

RESEARCH ARTICLE

Long-Range Correlations and Memory in the Dynamics of Internet Interdomain Routing

Maksim Kitsak^{1☯*}, Ahmed Elmokashfi^{2☯}, Shlomo Havlin³, Dmitri Krioukov^{1,4,5}

1 Department of Physics, Northeastern University, Boston, MA, United States of America, **2** Simula Research Lab, Oslo, Norway, **3** Department of Physics, Bar-Ilan University, Ramat Gan, Israel, **4** Department of Mathematics, Northeastern University, Boston, MA, United States of America, **5** Department of Electrical&Computer Engineering, Northeastern University, Boston, MA, United States of America

☯ These authors contributed equally to this work.

* m.kitsak@neu.edu



CrossMark
click for updates

OPEN ACCESS

Citation: Kitsak M, Elmokashfi A, Havlin S, Krioukov D (2015) Long-Range Correlations and Memory in the Dynamics of Internet Interdomain Routing. PLoS ONE 10(11): e0141481. doi:10.1371/journal.pone.0141481

Editor: Tobias Preis, University of Warwick, UNITED KINGDOM

Received: June 29, 2015

Accepted: October 8, 2015

Published: November 3, 2015

Copyright: © 2015 Kitsak et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data are available on Figshare (http://figshare.com/articles/Correlation_in_global_routing_dynamics/1549778). Technical questions regarding the data should be directed to Ahmed Elmokashfi (ahmed@simula.no).

Funding: This work was supported in part by NSF grants nos. CNS-1344289, CNS-1442999, CNS-0964236, CNS-1441828, and CNS-1039646; DARPA grant no. HR001112-10012; Norwegian Research Council grant no. 240850/O70; and Cisco Systems.

Competing Interests: This work was supported in part by NSF grants nos. CNS-1344289, CNS-

Abstract

Data transfer is one of the main functions of the Internet. The Internet consists of a large number of interconnected subnetworks or domains, known as Autonomous Systems (ASes). Due to privacy and other reasons the information about what route to use to reach devices within other ASes is not readily available to any given AS. The Border Gateway Protocol (BGP) is responsible for discovering and distributing this reachability information to all ASes. Since the topology of the Internet is highly dynamic, all ASes constantly exchange and update this reachability information in small chunks, known as routing control packets or BGP updates. In the view of the quick growth of the Internet there are significant concerns with the scalability of the BGP updates and the efficiency of the BGP routing in general. Motivated by these issues we conduct a systematic time series analysis of BGP update rates. We find that BGP update time series are extremely volatile, exhibit long-term correlations and memory effects, similar to seismic time series, or temperature and stock market price fluctuations. The presented statistical characterization of BGP update dynamics could serve as a basis for validation of existing and developing better models of Internet interdomain routing.

Introduction

On large scale, the Internet is a global system of approximately 40,000 interlinked computer networks connecting billions of users and devices worldwide [1]. These networks are called Autonomous Systems (ASes). ASes vary in size and function: they can be (i) Internet Service and/or Transit Providers (AT&T), (ii) Content Providers (Google), (iii) Enterprises (Harvard University), and (iv) Non-profit organizations [2]. Devices inside ASes are identified via unique Internet Protocol (IP) addresses, which are 32- or 128-bit numerical labels that act both as identifiers and locators of devices. An IP address is divided into two sections, a network section and a host section. The network section, which is known as IP prefix, identifies a group of hosts, while the host section identifies a particular device. An AS can include a number of IP prefixes.

1442999, CNS-0964236, CNS-1441828, and CNS-1039646; DARPA grant no. HR001112-10012; Norwegian Research Council grant no. 240850/O70; and Cisco Systems. Financial support from Cisco Systems does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

Each AS is administrated by a single entity, but a single organization may own and operate several ASes. ASes connect to each other via contractual agreements that govern the flow of data between and through them. This interconnection of ASes shapes the AS-level topology of the Internet, which facilitates connectivity between any pair of ASes and thus any pair of devices connected to the Internet (Fig 1a).

The information about how to reach devices within other ASes is not readily available to them. The exchange of this information is handled by specialized networked computers called routers. Performing routing requires signaling reachability information, comparing different possibilities, and maintaining a state that describes how to reach different IP prefixes. The Border Gateway Protocol (BGP) [3] is the globally deployed routing protocol that accomplishes this task. The BGP protocol can be summarized as follows. ASes advertise their IP prefixes to their neighbor ASes through BGP update messages. At each AS incoming BGP updates are processed by the BGP router and the resulting reachability information is then stored in routing tables.

The Internet is a dynamic system where participating networks and links between them do often experience configuration changes, failures, and restorations. BGP protocol reacts to changes in the Internet connectivity incrementally: BGP routers send update messages to their neighbor BGP routers. BGP update messages do not carry the information on the whole Internet connectivity state. Instead, they carry only the information concerning the affected IP

