1}^k 100/k} \\ &= \frac{k}{100} \left( \ln M! - \ln(100/k)! - \ln(M - 100/k)! \right). \end{aligned} \quad (8)$$

**Fig. 2** (a) Plots between  $\sigma(t)$  and degree of monopolization  $k$ . Each curve represents the relation between  $\sigma(t)$  and  $k$ . Filled squares numerical values for  $M = 1000$ , unfilled circles for  $M = 2000$ , filled circles for  $M = 3000$ , and unfilled triangle for  $M = 4000$ . (b) Plots between  $\sigma(t)$  and density of links  $p$ . Each curve represents the relation between  $\sigma(t)$  and  $k$ . Filled squares numerical values for  $M = 1000$ , unfilled circles for  $M = 2000$ , filled circles for  $M = 3000$ , and unfilled triangle for  $M = 4000$

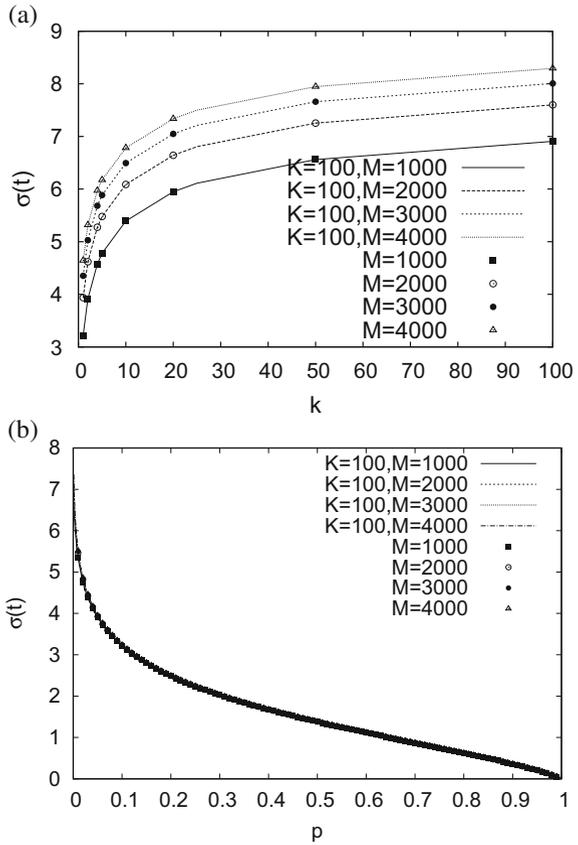


Figure 2a shows the relationship between  $\sigma(t)$  and the degree of monopolization at  $M = 1000, 2000, 3000,$  and  $4000$ . The network entropy per link  $\sigma(t)$  is small if a small population of nodes occupies a large number of links. The multiplication regime gives a large value of  $\sigma(t)$ . The value of  $\sigma(t)$  is a monotonically increasing function in terms of  $k$ . As  $M$  increases, the value of  $\sigma(t)$  increases. From this instance, we confirmed that  $\sigma(t)$  decreases with the degree of monopolization at  $A$  nodes.

Next, we confirm the dependency of  $\sigma(t)$  on the density of links. We assume that each element of an adjacency matrix  $C_{ij}(t)$  is given by an *i.i.d.* Bernoulli random variable with a successful probability of  $p$ . Then,  $m_i(t) = \sum_{j=1}^M C_{ij}(t)$  is sampled from an *i.i.d.* binomial distribution  $\text{Bin}(p, M)$ . In this case, one can approximate  $\sigma(t)$  as

$$\begin{aligned}
\sigma(t) &= \frac{\frac{1}{K} \sum_{i=1}^K \Sigma[m_i(t)]}{\frac{1}{K} \sum_{i=1}^K m_i(t)} \\
&\approx \frac{\langle \Sigma[m_1(t)] \rangle}{\langle m_1(t) \rangle} \\
&= \frac{1}{M} \sum_{k=1}^M \binom{M}{k} p^{k-1} (1-p)^{M-k} \ln \binom{M}{k}. \tag{9}
\end{aligned}$$

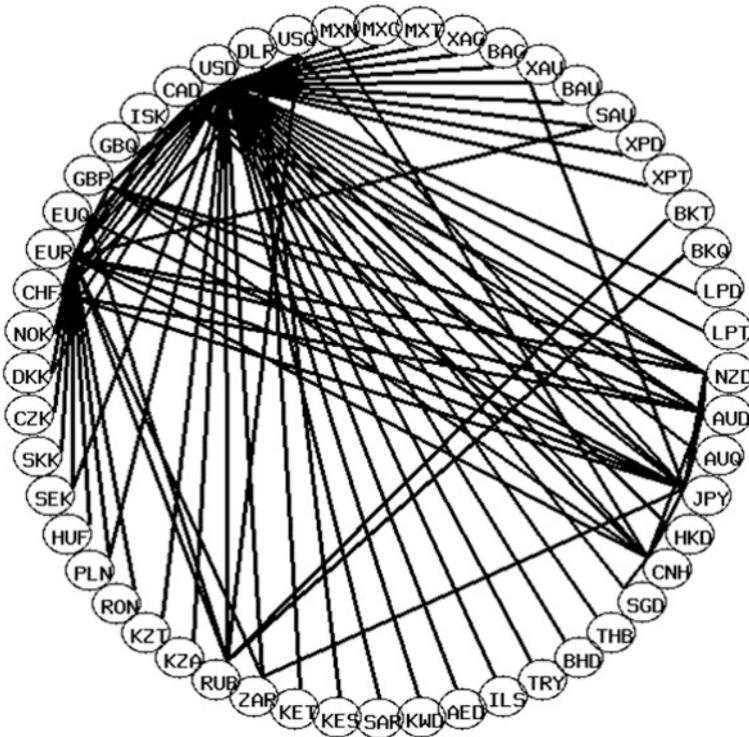
Figure 2b shows the plots of  $\sigma(t)$  versus  $p$  obtained from both Monte Carlo simulation with random links drawn from Bernoulli trials and Eq. (9). The number of links at each  $A$  node monotonically increases as  $p$  increases.  $\sigma(t)$  decreases as the density of links decreases. The dependence of the entropy per link on  $p$  is independent of  $M$ .

