frontiers in Physics

REVIEW
published: 03 December 2020
doi: 10.3389/fphy.2020.00286

Bitcoin Transaction Networks: An Overview of Recent Results

Nicoló Vallarano[1], *Claudio J. Tessone*[2,3]* *and Tiziano Squartini*[1]

[1] IMT School of Advanced Studies, NETWORKS Unit, Lucca, Italy, [2] UZH Blockchain Center, University of Zürich, Zurich, Switzerland, [3] URPP Social Networks, University of Zürich, Zurich, Switzerland

Cryptocurrencies are distributed systems that allow exchanges of native (and non-) tokens between participants. The availability of the complete historical bookkeeping opens up an unprecedented possibility: that of understanding the evolution of a cryptocurrency's network structure while gaining useful insights into the relationships between users' behavior and cryptocurrency pricing in exchange markets. In this article we review some recent results concerning the structural properties of the *Bitcoin Transaction Networks*, a generic name referring to a set of three different constructs: the *Bitcoin Address Network*, the *Bitcoin User Network*, and the *Bitcoin Lightning Network*. The picture that emerges is of a system growing over time, which becomes increasingly sparse and whose mesoscopic structural organization is characterized by the presence of an increasingly significant core-periphery structure. Such a peculiar topology is accompanied by a highly uneven distribution of bitcoins, a result suggesting that Bitcoin is becoming an increasingly centralized system at different levels.

Keywords: blockchain, bitcoin, complex networks, centralization, inequality, null models

OPEN ACCESS

Edited by:
Marco Alberto Javarone, University College London, United Kingdom

Reviewed by:
Chengyi Xia, Tianjin University of Technology, China
Matúš Medo, University of Electronic Science and Technology of China, China

***Correspondence:**
Claudio J. Tessone
claudio.tessone@business.uzh.ch

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *24 April 2020*
Accepted: *25 June 2020*
Published: *03 December 2020*

Citation:
Vallarano N, Tessone CJ and Squartini T (2020) Bitcoin Transaction Networks: An Overview of Recent Results. Front. Phys. 8:286. doi: 10.3389/fphy.2020.00286

1. INTRODUCTION

A cryptocurrency is an online payment system for which the storage and verification of transactions—and therefore the safeguarding of the system's consistency itself—are *decentralized*, i.e., do not require the presence of a trusted third party. This result can be achieved by securing financial transactions through a clever combination of cryptographic technologies [1].

Bitcoin, the first and most popular cryptocurrency, was introduced in 2008 by Satoshi Nakamoto [2]. It consists of a decentralized peer-to-peer network to which users connect to exchange property in the account units of the system, i.e., to perform transactions of *bitcoins*. Each transaction becomes part of a publicly available ledger, the *blockchain*, after having been validated by so-called *miners*, i.e., users who verify the validity of issued transactions according to the consensus rules that are part of the Bitcoin protocol [3, 4]. A new block, containing transactions known to the miner since the last block, is "mined" every 10 min on average, thereby adding new transactions to the blockchain; thus these transactions are "confirmed," in turn enabling users to spend the bitcoins they received through them[1]. The cryptography protocols that Bitcoin rests upon aim to prevent the so-called *double-spending problem*, i.e., the possibility of the same digital token being spent more than once in the absence of a central party that guarantees the validity of the transactions [1, 2]; remarkably, the transaction-verification mechanism Bitcoin relies on allows its entire transaction history to be openly accessible, a feature that, in turn, allows researchers to analyze Bitcoin transactions in different network representations.

[1]In practice, the so-called "6 confirmations" rule is followed: once a transaction is included in a block that is followed by at least six additional blocks [5], the transaction can be safely considered confirmed.

The gain in Bitcoin's popularity has given rise to new problems for its community, including (i) the lack of *scalability* of the transaction-verification method, i.e., the relatively low maximum number of transactions that can be verified per second, especially when compared with mainstream competitors, such as centralized payment networks, (ii) the increased *concentration of mining power* in mining pools, which implies that the verification mechanism in the network is becoming less and less distributed, and (iii) the tendency of users to *hoard*. In order to overcome these problems, which threaten the overall functioning of Bitcoin as a medium of exchange, new instruments have been adopted. Proposed in 2015 [6], the *Bitcoin Lightning Network* (BLN) is a "Layer 2" protocol that can operate on top of blockchain-based (Bitcoin-like) cryptocurrencies by creating bilateral channels for *off-chain* payments which are then settled concurrently on the blockchain, once the channels are closed. As both the transaction fees and the blockchain confirmation are no longer required, the network is spared from avoidable burdens; moreover, the key features of Bitcoin, i.e., its *decentralized architecture*, its *political organization*, and its *wealth distribution*, are no longer sacrificed, while the circulation of the native assets is enhanced.