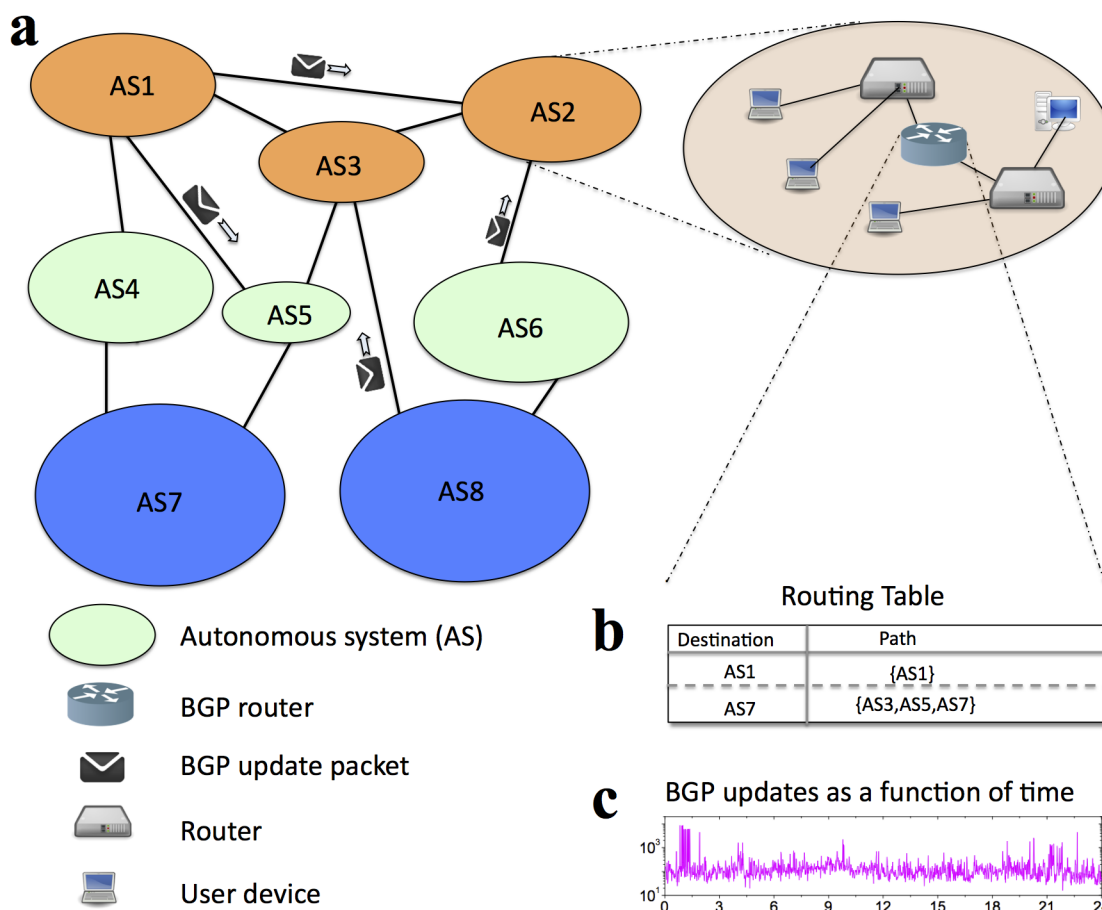


Fig 1. The Internet and BGP routing. **a**, On large scale the Internet is the product of interconnectivity among a large number of ASes (shown with ovals). In order to perform data transfer, ASes need to exchange the reachability information through BGP update messages. **b**, **c** BGP updates are processed by the BGP routers. **b**, The reachability information is stored in the routing table. **c**, Typical dynamics in the number of updates received by the BGP router.

doi:10.1371/journal.pone.0141481.g001

prefixes. Hence, to keep a consistent view of the network and, consequently, to be able to communicate with other networks, a BGP router must process incoming BGP updates in a timely manner and update its routing table accordingly (Fig 1b and 1c).

Current version of the BGP routing protocol was introduced in 1994. Since then, the deployment of the BGP routing protocol has sustained tremendous growth and it is arguably one of the main technological reasons behind the success of the Internet.

Nevertheless, there are two major concerns related to the fast rate of the Internet growth. On one hand, Internet growth implies the growth in the number of destinations for the BGP routing and, thus, results in the growth of routing table sizes. On the other hand, the growth of the Internet also leads to the growth in the number of BGP updates needed to maintain BGP routing [4]. Both factors are important, especially for routers at the core of the Internet. The growing size of routing tables requires increasingly larger and faster memory. At the same time, growing routing table sizes do not necessarily slow down data forwarding as long as address lookups are performed using high speed memories and constant-time matching algorithms [5]. Increasingly large amounts of BGP updates, on the other hand, is a more serious concern because processing BGP updates can be computationally heavy (updating routing state, generating more updates, checking import/export filters), and can trigger wide-scale instabilities [6].

Recent studies of BGP scalability range from measurements assessing the extent of the concern [7, 8] to studies suggesting radically new routing architectures [9, 10]. Elmokashfi et al. [7] analyzed the dynamics of BGP updates in four networks at the backbone of the Internet over a period of seven years and eight months. They have shown that on average the level of BGP updates is increasing, but not at an alarming rate: it was shown to grow at rates similar to the growth in the number of ASes. However, they have also illustrated that the dynamics of BGP updates is highly volatile even at large time scales, with peak rates exceeding the daily averages by several orders of magnitude.

The complexity of the inter-AS routing system makes it difficult to isolate different factors behind these fluctuations [11, 12]. An approach alternative to inferring this factors directly is to build a realistic model for the dynamics of BGP updates. To this end, one needs an in-depth statistical characterization of fluctuations in BGP update time series, which is the subject of this work.

We aim at improving our understanding of these fluctuations, which can help in validating existing models [13] and in developing better ones. To study the statistical properties of BGP updates, we use historical BGP update logs spanning a period of 8.5 years, collected by the RouteViews project [14] from the BGP routers of four ASes (*AT&T*, *NTT*, *III*, and *Tinet*). Throughout the manuscript we refer to these routers as monitors. A BGP update log is the time series of BGP updates arriving at the monitor recorded in 1 second intervals. The four ASes analyzed in this work are among the largest Internet Service Providers (ISPs). Therefore, their corresponding BGP update traffic is a reflection of BGP dynamics taking place in the core of the Internet, where the BGP update volatility is believed to reach maximum rates. (Detailed information on data collection and pre-processing can be found in the [S1 Text](#) Section II).

To put our study in a broader context we wish to note that many natural and economic systems have also been found to exhibit extreme fluctuations. Examples include DNA sequences [15] and heartbeat intervals [16], climate variability [17, 18], earthquakes [19, 20], stock markets [21–23], and languages [24, 25].