### 3 Empirical Analysis

The application of network analysis to financial time series has been advancing. Several researchers have investigated the network structure of financial markets (Bonanno et al. 2003; Gworek et al. 2010; Podobnik et al. 2009; Iori et al. 2008). Bonanno et al. examined the topological characterization of the correlation-based minimum spanning tree (MST) of real data (Bonanno et al. 2003). Gworek et al. analyzed the exchange rate returns of 38 currencies (including gold) and computed the characteristic path length and average weighted clustering coefficient of the MST topology of the graph extracted from the cross-correlations for several base currencies (Gworek et al. 2010). Podobnik et al. (2009) examined the cross-correlations between volume changes and price changes for the New York Stock Exchange, Standard and Poor's 500 index, and 28 worldwide financial indices. Iori et al. (2008) analyzed the network topology of the Italian segment of the European overnight money market and investigated the evolution of these banks' connectivity structure over the maintenance period. These studies collectively aimed to detect the susceptibility of network structures to macroeconomic situations.

Data collected from the ICAP EBS platform were used. The data period spanned May 28, 2007 to November 30, 2012 (ICAP 2013). The data included records for orders (BID/OFFER) and transactions for currencies and precious metals with a 1-s resolution. The total number of orders included in the data set is 520,973,843, and the total number of transactions is 58,679,809. The data set involved 94 currency pairs consisting of 39 currencies, 11 precious metals, and 2 basket currencies (AUD, NZD, USD, CHF, JPY, EUR, CZK, DKK, GBP, HUF, ISK, NOK, PLN, SEK, SKK, ZAR, CAD, HKD, MXC, MXN, MXT, RUB, SGD, XAG, XAU, XPD, XPT, TRY, THB, RON, BKT, ILS, SAU, DLR, KES, KET, AED, BHD, KWD, SAR, EUQ,

## 52 currencies, 94 currency pairs



**Fig. 3** A network representation of 94 currency pairs included in the data sets. Nodes show currencies and links currency pairs. If there is a link between two currencies, it is shown that the currency pair consisting of them is included in the data set

USQ, CNH, AUQ, GBQ, KZA, KZT, BAG, BAU, BKQ, LPD, and LPT).<sup>1</sup> Figure 3 shows a network representation of 94 currency pairs.

<sup>1</sup>AED, United Arab Emirates dirham; AUD, Australian dollar; AUQ, Australian dollar (small amount); BAG, gold (bank); BAU, silver (bank); BHD, Bahraini dinar; BKT, basket of USD/EUR; BKQ, basket of USD/EUR (small amount); CAD, Canadian dollar; CHF, Swiss franc; CNH, Chinese yuan; CZK, Czech koruna; DKK, Danish krone; EUR, EU euro; EUQ, EU euro (small amount); GBP, British sterling; GBQ, British sterling (small amount); HKD, Hong Kong dollar; HUF, Hungarian forint; ILS, Israeli new shekel; ISK, Iceland krona; JPY, Japanese yen; KES, Kenyan shilling; KET, Kenyan shilling (small amount); KZA, Kazakhstani tenge (small amount); KZT, Kazakhstani tenge; LPD, palladium (London); LPT, platinum (London); MXN, Mexican peso; MXQ, Mexican peso (small amount); MXT, Mexican peso (special deals); NOK, Norwegian krone; NZD, New Zealand dollar; PLN, Poland zloty; RON, Romanian leu; RUB, Russian ruble; SAR, Saudi Arabian riyal; SGD, Singapore dollar; SEK, Swedish krona; SKK, Slovak koruna; SAU, silver (small amount); TRY, Turkish lira; THB, Thai baht; USD/DLR, US dollar; USQ, US dollar (small amount); ZAR, South African rand; XAU, gold; SAU, gold (small amount); XAG, silver; XPD, palladium; and XPT, platinum.

### 3.1 The Total Number

The number of quotations and transactions in each currency pair was extracted from the raw data. Let  $m_{X,i}(t)$  ( $t = 0, \dots; i = 1, \dots, K$ ) be the number of quotations ( $X = P$ ) or transactions ( $X = D$ ) within every minute ( $\delta = 1$  [min]) for a currency pair  $i$  ( $K = 94$ ) at time  $t$ . Let  $c_X(t)$  be denoted as the total number of quotations ( $X = P$ ) and transactions ( $X = D$ ), which is defined as

$$c_X(t) = \sum_{i=1}^K m_{X,i}(t). \tag{10}$$

Let us consider the maximum value of  $c_X(t)$  in each week:

$$w_X(s) = \max_{t \in W(s)} \{c_X(t)\}, \tag{11}$$

where  $W(s)$  ( $s = 1, \dots, T$ ) represents a set of times included in the  $s$ -th week. A total of 288 weeks are included in the data set ( $T = 288$ ). Figure 4 shows the maximum values  $c_X(t)$  for the period from May 28, 2007 to November 30, 2012.