Bitcoin is almost 10 years old; however, while a large amount of literature concerning either the purely financial or the purely engineering aspects of it exists (e.g., prediction of the exchange rate between Bitcoin and the US dollar [7], statistical properties of the exchange rate [8], statistical properties of Bitcoin daily log-returns [9], comparison of Bitcoin volatility with that of the exchange rates of major global currencies [10, 11], identification of factors influencing the Bitcoin price [12], predictability of the Bitcoin price via machine-learning techniques [13], the interplay between social interactions and movements of the Bitcoin price [14, 15], and the problem of de-anonymization of Bitcoin users [16–20]), only recently have researchers started to investigate the *structural* properties of Bitcoin. In Kondor et al. [21], the authors consider the network of transactions between addresses at the weekly time scale, showing the emergence of power-law distributions and that the number of incoming transactions reflects the wealth of nodes; in Javarone and Wright [22], the network of transactions between users is studied at the macroscale, in order to check for its small-worldness; in Parino et al. [23], the authors investigate the network of international Bitcoin flows, identifying socio-economic factors that drive Bitcoin adoption across countries. In general, however, the works analyzing Bitcoin from a network perspective provide a quite limited view of its evolution, focusing either on a single representation of the network or on a relatively short period of time; even those studies that address the problem from a wider perspective [24, 25] are often limited to a purely descriptive analysis and do not compare empirical observations with the outcomes of proper models.

In this article we summarize the results of three papers [26–28], providing a comprehensive overview of the empirical traits that characterize Bitcoin evolution and framing them within models rooted in statistical physics. In Bovet et al. [26], the authors analyzed the local properties of two Bitcoin representations, the *Bitcoin Address Network* (BAN) and the *Bitcoin User Network* (BUN), and looked for the presence of correlations between (exogenous) price movements and (endogenous) changes in the topological structure of the networks. In Bovet et al. [27], the mesoscale structure of the BUN is examined; particular attention is paid to identifying the best network model able to describe the structure; in addition, the same exercise as above is carried out, i.e., comparison of the evolution of purely structural properties and the appearance of price bubbles in a cyclical fashion. Lastly, in Lin et al. [28], the evolution of the BLN's topology is investigated, revealing that the BLN is becoming an increasingly centralized system and that the "capital" is becoming increasingly unevenly distributed.

2. DATA

As previously mentioned, Bitcoin relies on a decentralized public ledger, the blockchain, that records all transactions between Bitcoin users. A transaction is a set of input and output addresses; the output addresses that are "unspent," i.e., not yet recorded on the ledger as input addresses, can be claimed—and therefore spent—only by the owner of the corresponding cryptographic key. This is the reason one speaks of *pseudonimity*: an observer of the blockchain can see all unspent *addresses* but cannot link them to the actual owners.

2.1. The Bitcoin Address Network (BAN)

The BAN is the simplest network that can be constructed from the blockchain records. From a technical point of view, it is a directed weighted graph whose nodes represent addresses; the directions and weights of the links between nodes are provided by the input-output relationships defining the transactions recorded on the blockchain. The BAN was considered over a period of 9 years, from 9th January 2009 to 18th December 2017, at the end of which the data set consisted of 304,111,529 addresses, between which a total number of 283,028,575 transactions were performed. In terms of traded volume, the transactions between addresses amounted to 4,432,597,496 bitcoins.

2.2. The Bitcoin User Network (BUN)

Since the same owner may control several addresses [18], one can define a network of "users" whose nodes are *clusters of addresses*. These clusters are derived by implementing different *heuristics* provided by the state-of-the-art literature [16, 17, 29, 30]. The "user networks" thus obtained should not be regarded as perfect representations of the actual networks of users, but rather as attempts to group addresses while minimizing the presence of false positives. Two heuristics have been employed here: the multi-input one (based on the assumption that addresses appearing as input to the same transaction are controlled by the same user) and the change address one (based on the assumption that a new address appearing as output of a transaction and with the smallest amount of transferred money must belong to the input user). The BUN was considered over the same period as the BAN (i.e., 9 years from 9th January 2009 to 18th December 2017), at the end of which the data set consisted of 16,749,939 users, between which a total number of 224,620,265 transactions took place. In terms of traded volume, the transactions between users amounted to 3,114,359,679 bitcoins.