Analysis

First we highlight the volatility of BGP updates series by reproducing the results of previous works [7]. We plot the average rate of BGP updates received by the *NTT* monitor on May 28th,

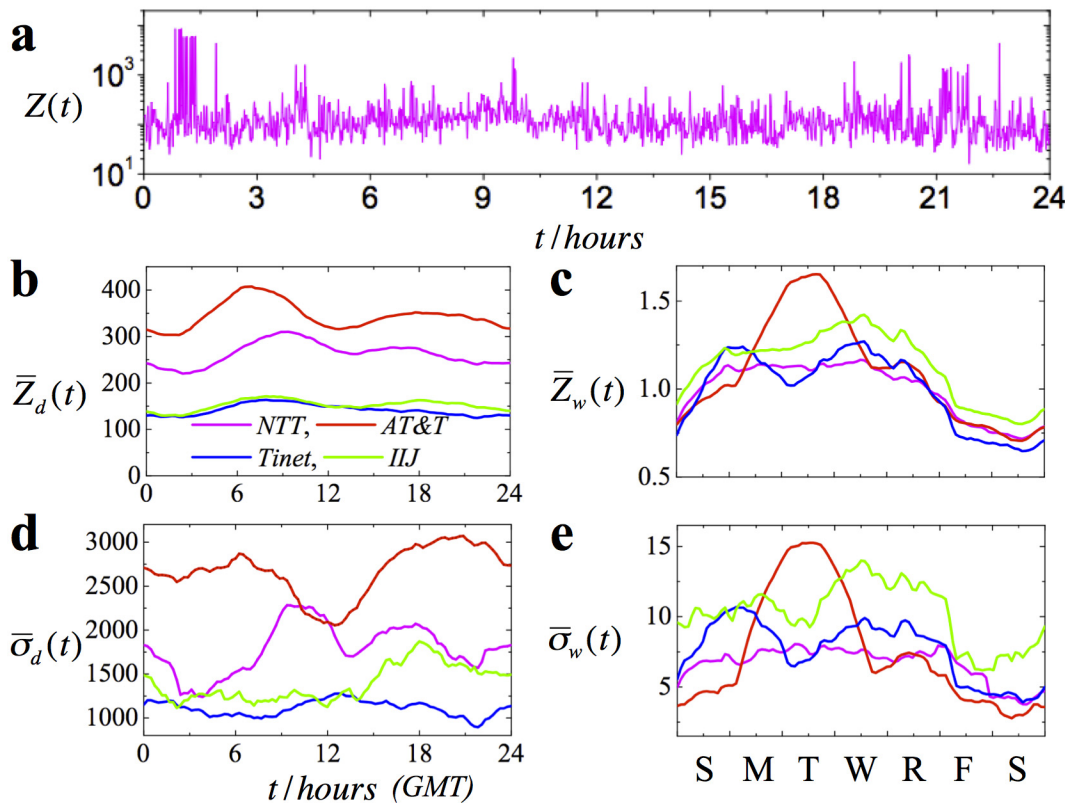


Fig 2. Time series of BGP updates. **a**, The number of updates received by the *NTT* monitor on May 28th, 2010, from 00.00 to 24.00 GMT. **b**, the intra-day pattern, $\bar{Z}_d(t)$, and, **d**, the standard deviation from the intra-day pattern, $\bar{\sigma}_d(t)$, of BGP updates measured for *NTT*, *IJ*, *Tinet*, and *AT&T* monitors. **c**, the intra-week pattern, $\bar{Z}_w(t)$, and, **e**, the standard deviation from the intra-week pattern, $\bar{\sigma}_w(t)$, of BGP updates measured for *NTT*, *IJ*, *Tinet*, and *AT&T* monitors.

doi:10.1371/journal.pone.0141481.g002

2010, from 00: 00 to 24: 00 Greenwich Mean Time (GMT). As seen from Fig 2a, in 1 minute interval the *NTT* monitor receives on the average several hundred updates, while extreme fluctuations occasionally produce 10^4 updates per minute. BGP updates are largely driven by two sources: spontaneous BGP events and maintenance sessions. The former consist of mostly spontaneous updates, such as misconfigurations, duplicate announcements and special events. Maintenance sessions, on the other hand, are periodic by nature and happen at certain times of the day on particular days of the week.

In order to separate the two sources of fluctuations we calculate the intra-day and intra-week patterns for the BGP update time series. The intra-day pattern, \bar{Z}_d , is then defined as the number of events taking place at a specific time of the day, t_{day} , averaged throughout the observation period:

$$\bar{Z}_d(t_{day}) = \frac{1}{N_d} \sum_{i=1}^{N_d} Z^i(t_{day}), \tag{1}$$

where N_d is the total number of days in the observation period, and $Z^i(t_{day})$ is the number of events at day i at t_{day} . The intra-week pattern $\bar{Z}_w(t_{week})$ is defined in a similar way after first normalizing the time series with the intra-day pattern.

$$\tilde{Z}(t) \equiv \frac{Z(t)}{\bar{Z}_d(t_{day}(t))}, \tag{2}$$

$$\bar{Z}_w(t_{week}) = \frac{1}{N_w} \sum_{i=1}^{N_w} \tilde{Z}^i(t_{week}) \tag{3}$$

Here N_w is the number of weeks in the observational period and $\tilde{Z}^i(t_{week})$ is the normalized number of events at week i at time of the week t_{week} (see [Methods](#) for details).

As seen from [Fig 2b](#), the intraday BGP update patterns reach maximum values in the interval from approximately 06: 00 to 10: 00 GMT, which is typical time for scheduling maintenance tasks [26]. The intraweek patterns, in their turn, are characterized by higher values during weekdays and smaller values during weekends. (see [Fig 2c](#)). We note that the standard deviations of the intraday and intraweek patterns, $\bar{\sigma}_d(t)$ and $\bar{\sigma}_w(t)$, tend to exceed the corresponding average values of the intra-day and the intra-week patterns by an order of magnitude, which is consistent with the extreme burstiness of the BGP updates ([Fig 2d and 2e](#)).