According to the extreme value theorem, the probability density for maximum values can be assumed to be a Gumbel density:

$$P(w_X; \mu_X, \rho_X) = \frac{1}{\rho_X} \exp\left(-\frac{w_X - \mu_X}{\rho_X} - e^{-\frac{w_X - \mu_X}{\rho_X}}\right), \tag{12}$$

where  $\mu_X$  and  $\rho_X$  are the location and scale parameters, respectively. Under the assumption of the Gumbel density, these parameters are estimated with the maximum likelihood procedure. The log-likelihood function for  $T$  observations  $w_X(s')$  ( $s' = 1, \dots, T$ ) under Eq. (12) is defined as

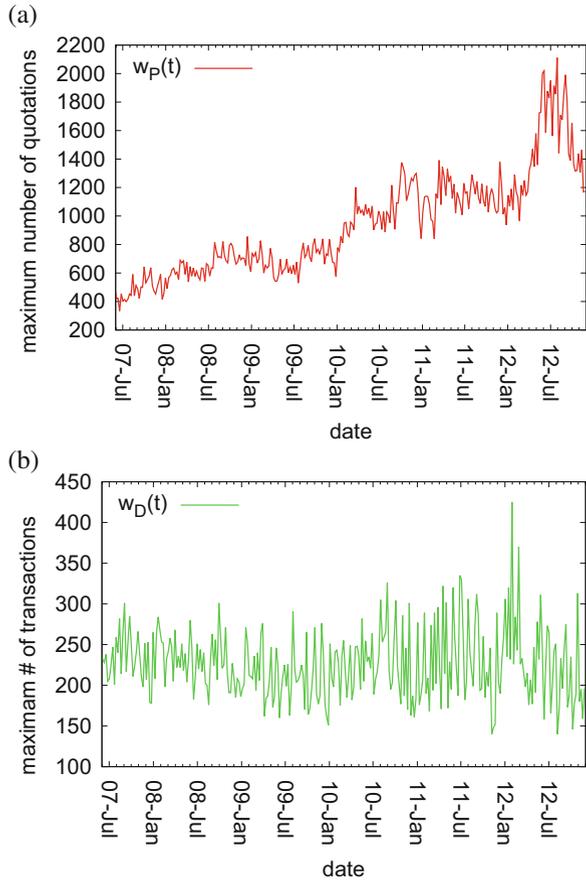
$$l(\mu_X, \rho_X) = \sum_{s'=1}^T \ln\left[\frac{1}{\rho_X} \exp\left(-\frac{w_X(s') - \mu_X}{\rho_X} - e^{-\frac{w_X(s') - \mu_X}{\rho_X}}\right)\right]. \tag{13}$$

The maximum likelihood estimators are obtained by maximizing the log-likelihood function. Partially differentiating  $l(\mu_X, \rho_X)$  in terms of  $\mu_X$  and  $\rho_X$  and setting them to zero, one has its maximum likelihood estimators as

$$e^{-\frac{\hat{\mu}_X}{\hat{\rho}_X}} = \frac{1}{T} \sum_{s'=1}^T e^{-\frac{w_X(s')}{\hat{\rho}_X}}, \tag{14}$$

$$\hat{\rho}_X = \frac{1}{T} \sum_{s'=1}^T w_X(s') - \frac{\sum_{s'=1}^T e^{-\frac{w_X(s')}{\hat{\rho}_X}} w_X(s')}{\sum_{s'=1}^T e^{-\frac{w_X(s')}{\hat{\rho}_X}}} \tag{15}$$

**Fig. 4** (a) The maximum values of the number of quotations within 1 min in every week. (b) The maximum values of the number of transactions within 1 min in every week



The parameters are estimated as  $\hat{\mu}_P = 772.179499$ ,  $\hat{\rho}_P = 281.741815$ ,  $\hat{\mu}_D = 206.454884$ , and  $\hat{\rho}_D = 35.984804$ .

The Kolmogorov–Smirnov (KS) test is conducted to determine the statistical significance of the estimated distributions. The KS test is a popular statistical method of assessing the difference between observations and its assumed distribution by  $p$ -value, which is a measure of probability where a difference between the two distributions happens by chance. Large  $p$ -values imply that the observations are sampled from the assumed distribution in the null hypothesis with high significance. Let  $w_X(s)$  ( $s = 1, \dots, T$ ) be  $T$  observations, and let  $K_T$  be a test statistic

$$K_T = \sup_{1 \leq s' \leq T} \sqrt{T} \left| F_T(w_X(s')) - F(w_X(s')) \right|, \tag{16}$$

where  $0 \leq F(v) \leq 1$  is an assumed cumulative distribution in a null hypothesis and  $F_T(v)$  an empirical one based on  $T$  observations such that  $F_T(v) = k/T$ , in which  $k$  represents the number of observations satisfying  $v_X(s) \leq v$  ( $s = 1, \dots, T$ ). The  $p$ -value is computed from the Kolmogorov–Smirnov distribution.

**Table 1** The  $p$ -values of statistical tests under a stationary assumption of the Gumbel distribution for the maximum values

$p$ -val (P)	KS-val (P)	$p$ -val (D)	KS-val (D)
0.041374	1.392521	0.586818	0.774087

The KS test is conducted under the assumption of the Gumbel distribution for the maximum value corresponding to Eq. (19):

$$F(w_X; \mu_X, \rho_X) = \exp\left[-\exp\left(-\frac{w_X - \mu_X}{\rho_X}\right)\right]. \tag{17}$$

The  $p$ -values of the KS test are shown in Table 1. The stationary Gumbel assumption cannot explain the maximum values for quotes with a 5% significance level in the KS test. The stationary Gumbel assumption may not be accepted in the case of the block maximum number of quotes. The dominant reason is the strong nonstationarity of the maximum number of quotes. During the last 5 years, the currencies and pairs quoted in the electronic brokerage market increased. The mean value of the total number constantly increased. In fact, the maximum number of quotations  $w_P(t)$  reached the maximum value on 30 July, 2012. The nonstationarity breaks the assumption of the extreme value theorem.