2.3. The Bitcoin Lightning Network (BLN)

The BLN is constructed similarly to the way the BAN is defined. It is a directed weighted graph whose nodes are addresses exchanging bitcoins on "Layer 2." Three different representations of the BLN have been studied so far: the daily, the weekly, and the daily-block representations. While a daily/weekly snapshot includes all channels that were found to be active during that day/week, a daily-block snapshot consists of all channels that were found to be active at the time the first block of the day was released (hence, the transactions considered for the daily-block representation are a subset of those constituting the daily representation). The BLN was considered over a period of 18 months, from 14th January 2018 to 13th July 2019, at the end of which the network consisted of 8,216 users, 122,517 active channels, and 2732.5 transacted bitcoins.

2.4. Notation

Although information about the magnitude of transactions is available, the BAN and the BUN were analyzed as binary directed networks; as such, they are completely specified by their binary asymmetric adjacency matrices, $\mathbf{A}^{(t)}_{\mathrm{BAN}}$ and $\mathbf{A}^{(t)}_{\mathrm{BUN}}$, at time t. The generic entry $a^{(t)}_{ij}$ is equal to 1 if at least one transaction between address (user) i and address (user) j takes place, i.e., if bitcoins are transferred from address (user) i to address (user) j, during the time snapshot t and is equal to 0 otherwise. The BLN, on the other hand, is a weighted undirected network, represented by a symmetric matrix $\mathbf{W}^{(t)}_{\mathrm{BLN}}$ whose generic entry $w^{(t)}_{ij} = w^{(t)}_{ji}$ indicates the total amount of money exchanged between i and j, across all channels, at time t; here, we will focus mainly on its binary projection $\mathbf{B}^{(t)}_{\mathrm{BLN}}$, whose generic entry is $b^{(t)}_{ij} = b^{(t)}_{ji} = 1$ if $w^{(t)}_{ij} = w^{(t)}_{ji} > 0$ and $b^{(t)}_{ij} = b^{(t)}_{ji} = 0$ otherwise.

3. RESULTS

3.1. The Bitcoin Address Network and the Bitcoin User Network

Let us start by reviewing results on the BAN and the BUN at the weekly time scale. Similar results have been obtained for the BAN and the BUN at the daily time scale [26].

3.1.1. Basic Statistics

We begin by commenting on the evolution of some basic statistics characterizing the BAN and BUN that, as noted elsewhere [19], have started to evolve in a more stationary fashion since mid-2011. As **Figure 1** shows, both the number of nodes N and the number of links $L = \sum_i \sum_{j(\neq i)} a_{ij}$ increase steadily over time, irrespective of the representation considered; the link density $d = \frac{L}{N(N-1)}$, however, decreases, meaning that the system is becoming sparser. The dependence of d on N can be better specified, from a mathematical point of view, upon observing that the average degree $\overline{k^{\mathrm{in}}} = \overline{k^{\mathrm{out}}} = \frac{\sum_i \sum_{j(\neq i)} a_{ij}}{L} = \frac{L}{N}$ is either constant (for the BUN) or limited (for the BAN) over time [26]; hence, it follows that $L \propto N$ and $d \sim N^{-1}$.

3.1.2. Degree Distributions

Generally speaking, both out-degrees and in-degrees are characterized by heavy-tailed distributions, indicating that a large number of low-connectivity nodes coexist with a few hubs whose degree is several orders of magnitude greater. A visual inspection of the functional form of the degree distributions suggests that the out-degrees distribution follows a power law [26, 31]. To test this hypothesis Bovet et al. [26] employed an algorithm based on a double Kolmogorov-Smirnov statistical test [32]; they found that the hypothesis above cannot be rejected, at a 0.05 confidence level, for almost half of the considered snapshots.

Of particular interest is the evolution of the out-degrees standard deviation, especially in regard to its informativeness about exogenous events. As an example, consider the failure in February 2014 of Mt. Gox, a quasi-monopolist exchange market at the time. This event deeply affected the overall Bitcoin structure: the percentage of snapshots for which the null hypothesis (that the out-degrees distribution follows a power law) can be rejected was $\simeq$ 50% *before* February 2014 and dropped to $\simeq$ 25% *afterwards*.

Bovet et al. [26] also argued that the presence of heavy-tailed distributions may be explained by a mechanism similar to that of preferential attachment: new, or occasional, users "preferentially" connect to already well-connected nodes (exchange markets, utility providers, etc.), thus leading to the formation of super-hubs. Elsewhere it has been argued that the related mechanism known as "fittest-gets-richer" or "good-gets-richer" [33] may be also at work, with the computational resources of a node playing the role of its fitness [22].