To characterize the volatility of the BGP updates we analyze the distribution of the number of BGP updates received by the monitor in 1 minute intervals. [Fig 3a](#) confirms the volatile nature of BGP updates. We find that all monitors are characterized by similar distributions $P(Z)$. Although the average number of BGP updates received per minute is quite small ($\bar{Z}_{NTT} = 250$), the peak values may occasionally exceed 10^5 BGP updates per minute. The distributions of the number of BGP updates, $P(Z)$, are positively skewed (measured skewness values are: $\gamma_1(AT\&T) = 45.7$, $\gamma_1(IIJ) = 121.2$, $\gamma_1(Tinet) = 49.1$, $\gamma_1(NTT) = 69.1$) and the distribution tails scale as a power-law, $P(Z) \sim Z^{-\mu}$ with $\mu = 2.51 \pm 0.11$ ($p = 0.992$ for *IIJ*, see [S1 Text](#) Section V for details). We also note that the observed power-law behavior of the tail of $P(Z)$ seems to be independent of the aggregation window size ([Fig 3b](#)).

The power-law distribution of the number of BGP updates implies that BGP routers should be able to cope with surges in the number of updates exceeding the corresponding average levels by several orders of magnitude. To understand how and when these surges occur we analyze correlation patterns of the BGP updates. We employ three standard methods traditionally used in the time-series analysis: auto-correlation function (ACF), power spectrum (PS), and the linear detrended fluctuation analysis (DFA1) (see [Methods](#), [S1 Text](#) Section IV, and Ref. [27] for details).

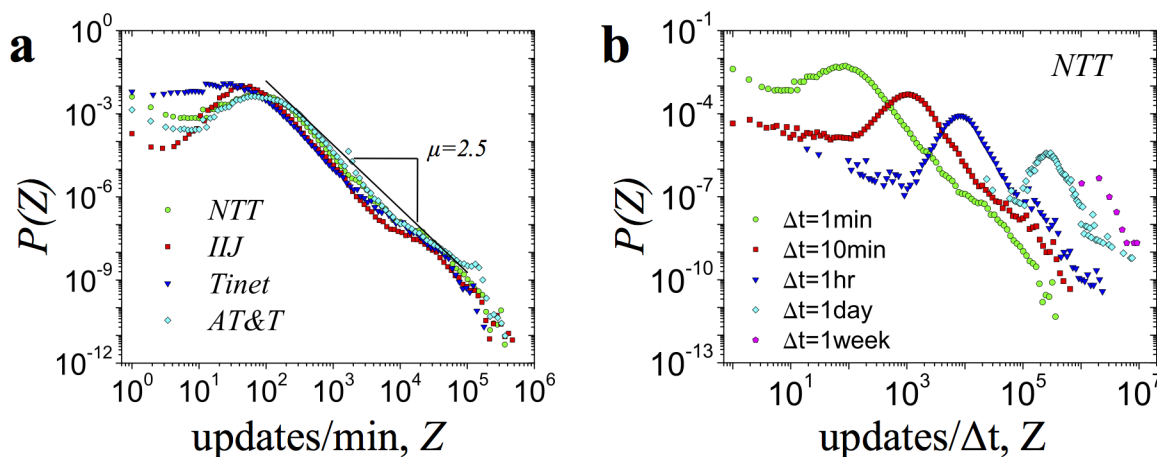


Fig 3. Extreme events in BGP dynamics. **a**, The distribution of the number of BGP updates received by the 4 monitors in 1 minute interval, All monitors collapse onto a single master curve. Power law regression fit yields a slope of $\mu = 2.3$ **b**, The distribution of number of updates $P(z)$ received by the *NTT* monitor calculated for aggregation window sizes $\Delta t = 1min$, $\Delta t = 10min$, $\Delta t = 1hour$, $\Delta t = 1day$ and $\Delta t = 1week$.

doi:10.1371/journal.pone.0141481.g003

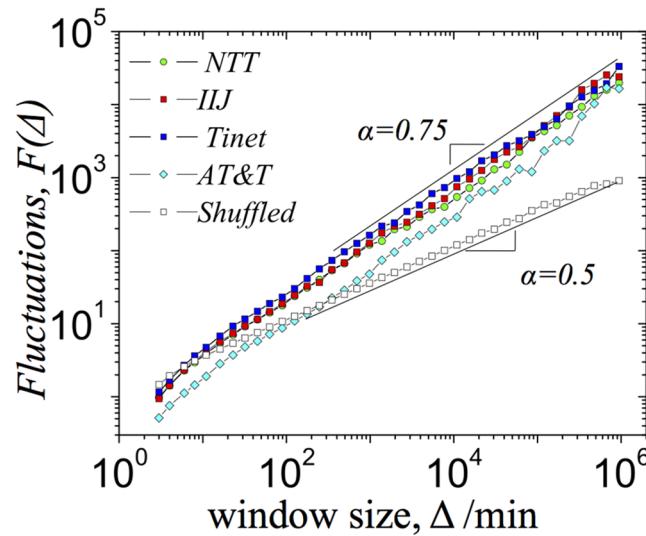


Fig 4. Correlations in the BGP update times series. Fluctuations of the detrended BGP update time series as a function of window size.

doi:10.1371/journal.pone.0141481.g004

Even though most of BGP update events last less than 1 minute, the duration of some of them may exceed several minutes [28]. Thus, to avoid possible correlations associated with long BGP updates in our subsequent analysis we use larger aggregation window size of $\Delta(t) = 10min$. Further, to eliminate possible spurious effects and correlations attributed to periodic activities we also normalize the BGP update data with both intra-week and intra-day patterns:

$$z(t) \equiv \frac{Z(t)}{\bar{Z}_w(t_{week}(t))\bar{Z}_d(t_{day}(t))}, \quad (4)$$

All three methods indicate the presence of long-range correlations in the BGP update time-series (see Fig 4 and S1 Fig). Specifically, we find that DFA1 performed for *NTT*, *IJ* and *Tinet* and *AT&T* indicates that fluctuations grow as a power-law with aggregation window size Δ , $F(\Delta) \sim \Delta^\alpha$, where $\alpha = 0.75$ (Fig 4). To highlight the effects of long-range correlations in the BGP updates time series we also performed DFA1 for the randomized counterparts of the BGP updates (see Methods). In the randomized case we obtained $F_{random}(\Delta) \sim \Delta^\alpha$ with $\alpha = 0.5$, which corresponds to the uncorrelated time series (Fig 4). Similar results are obtained by ACF and PS analysis. The autocorrelation function of the BGP updates decays as a power law over several orders of magnitude for all monitors, $ACF(\Delta z) \sim z^{-\gamma}$ (S1a Fig). We obtain similar γ values for three monitors: $\gamma = 0.5$ for *NTT*, and $\gamma = 0.4$ for *IJ* and *Tinet* monitors. The power spectrum density, $S(f)$, also decays as a power-law with frequency, $S(f) \sim f^{-\beta}$, where $\beta = 0.6$ for all monitors (S1b Fig). We note that the obtained values of correlation exponents approximately conform with expected relations, $\gamma = 1 - \beta$, $\alpha = \frac{\beta+1}{2}$, and $\gamma = 2(1 - \alpha)$ [17, 29–31].