It is confirmed that the stationary Gumbel assumption can be accepted for the block maxima of transactions in each week using the KS test with a 5% significance level. The maximum number of transactions  $w_D(t)$  was reached on on January 30, 2012. This period seems to be related to the extreme synchrony.

### 3.2 Network Entropy Per Link

The proposed method based on statistical–physical entropy is applied to measure the states of the foreign exchange market. The relationship between a bipartite network structure and macroeconomic shocks or crises was investigated, and the occurrence probabilities of extreme synchrony were inferred. We compute a statistical–physical entropy per link from  $m_{X,i}(t) (X \in \{P, D\})$  with Eqs. (2) and (3), which are denoted as  $\sigma_X(t)$ ,  $\sigma_P(t)$  and  $\sigma_D(t)$ .

Since small values of  $\sigma_X(t)$  correspond to a concentration of links at a few nodes or a dense network structure, let us consider the minimum value of  $\sigma_X(t)$  every week:

$$v_X(s) = \min_{t \in W(s)} \{\sigma_X(t)\}, \tag{18}$$

where  $W(s)$  ( $s = 1, \dots, T$ ) represents a set of times included in the  $s$ -th week. A total of 288 weeks are included in the data set ( $T = 288$ ). According to the extreme value theorem, the probability density for minimum values can be assumed to be the Gumbel density:

$$P(v_X; \mu_X, \rho_X) = \frac{1}{\rho_X} \exp\left(\frac{v_X + \mu_X}{\rho_X} - e^{\frac{v_X + \mu_X}{\rho_X}}\right), \tag{19}$$

where  $\mu_X$  and  $\rho_X$  are the location and scale parameters, respectively. Under the assumption of the Gumbel density, these parameters are estimated with the maximum likelihood procedure. The log-likelihood function for  $T$  observations  $v_X(s')$  ( $s' = 1, \dots, T$ ) under Eq. (19) is defined as

$$l(\mu_X, \rho_X) = \sum_{s'=1}^T \ln\left[\frac{1}{\rho_X} \exp\left(\frac{v_X(s') + \mu_X}{\rho_X} - e^{\frac{v_X(s') + \mu_X}{\rho_X}}\right)\right]. \tag{20}$$

Partially differentiating  $l(\mu_X, \rho_X)$  in terms of  $\mu_X$  and  $\rho_X$  and setting them to zero yields its maximum likelihood estimators as

$$e^{-\frac{\hat{\mu}_X}{\hat{\rho}_X}} = \frac{1}{T} \sum_{s'=1}^T e^{\frac{v_X(s')}{\hat{\rho}_X}}, \tag{21}$$

$$\hat{\rho}_X = \frac{\sum_{s'=1}^T e^{\frac{v_X(s')}{\hat{\rho}_X}} v_X(s')}{\sum_{s'=1}^T e^{\frac{v_X(s')}{\hat{\rho}_X}}} - \frac{1}{T} \sum_{s'=1}^T v_X(s'). \tag{22}$$

The parameter estimates are computed as  $\hat{\mu}_P = -4.865382$ ,  $\hat{\rho}_P = 0.110136$ ,  $\hat{\mu}_D = -5.010175$ , and  $\hat{\rho}_D = 0.120809$  with Eqs. (21) and (22).

The KS test is conducted for the Gumbel distribution for the minimum values corresponding to Eq. (19):

$$F(v_X; \mu_X, \rho_X) = 1 - \exp\left[-\exp\left(\frac{v_X + \mu_X}{\rho_X}\right)\right]. \tag{23}$$

The  $p$ -value of the distribution is shown in Table 2. The stationary Gumbel assumption cannot explain the synchronizations observed in both quotes and transactions completely with a 5% significance level. The stationary Gumbel assumption is rejected because there is a stationary assumption to derive the extreme value distribution. If we can weaken this assumption, then the goodness of fit may be improved.

**Table 2** The  $p$ -values of statistical tests under a stationary Gumbel ssumption

$p$ -val (P)	KS-val (P)	$p$ -val (D)	KS-val (D)
0.001393	1.906528	0.019241	1.523791

### 4 Probability of Extreme Synchrony

The literature detecting structural breaks or change points in an economic time series (Goldfeld and Quandt 1973; Preis et al. 2011; Scalas 2007; Cheong et al. 2012) points out that nonstationary time series are constructed from locally stationary segments sampled from different distributions. Goldfeld and Quandt conducted a pioneering work on the separation of stationary segments (Goldfeld and Quandt 1973). Recently, a hierarchical segmentation procedure was also proposed by Choeng et al. under the Gaussian assumption (Cheong et al. 2012). We applied this concept to define the segments for  $v_X(s')$  ( $s' = 1, \dots, T$ ).

Let us consider the null model  $L_1$ , which assumes that all the observations  $v_X(s')$  ( $s' = 1, \dots, T$ ) are sampled from a stationary Gumbel density parameterized as  $\mu$  and  $\rho$ . An alternative model  $L_2(s)$  assumes that the left observations  $v_X(s')$  ( $s' = 1, \dots, s$ ) are sampled from a stationary Gumbel density parameterized as  $\mu_L$  and  $\rho_L$  and that the right observations  $v_X(s')$  ( $s' = s + 1, \dots, T$ ) are sampled from a stationary Gumbel density parameterized as  $\mu_R$  and  $\rho_R$ .