The appearance of long-range correlations in BGP update time series indicates that at a given time the state of a particular BGP router is determined by its previous states. Consequently, long-range correlations may imply the presence of memory effects in the inter-domain Internet routing. To probe for the latter we ask, what is a typical time interval τ separating two large events. Formally, we define a return interval $\tau(q)$ as a time separation between two consecutive events $z(t_1)$ and $z(t_2)$, such that $z(t_1) > q$ and $z(t_2) > q$ (see Fig 5a). The evidence of memory in BGP update time series is seen in Fig 5b, which displays typical sequence of 500

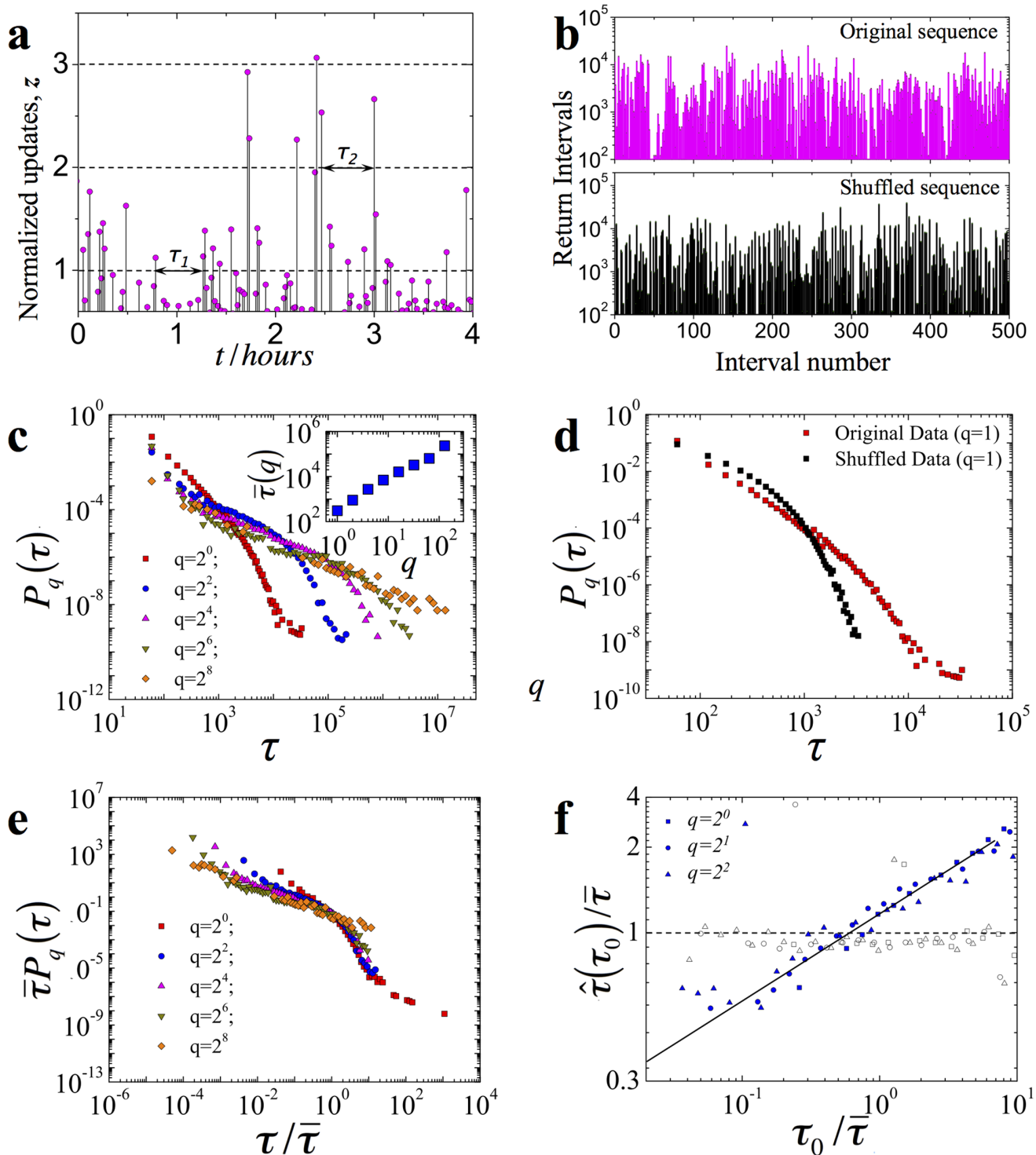


Fig 5. Return interval statistics of BGP updates. **a**, Schematic illustration of the BGP update return intervals. Shown are the intervals τ_1 and τ_2 calculated for threshold $q = 1$ and $q = 2$ respectively. **b**, Typical sequence of 500 BGP update return intervals for *NTT*, where $q = 4$, calculated for (magenta) original and (black) shuffled data. **c**, The distribution function $P_q(\tau)$ of BGP update return intervals of the *NTT*, calculated for different values of q . The inset depicts the average return interval $\bar{\tau}$ as a function of threshold q . **d**, $P_q(\tau)$ for BGP update return intervals of the *NTT* monitor calculated for $q = 1$. Original data is shown with red while shuffled data is shown with black. **e**, Scaled plots of the BGP return intervals for the *NTT* monitor. **f**, The mean conditional return interval $\hat{\tau}$ as a function of preceding return interval τ_0 for the *NTT* monitor. Both $\hat{\tau}$ and τ_0 are normalized with the mean return interval ($\bar{\tau}$). For BGP updates without memory we expect $\hat{\tau}(\tau_0) = 1$, as supported by the open symbols obtained for shuffled return interval data.

doi:10.1371/journal.pone.0141481.g005

consecutive return intervals for the *NTT* monitor. The original return interval data (shown in magenta) is characterized by “patches” of extreme return intervals, while there is no such “patches” in the shuffled data (shown in black) obtained by randomizing the time-order of the original series of BGP updates.

To further explore memory effects we analyze the distribution of return intervals $P_q(\tau)$ for the *NTT* monitor (Fig 5c). We note that $P_q(\tau)$ decays slower than the Poisson distribution, which is expected for uncorrelated data (Fig 5d). As q increases, the decay of $P_q(\tau)$ becomes slower and the average return interval $\bar{\tau}(q)$ increases implying that the larger events become increasingly rare (see the inset of Fig 5c). We also note that, independent of q , all the distributions $P_q(\tau)$, upon proper rescaling, collapse to a single master curve:

$$P_q(\tau) = \frac{1}{\bar{\tau}} f\left(\frac{\tau}{\bar{\tau}}\right), \tag{5}$$

where $f(x)$ does not depend on the threshold value (see Fig 5e and S3 Fig). The resulting master curve $f(x)$ fits a stretched exponential $\exp(-x^{-\gamma})$ with exponent $\gamma = 0.5$ ($p = 0.08$, see S1 Text Section V), which approximately matches the observed autocorrelation exponent $\gamma = 0.5$ [32]. We note that the observed scaling of $P_q(\tau)$ holds not only for *NTT* but also for the other three analyzed monitors (see S1 Text Section IV and S2 Fig).

Finally, to test memory effects directly we measure the average return interval $\hat{\tau}$ following immediately after return intervals of fixed duration τ_0 . Fig 5f and S3 Fig depict $\hat{\tau}$ as a function of τ_0 for three possible values of threshold q (filled symbols). We observe that $\hat{\tau}$ increases as a function of τ_0 , indicating that on the average longer (shorter) return intervals tend to follow longer (shorter) intervals. In contrast, $\hat{\tau}$ is independent of preceding return interval τ_0 for randomized data (open symbols in Fig 5f and S3 Fig).

Discussion

In this work, we investigated the statistical properties of BGP updates. Complementing previous studies, we confirmed that the rate of BGP updates is highly volatile, with extreme events at times exceeding the average rates by up to 4 orders of magnitude. We established that the distribution in the number of BGP updates received by a BGP monitor in a given time window is characterized by a power-law tail with exponent $\mu = 2.5$. We also found (using three independent methods) that the BGP update time series exhibit long-range correlations. The analysis of the return interval data revealed the universal scaling in the distribution of return intervals $P_q(\tau)$. We also found memory effects in the return interval data. Small (or large) return intervals separating BGP update events are more likely to be followed by small (or large) intervals.

The observed volatility and correlation properties of the BGP update dynamics place inter-domain Internet routing into the same class of phenomena as earthquakes [19, 20], climate [18], stock markets [21–23] and languages [24, 25]. Unlike these systems, however, the Internet routing is a fully engineered system. The observed dynamical similarities between these stochastic systems imply that the key mechanisms underlying Internet routing are in a certain way similar to the mechanisms governing the dynamics of stock markets or seismic movements in the Earth crust.

As with stock market price dynamics, one would wish to be able to predict BGP dynamics, or at least extreme events in it. To this end, one could benefit from the return interval scaling. The established scaling of $P_q(\tau)$ may allow one to approximate the statistics of return intervals for large events (characterized by large q values) using the much richer statistics of return intervals of smaller events.

The observed long-range correlations and memory effects indicate that the communication patterns between BGP routers are an outcome of an interplay between certain semi-deterministic processes. Such processes are well known at the low level of the operation of an individual BGP router (e.g. BGP route selection process). Yet this knowledge is as helpful as the knowledge about the dynamical properties of an individual molecule in a gas—when studying the properties of this gas (or the Internet in our case), some molecular details do matter, but most details are irrelevant.

Therefore the identification of a proper level of abstraction in modeling the dynamics of BGP routing is an important problem for understanding Internet dynamics. The statistical analysis of the BGP update time series that we have conducted here should serve as a basis for validation of existing models and for developing better ones.

Materials and Methods

Intraday and Intra-week Patterns

Consider series $Z(t)$, where Z is the number of events taking place at time t , and t is specified as UNIX timestamps. We first define functions $t_{day}(t)$ and $t_{week}(t)$ which map Unix timestamps t to respectively specific time of the day or specific time of the week ($t_{day} \in [0: 00, 24: 00]$, $t_{week} \in [Sunday, 0: 00, Saturday, 24: 00]$). Both t_{day} and t_{week} are calculated corresponding to the GMT time zone.

The intra-day pattern, \bar{Z}_d , is then defined as the number of events taking place at a specific time of the day, t_{day} , averaged throughout the observation period:

$$\bar{Z}_d(t_{day}) = \frac{1}{N_d} \sum_{i=1}^{N_d} Z^i(t_{day}), \tag{6}$$

where N_d is the total number of days in the observation period, and $Z^i(t_{day})$ is the number of events at day i at t_{day} . The intra-week pattern $\bar{Z}_w(t_{week})$ is defined in a similar way after first normalizing the time series with the intra-day pattern.

$$\tilde{Z}(t) \equiv \frac{Z(t)}{\bar{Z}_d(t_{day}(t))}, \tag{7}$$

$$\bar{Z}_w(t_{week}) = \frac{1}{N_w} \sum_{i=1}^{N_w} \tilde{Z}^i(t_{week}). \tag{8}$$

Here N_w is the number of weeks in the observational period and $\tilde{Z}^i(t_{week})$ is the normalized number of events at week i at time of the week t_{week} .