Denoting likelihood functions as

$$L_1(\mu, \rho) = \prod_{s'=1}^T P(v_X(s'); \mu, \rho), \tag{24}$$

$$L_2(s; \mu_L, \rho_L, \mu_R, \rho_R) = \prod_{s'=1}^s P(v_X(s'); \mu_L, \rho_L) \times \prod_{s'=s+1}^T P(v_X(s'); \mu_R, \rho_R), \tag{25}$$

the difference between the log-likelihood functions can be defined as

$$\Delta(s) = \log L_2(s) - \log L_1. \tag{26}$$

$\Delta(s)$  can be approximated as the Shannon entropy  $H[p] = - \int_{-\infty}^{\infty} dv \log p(v)p(v)$ :

$$\Delta(s) \approx TH[P(v_X; \mu, \rho)] - sH[P(v_X; \mu_L, \rho_L)] - (T - s)H[P(v_X; \mu_R, \rho_R)]. \tag{27}$$

Since the Shannon entropy of the Gumbel density expressed in Eq. (12) is calculated as

$$H[P(v_X; \mu_X, \rho_X)] = \ln \rho_X - \gamma_E + 1, \tag{28}$$

where  $\gamma_E$  represents Euler's constant, defined as

$$\gamma_E = \int_0^\infty \ln t e^{-t} dt, \tag{29}$$

we obtain

$$\Delta(s) \approx T \ln \rho - s \ln \rho_L - (T - s) \ln \rho_R. \tag{30}$$

In the context of model selection, several information criteria are proposed. The information criterion provides both goodness of fit of the model to the data and model complexity. For the sake of simplicity, we use the Akaike information criterion (AIC) to determine the adequate model. The AIC for a model with the number of parameters  $K$  and the maximum likelihood of  $L$  is defined as

$$AIC = -2 \ln L + 2K. \tag{31}$$

We can compute the difference in AIC between model  $L_2$  and model  $L_1(s)$  as

$$\begin{aligned} \Delta_{AIC}(s) &= \text{AIC of } L_2(s) - \text{AIC of } L_1 \\ &\approx -2(T \ln \hat{\rho} - s \ln \hat{\rho}_L - (T - s) \ln \hat{\rho}_R) + 4, \\ &= -2\Delta(s) + 4 \end{aligned} \tag{32}$$

since the number of parameters of  $L_1$  is 2, that of  $L_2(s)$  is 4, and the maximum likelihood is obtained by using their maximum likelihood estimators calculated from

$$\hat{\rho} = \frac{\sum_{s'=1}^T e^{\frac{v_X(s')}{\hat{\rho}_X}} v_X(s')}{\sum_{s'=1}^T e^{\frac{v_X(s')}{\hat{\rho}_X}}} - \frac{1}{T} \sum_{s'=1}^T v_X(s') \tag{33}$$

$$\hat{\rho}_L = \frac{\sum_{s'=1}^s e^{\frac{v_X(s')}{\hat{\rho}_L}} v_X(s')}{\sum_{s'=1}^s e^{\frac{v_X(s')}{\hat{\rho}_L}}} - \frac{1}{s} \sum_{s'=1}^s v_X(s') \tag{34}$$

$$\hat{\rho}_R = \frac{\sum_{s'=s+1}^T e^{\frac{v_X(s')}{\hat{\rho}_R}} v_X(s')}{\sum_{s'=s+1}^T e^{\frac{v_X(s')}{\hat{\rho}_R}}} - \frac{1}{T - s} \sum_{s'=s+1}^T v_X(s') \tag{35}$$

Therefore,  $P(v_X; \mu_L, \rho_L)$  is maximally different from  $P(v_X; \mu_R, \rho_R)$  when  $\Delta(s)$  assumes a maximal value. This spectrum has a maximum at some time  $s^*$ , which is denoted as

$$\Delta_{AIC}^* = \Delta_{AIC}(s^*) = \max_s \Delta_{AIC}(s). \tag{36}$$

The segmentation can be used recursively to separate the time series into further smaller segments. We do this iteratively until all segment boundaries have converged onto their optimal segment, defined by a stopping (termination) condition.

Several termination conditions were discussed in previous studies (Cheong et al. 2012). Assuming that  $\Delta_0 > 0$ , we terminate the iteration if  $\Delta_{AIC}^*$  is less than a typical conservative threshold of  $\Delta_0 = 10$ , while the procedure is recursively conducted if  $\Delta_{AIC}^*$  is larger than  $\Delta_0$ . We checked the robustness of this segmentation procedure for  $\Delta_0$ .  $\Delta_0$  gives a statistical significance level of termination. The value of  $\Delta_0$  is related to statistical significance. According to Wilks theorem,  $-2\Delta(s)$  follows a chi-squared distribution with a degree of freedom  $r$ , where  $r$  is given by the difference between the number of parameters assumed in the null hypothesis and one in the alternative hypothesis. In this case,  $r = 2$ . Hence, the cumulative distribution function of  $\Delta_{AIC}^*$  may follow

$$\Pr[\Delta_{AIC}^* > x] = 1 - \gamma\left(1, \frac{x - 4}{2}\right), \tag{37}$$

where  $\gamma(x, a)$  is the regularized incomplete gamma function defined as

$$\gamma(a, x) = \frac{\int_0^x t^{a-1} e^{-t} dt}{\int_0^\infty t^{a-1} e^{-t} dt}. \tag{38}$$

Therefore, setting the threshold  $\Delta_0 = 10$  implies that the segmentation procedure is tuned as a 4.928% statistical significance level.

Let the number of segments be  $L_X$ , the parameter estimates be  $\{\mu_{X,j}, \rho_{X,j}\}$  at the  $j$ -th segment, and the length of the  $j$ -th segment be  $\tau_{X,j}$ , where  $\sum_{j=1}^{R_X} \tau_{X,j} = T$ . The cumulative probability distribution for  $v_X(s)$  ( $s = 1, \dots, T$ ) may be assumed to be a finite mixture of Gumbel distributions:

$$\begin{aligned} \Pr(V_X \leq v_X) &= \int_{-\infty}^{v_X} \sum_{j=1}^{R_X} \frac{\tau_{X,j}}{T} P(v'_X; \mu_{X,j}, \rho_{X,j}) dv'_X \\ &= \sum_{j=1}^{R_X} \frac{\tau_{X,j}}{T} \left\{ 1 - \exp\left[-e^{\frac{v_X + \mu_{X,j}}{\rho_{X,j}}}\right] \right\}, \end{aligned} \tag{39}$$

Tables 3 and 4 show parameter estimates of  $v_P(s)$  and  $v_D(s)$  using the recursive segmentation procedure. Figure 5 shows the temporal development of  $v_P(s)$  and  $v_D(s)$  from May 28, 2007 to November 30, 2012.  $R_P = 6$  and  $R_D = 6$  are obtained from  $v_X(s)$  using the proposed segmentation procedure. During the observation period, the global financial system suffered from the following significant macroeconomic shocks and crises: (I) the BNP Paribas shock (August 2007), (II) the Bear Stearns shock (February 2008), (III) the Lehman shock (September 2008 to March 2009), (IV) the European sovereign debt crisis (April to May 2010), (V) the East

**Table 3** Parameter estimates obtained from the weekly minimum values of network entropy for quotations with the recursive segmentation procedure

$j$	Start date	End date	$\tau/T$	$\hat{\mu}_{X,j}$	$\hat{\rho}_{X,j}$
1	May 28, 2007	Oct. 15, 2007	0.072917	-4.965486	0.078571
2	Oct. 22, 2007	Aug. 17, 2009	0.333333	-4.829061	0.080558
3	Aug. 24, 2009	Oct. 21, 2011	0.274306	-4.880156	0.099532
4	Feb. 28, 2011	Jul. 30, 2012	0.260417	-4.798364	0.058265
5	Aug. 6, 2012	Sep. 10, 2012	0.020833	-4.856415	0.121566
6	Sep. 17, 2012	Nov. 26, 2012	0.038194	-5.031930	0.071871

**Table 4** Parameter estimates obtained from the weekly minimum values of network entropy for transactions with the recursive segmentation procedure

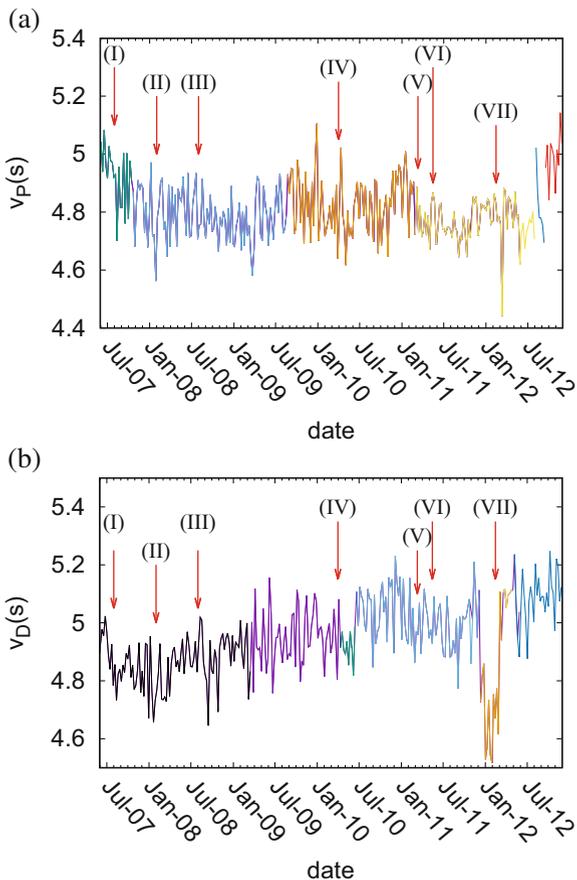
$j$	Start date	End date	$\tau/T$	$\hat{\mu}_{X,j}$	$\hat{\rho}_{X,j}$
1	May 28, 2007	Mar. 16, 2009	0.329861	-4.903660	0.072684
2	Mar. 23, 2009	Jun. 14, 2010	0.225694	-4.990569	0.088420
3	Jun. 21, 2010	Dec. 5, 2011	0.267361	-5.065305	0.087721
4	Dec. 12, 2011	Mar. 12, 2012	0.048611	-4.826384	0.174328
5	Mar. 19, 2012	Apr. 30, 2012	0.024306	-5.108679	0.011079
6	May 7, 2012	Nov. 26, 2012	0.104167	-5.131156	0.079630

Japan tsunami (March 2011), (VI) the United States debt-ceiling crisis (May 2011), and (VII) the Bank of Japan's 10 trillion JPY gift on Valentine's Day (February 2012).