The standard deviations of the intraday and intra-week patterns are defined as

$$\sigma_d(t_{day}) \equiv \sqrt{\frac{1}{N_d} \sum_{i=1}^{N_d} (Z^i(t_{day}) - \bar{Z}_d(t_{day}))^2}, \tag{9}$$

$$\sigma_w(t_{week}) \equiv \sqrt{\frac{1}{N_w} \sum_{i=1}^{N_w} (\tilde{Z}^i(t_{week}) - \bar{Z}_w(t_{week}))^2} \tag{10}$$

Detrended Fluctuation Analysis

Detrended Fluctuation Analysis is a method designed to study correlations in time series [27]. Here we employ the linear version of the DFA, defined as follows. We first calculate the cumulative BGP update time series:

$$y(t) = \sum_{t'=t_i}^t (z(t') - \bar{z}), \tag{11}$$

where t_i is the initial time value in the series, $z(t)$ is the original time series and \bar{z} is its average value. The cumulative time series $y(t)$ is then divided into boxes of equal size Δ . In each box, a least squares linear fit to the $y(t)$ data is performed, representing the trend in that box. That is, for each box Δ we determine linear approximation for the corresponding piece of the time series:

$$y_{\Delta}(t) = m_{\Delta}t + b_{\Delta}, \tag{12}$$

where m_{Δ} and b_{Δ} are the slope and the intercept of the straight line. Next we detrend the integrated time series, $y(t)$, by subtracting the local trend, $y_{\Delta}(t)$, in each box. The root-mean-square fluctuation of this integrated and detrended time series is calculated:

$$F(\Delta) = \sqrt{\frac{1}{N} \sum_{t=t_i}^{t_f} [y(t) - y_{\Delta}(t)]^2}, \tag{13}$$

where N is the total number of points in the original time series, t_i and t_f are respectively the initial and final time values in the series.

This fluctuation measurement process is repeated at a range of different box sizes Δ . The fluctuations typically exhibit a power law scaling as a function of box size:

$$F(\Delta) \sim \Delta^{\alpha}, \tag{14}$$

depending on the observed exponent α one can distinguish anti-correlated fluctuations ($\alpha < 1/2$), uncorrelated fluctuations ($\alpha = 1/2$), and correlated fluctuations ($\alpha > 1/2$).

Data Randomization

To assess the significance of correlations and memory effects in the BGP update time series we compare original results to those obtained for randomized (shuffled) datasets. In all experiments the randomization is performed at the most basic level: for a given time series $Z(t)$ we obtain its randomized (shuffled) counterpart by randomly rearranging time stamps attributed to each element in the series. Shuffled data is subsequently normalized and binned using the same procedures as those applied to original data.

Supporting Information

S1 Text.

(PDF)

S1 Fig. Correlations in the BGP update times series. **a**, The autocorrelation function, ACF and **b**, The Power Spectrum $S(f)$.

(TIFF)

S2 Fig. (Left column) The distribution of return intervals, $P_q(\tau)$, for BGP updates of **a**, *NTT*, **c**, *AT&T*, **e**, *Tinet*, and **g**, *III* monitors. The distributions are calculated for different values of

threshold q . (Right column) Rescaled plots of the BGP return intervals of **b**, *NTT*, **d**, *AT&T*, **f**, *Tinet*, and **h**, *IJJ* monitors.

(TIFF)

S3 Fig. The mean conditional interval $\hat{\tau}(\tau_0)$ divided by $\bar{\tau}$ as a function of $\frac{\tau_0}{\bar{\tau}}$ for **a** *NTT*, **b** *AT&T*, **c** *Tinet*, and **d** *IJJ* monitors. In time series without memory, $\hat{\tau}(\tau_0) = 1$, indicated by the open symbols that show the shuffled return interval data.

(TIFF)

S4 Fig. KS goodness of fit tests for **a**, the distribution of number of BGP updates, $P(z)$ for the *NTT* monitor, and **b**, the distribution of return intervals $P_q(\tau)$ for the *NTT* monitor.

(TIFF)

Acknowledgments

We thank H. E. Stanley, kc claffy, and D. Rybski for useful discussions and suggestions.

Author Contributions

Conceived and designed the experiments: MK AE SH DK. Performed the experiments: MK AE. Analyzed the data: MK AE. Wrote the paper: MK AE DK.