Before entering these global affairs, both  $v_P(s)$  and  $v_D(s)$  took large values. Note that, during the (I) Paribas shock, the (II) Bear Stearns and the (III) Lehman shock  $v_P(s)$  and  $v_D(s)$  took smaller values than they did during the previous term. This implies that a global shock may drive many participants and that these participants may trade the same currencies at the same time. The smallest values  $v_P(s)$  and  $v_D(t)$  correspond to the days of the (II) Bear Stearns shock, the (III) Lehman shock, and the (VI) Euro crisis. These days are generally related to the start or the end of macroeconomic shocks or crises. The period from December 2011 to March 2012 shows that the values of  $v_D(s)$  are smaller than they were during other periods. This result implies that, during period, singular patterns appeared in the transactions.

Figure 6 shows both the empirical and estimated cumulative distribution functions of  $v_P(s)$  and  $v_D(s)$ . The estimated cumulative distributions are drawn from Eq. (39) with parameter estimates. The KS test verifies this mixing assumption. The distribution estimated by the finite mixture of Gumbel distributions for quotes is well fitted, as shown in Table 5. From the  $p$ -values, the mixture of Gumbel distributions for quotations accepts the null hypothesis that  $v_P(s)$  is sampled from the mixing distribution with a 5% significance level. The mixture of Gumbel distributions for transactions also accepts the null hypothesis that  $v_D(s)$  is sampled from the mixing distribution with a 5% significance level. Extrapolation of cumulative distribution function also provides a guideline of the future probability of extreme events. The finite mixture Gumbel distributions with parameter estimates may be used as an inference of probable extreme synchrony.

**Fig. 5** Temporal development of (a)  $v_P(s)$  and (b)  $v_D(s)$  from May 28, 2007 to November 30, 2012. Each color corresponds to a segment which is shown in Tables 3 and 4

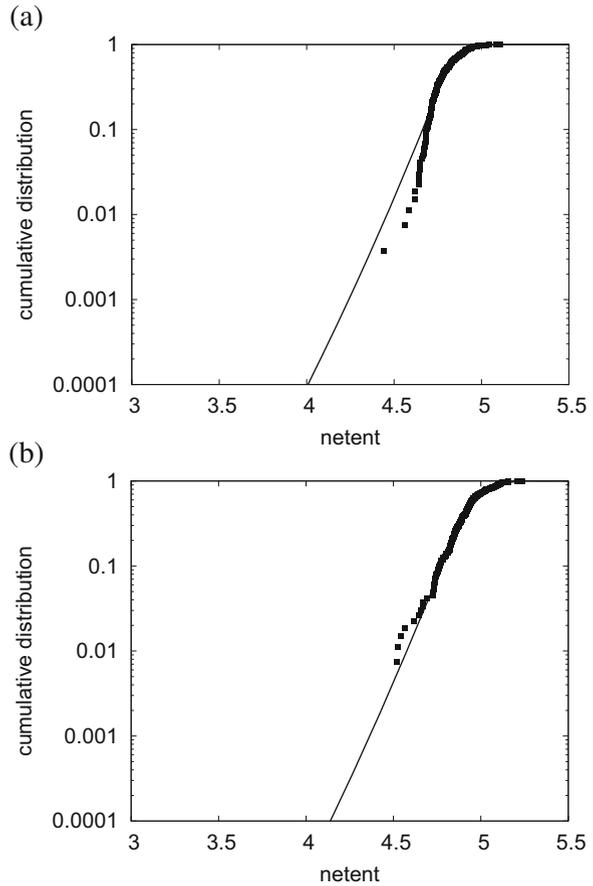


## 5 Conclusion

A method based on the concept of “entropy” in statistical physics was proposed to quantify states of a bipartite network under constraints. The statistical–physical network entropy of a bipartite network was derived under the constraints for the number of links at each group node. Both numerical and theoretical calculations for a binary bipartite graph with random links showed that the network entropy per link can capture both the density and the concentration of links in the bipartite network. The proposed method was applied to measure the structure of bipartite networks consisting of currency pairs and participants in the foreign exchange market.

An empirical investigation of the total number of quotes and transactions was conducted. The nonstationarity of the number of quotes and transactions strongly affected the extreme value distributions. The empirical investigation confirmed that the entropy per link decreased before and after the latest global shocks that have

**Fig. 6** Cumulative distribution functions for the minimum values of the entropy per link in each week (a)  $v_P(s)$  and (b)  $v_D(s)$ . Filled squares represent the empirical distribution of  $v_P(s)$ , and unfilled circles represent the empirical distribution of  $v_D(s)$ . A solid curve represents the estimated distribution of  $v_P(s)$ , and a dashed curve represents the estimated distribution of  $v_D(s)$



**Table 5** The  $p$ -values of statistical tests under the assumption of a finite mixture of Gumbel distributions

$p$ -val (P)	KS-val (P)	$p$ -val (D)	KS-val (D)
0.183793	1.092317	0.829013	0.625372

influenced the world economy. A method was proposed to determine segments with recursive segmentation based on the Akaike information criterion between Gumbel distributions with different parameters. Under the assumption of a finite mixture of Gumbel distributions, the estimated distributions were verified by the Kolmogorov–Smirnov test. The finite mixture of Gumbel distributions can estimate the occurrence probabilities of extreme synchrony of a nonstationary system extracted as a bipartite network. The extrapolation of the extreme synchrony can be done based on the estimated mixture of Gumbel distributions.

**Acknowledgements** This work is supported by Japan Science and Technology Agency (JST) PRESTO Grant Number JPMJPR1504, Japan. The work was also financially supported by the Grant-in-Aid for Young Scientists (B) by the Japan Society for the Promotion of Science (JSPS) KAKENHI (#23760074).