References

1. Elmokashfi A, Kvalbein A, Dovrolis C. BGP Churn Evolution: A Perspective From the Core. *IEEE/ACM Transactions on Networking*. 2011 Oct;(99).
2. Hawkinson J, Bates T. Guidelines for creation, selection, and registration of an Autonomous System (AS); 1996. RFC1930.
3. Rekhter Y, Li T, Hares S. A Border Gateway Protocol 4 (BGP-4); 2006. RFC4271.
4. Meyer D, Zhang L, Fall K. Report from the IAB Workshop on Routing and Addressing; 2007. <http://tools.ietf.org/id/draft-iab-raws-report-02.txt>.
5. Varghese G. *Network Algorithmics*. Morgan Kaufmann, San Fransisco.; 2004.
6. Wang L, Zhao X, Pei D, Bush R, Massey D, Mankin A, et al. Observation and analysis of BGP behavior under stress. In: *Internet Measurement Workshop*; 2002. p. 183–195.
7. Elmokashfi A, Kvalbein A, Dovrolis C. BGP churn evolution: a perspective from the core. *Networking, IEEE/ACM Transactions on*. 2012; 20(2):571–584. doi: [10.1109/TNET.2011.2168610](https://doi.org/10.1109/TNET.2011.2168610)
8. Cittadini L, Muhlbauer W, Uhlig S, Bush R, Francois P, Maennel O. Evolution of internet address space deaggregation: myths and reality. *Selected Areas in Communications, IEEE Journal on*. 2010; 28(8):1238–1249. doi: [10.1109/JSAC.2010.101002](https://doi.org/10.1109/JSAC.2010.101002)
9. Farinacci D, Fuller V, Meyer D, Lewis D. Locator/ID Separation Protocol (LISP); 2012. Draft-ietf-lisp-22. IETF Internet Draft.
10. Boguñá M, Papadopoulos F, Krioukov D. Sustaining the Internet with hyperbolic mapping. *Nat Commun*. 2010 09; 1. PMID: [20842196](https://pubmed.ncbi.nlm.nih.gov/20842196/)
11. Caesar M, Subramanian L, Katz RH. Towards localizing root causes of BGP dynamics. U.C.Berkely; 2003. UCB/CSD-04-1302.
12. Feldmann A, Maennel O, Mao ZM, Berger A, Maggs B. Locating Internet routing instabilities. *ACM SIGCOMM Computer Communication Review*. 2004; 34(4):205–218. doi: [10.1145/1030194.1015491](https://doi.org/10.1145/1030194.1015491)
13. Valler NC, Butkiewicz M, Prakash BA, Faloutsos M, Faloutsos C. Non-binary information propagation: Modeling BGP routing churn. In: *NetSciCom 2011, (colocated with IEEE INFOCOM)*; 2011. p. 900–905.
14. Routeviews Project Page; <http://www.routeviews.org>.
15. Peng C, Buldyrev S, Goldberger A, Havlin S, Sciortino F, Simons M, et al. Long-range correlations in nucleotide sequences. *Nature*. 1992; 356(6365):168–170. doi: [10.1038/356168a0](https://doi.org/10.1038/356168a0) PMID: [1301010](https://pubmed.ncbi.nlm.nih.gov/1301010/)
16. Peng CK, Mietus J, Hausdorff J, Havlin S, Stanley HE, Goldberger A. Long-range anticorrelations and non-Gaussian behavior of the heartbeat. *Physical review letters*. 1993; 70(9):1343. doi: [10.1103/PhysRevLett.70.1343](https://doi.org/10.1103/PhysRevLett.70.1343)

17. Koscielny-Bunde E, Bunde A, Havlin S, Roman HE, Goldreich Y, Schellnhuber HJ. Indication of a Universal Persistence Law Governing Atmospheric Variability. *Phys Rev Lett*. 1998 Jul; 81:729–732. doi: [10.1103/PhysRevLett.81.729](https://doi.org/10.1103/PhysRevLett.81.729)
18. Eichner JF, Koscielny-Bunde E, Bunde A, Havlin S, Schellnhuber HJ. Power-law persistence and trends in the atmosphere: A detailed study of long temperature records. *Phys Rev E*. 2003 Oct; 68:046133. doi: [10.1103/PhysRevE.68.046133](https://doi.org/10.1103/PhysRevE.68.046133)
19. Lennartz S, Livina VN, Bunde A, Havlin S. Long-term memory in earthquakes and the distribution of interoccurrence times. *EPL (Europhysics Letters)*. 2008; 81(6):69001. doi: [10.1209/0295-5075/81/69001](https://doi.org/10.1209/0295-5075/81/69001)
20. Bunde A, Kropp J, Schellnhuber HJ. The science of disasters: climate disruptions, heart attacks, and market crashes. vol. 2. Springer Science & Business Media; 2002.
21. Yamasaki K, Muchnik L, Havlin S, Bunde A, Stanley HE. Scaling and memory in volatility return intervals in financial markets. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(26):9424–9428. doi: [10.1073/pnas.0502613102](https://doi.org/10.1073/pnas.0502613102) PMID: [15980152](https://pubmed.ncbi.nlm.nih.gov/15980152/)
22. Preis T, Reith D, Stanley HE. Complex dynamics of our economic life on different scales: insights from search engine query data. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. 2010; 368(1933):5707–5719. doi: [10.1098/rsta.2010.0284](https://doi.org/10.1098/rsta.2010.0284)
23. Preis T, Kenett DY, Stanley HE, Helbing D, Ben-Jacob E. Quantifying the behavior of stock correlations under market stress. *Scientific reports*. 2012; 2(752).
24. Gao J, Hu J, Mao X, Perc M. Culturomics meets random fractal theory: insights into long-range correlations of social and natural phenomena over the past two centuries. *Journal of The Royal Society Interface*. 2012; 9(73):1956–1964. doi: [10.1098/rsif.2011.0846](https://doi.org/10.1098/rsif.2011.0846)
25. Petersen AM, Tenenbaum JN, Havlin S, Stanley HE, Perc M. Languages cool as they expand: Allometric scaling and the decreasing need for new words. *Scientific Reports*. 2012; 2(943).
26. Francois P, Coste PA, Decraene B, Bonaventure O. Avoiding disruptions during maintenance operations on BGP sessions. *IEEE Transactions on Network and Service Management*. 2007; 4(3):1–11. doi: [10.1109/TNSM.2007.021102](https://doi.org/10.1109/TNSM.2007.021102)
27. Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL. Mosaic organization of DNA nucleotides. *Physical Review E*. 1994; 49(2):1685. doi: [10.1103/PhysRevE.49.1685](https://doi.org/10.1103/PhysRevE.49.1685)
28. Labovitz C, Ahuja A, Bose A, Jahanian F. Delayed Internet routing convergence. *ACM SIGCOMM Computer Communication Review*. 2000; 30(4):175–187. doi: [10.1145/347057.347428](https://doi.org/10.1145/347057.347428)
29. Beran J. *Statistics for long-memory processes*. vol. 61. Chapman and Hall, New York; 1994.
30. Barabasi AL, Stanley HE. *Fractal Concepts in Surface Growth*. Cambridge University Press; 1995.
31. Kantz H, Schreiber T. *Nonlinear time series analysis*. vol. 7. Cambridge university press; 2004.
32. Eichner JF, Kantelhardt JW, Bunde A, Havlin S. Statistics of return intervals in long-term correlated records. *Phys Rev E*. 2007 Jan; 75:011128. doi: [10.1103/PhysRevE.75.011128](https://doi.org/10.1103/PhysRevE.75.011128)