## References

- Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74:47–97
- Andersen TG (1996) Return volatility and trading volume: an information flow interpretation of stochastic volatility. *J Financ* 51:169–204
- Bianconi G (2009) Entropy of network ensembles. *Phys Rev E* 79:036114; Anand K, Bianconi G (2009) Entropy measures for networks: toward an information theory of complex topologies. *Phys Rev E* 80:045102
- Bonanno G, Caldarelli G, Lillo F, Mantegna RN (2003) Topology of correlation-based minimal spanning trees in real and model markets. *Phys Rev E* 68:046130
- Carbone A, Stanley HE (2007) Scaling properties and entropy of long-range correlated time series. *Physica A* 304:21–24
- Carbone A, Kaniadakis G, Scarfone AM (2007) *Eur Phys J B* 57:121
- Cheong SA, Fornia RP, Lee GHT, Kok JL, Yim WS, Xu DY, Zhang Y (2012) The Japanese economy in crises: a time series segmentation study. *Econ E-J* 2012-5. <http://www.economics-ejournal.org>
- Chmiel AM, Sienkiewicz J, Suchecki K, Hołyst JA (2007) Networks of companies and branches in Poland. *Physica A* 383:134–138
- Clark P (1973) A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* 41:135–155
- Dehmer M, Mowshowitz A (2011) A history of graph entropy measures. *Inf Sci* 181:57–78
- de Menezes MA, Barabási A-L, Fluctuations in network dynamics. *Phys Rev Lett* 92 (2004) 028701.
- Goldfeld SM, Quandt RE (1973) A Markov model for switching regressions. *J Econometrics* 1:3–15
- Guillaume J-L, Latapy M (2006) Bipartite graphs as models of complex networks. *Physica A* 371:795–813
- Gworek S, Kwapien J, Drożdż S (2010) Sign and amplitude representation of the forex networks. *Acta Phys Pol A* 117:681–687
- ICAP (2013). The data is purchased from ICAP EBS: <http://www.icap.com>
- Iori G, Masi GD, Precup OV, Gabbi G, Caldarelli G (2008) A network analysis of the Italian overnight money market. *J Econ Dyn Control* 32:259–278
- Lambiotte R, Ausloos M, Thelwall M (2007) Word statistics in Blogs and RSS feeds: towards empirical universal evidence. *J Informetrics* 1:277–286
- Lämmer S, Gehlsen B, Helbing D (2006) Scaling laws in the spatial structure of urban road networks. *Physica A* 363:89–95
- Lamoureux CG, Lastrapes WD (1990) Heteroskedasticity in stock return data: volume versus GARCH effects. *J Financ* 45:221–229
- Liesenfeld R (1998) Dynamic bivariate mixture models: modeling the behavior of prices and trading volume. *J Bus Econ Stat* 16:101–109
- Mandelbrot BB, Taylor H (1967) On the distribution of stock price differences. *Oper Res* 15:1057–1062
- Milaković M, Alfrano S, Lux T (2010) The small core of the German corporate board network. *Comput Math Organ Theory* 16:201–215

- Miura W, Takayasu H, Takayasu M (2012) Effect of coagulation of nodes in an evolving complex network. *Phys Rev Lett* 108:168701
- Mowshowitz A (1968) Entropy and the complexity of graphs: I. An index of the relative complexity of a graph. *Bull Math Biophys* 30:175–204
- Podobnik B, Horvatic D, Petersen AM, Stanley HE (2009) Cross-correlations between volume change and price change. *Proc Natl Acad Sci U S A* 106:22079–22084
- Preis T, Schneider JJ, Stanley HE (2011) Switching processes in financial markets. *Proc Natl Acad Sci U S A* 108:7674–7678
- Rashevsky N (1955) Life, information theory, and topology. *Bull Math Biophys* 17:229–235
- Richardson M, Smith T (1994) A direct test of the mixture of distributions hypothesis: measuring the daily flow of information. *J Financ Quant Anal* 29:101–116
- Sato A-H (2007) Frequency analysis of tick quotes on the foreign exchange market and agent-based modeling: a spectral distance approach. *Physica A* 382:258–270
- Sato A-H (2010) Comprehensive analysis of information transmission among agents: similarity and heterogeneity of collective behavior. In: Chen S-H et al (eds) *Agent-based in economic and social systems VI: post-proceedings of the AESCS international workshop 2009, agent-based social systems*, vol 8. Springer, Tokyo, pp 1–17
- Sato A-H (2017) Inference of extreme synchrony with an entropy measure on a bipartite network. In: 2017 IEEE 41st annual computer software and applications conference (COMPSAC), pp. 766–771
- Sato A-H, Holyst JA (2008) Characteristic periodicities of collective behavior at the foreign exchange market. *Eur Phys J B* 62:373–380
- Scalas E (2007) Mixtures of compound Poisson processes as models of tick-by-tick financial data. *Chaos, Solitons Fractals* 34:33–40
- Tauchen T, Pitts M (1983) The price variability-volume relationship on speculative markets. *Econometrica* 51:485–505
- Trucco E (1956) A note on the information content of graphs. *Bull Math Biophys* 18:129–135
- Tumminello M, Micciché S, Lillo F, Pillo J, Mantegna RN (2011) Statistically validated networks in bipartite complex systems. *PLoS One* 6:e17994
- Watanabe T (2000) A nonlinear filtering approach to stochastic volatility models with an application to daily stock returns. *J Bus Econ Stat* 18:199–210
- Wilhelm T, Hollunder J (2007) Information theoretic description of networks. *Physica A* 385:385–396