3.1.3. Bitcoin Structure vs. Bitcoin Price

The result concerning the evolution of the out-degrees distribution suggests that the Bitcoin network structure indeed carries signatures of exogenous events. As in this case, the non-structural quantity *par excellence* is represented by the *price* of the currency, it may be of interest to look for the presence of correlations between the evolution of the price and the evolution of purely topological quantities. Justification for such an analysis rests upon the simple consideration that the price of a cryptocurrency is ultimately related to the behavior of users whose "network" activity translates into that of establishing connections with other nodes, whence our expectation of finding some traces of the aforementioned correlations.

The simplest analysis is based on drawing scatterplots of the network size and network link density vs. the Bitcoin price (in USD). As **Figure 2** shows, a clear trend appears, indicating that the price and the network size N (respectively, the link density d) are overall positively (respectively, negatively) correlated throughout the entire Bitcoin history. Notice, however, the trend inversion that occurs immediately after the Mt. Gox failure; it is a consequence of the prolonged price decrease observed in 2014–2016, during which the network size increased by almost one order of magnitude.

To further confirm the presence of a double regime, Bovet et al. [26] inspected the correlation between the moments of the out-degrees distribution and the Bitcoin price over time. To this end, the so-called "ratio of price to its moving average" (RPMA)

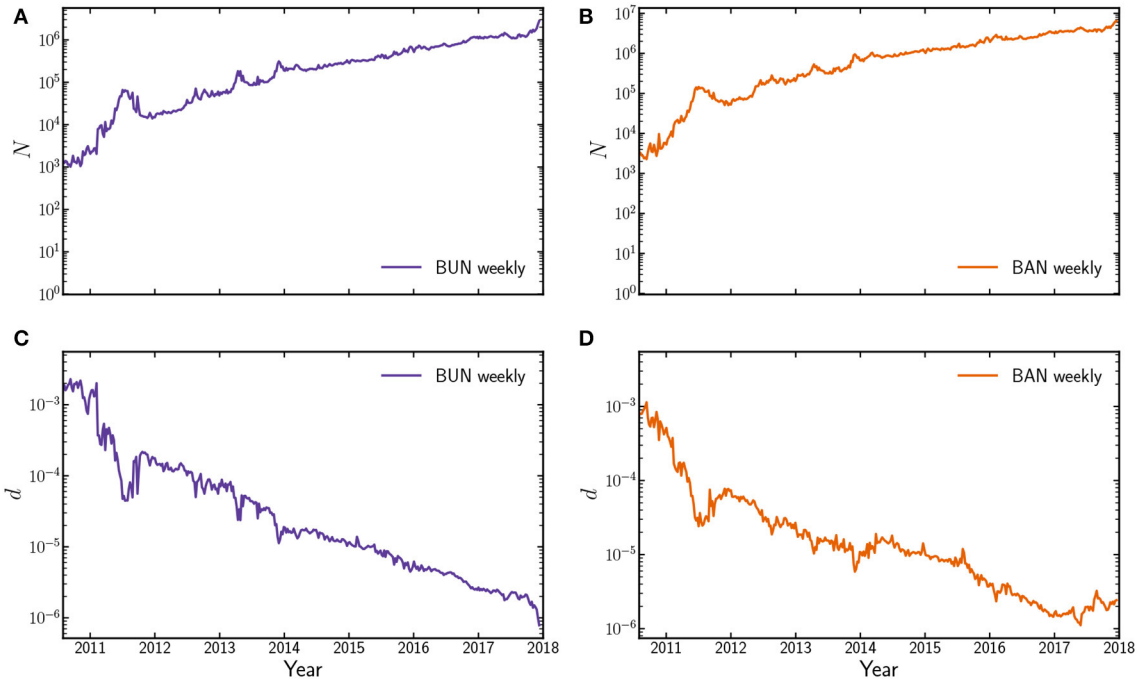

FIGURE 1 | Evolution of basic statistics, i.e., the number of nodes **(A,B)** and the link density **(C,D)** for two Bitcoin network representations, i.e., the BAN **(A,C)** and the BUN **(B,D)** at the weekly time scale, from July 2010 to 18th December 2017 (i.e., for networks with at least 200 nodes). Although the network size increases, it becomes sparser (irrespectively from the considered representation). Similar results are observed for the BAN and the BUN at the daily time scale. See also Bovet et al. [26].

indicator was defined:

$$\mathrm{RPMA}_t = 100 \log_{10} \left(\frac{P_t}{\frac{1}{\tau} \sum_{s=t-1-\tau}^{t-1} P_s} \right), \qquad (1)$$

where τ represents a tunable temporal parameter. As shown in Bovet et al. [26], the standard deviation and the kurtosis diverge as the network size becomes larger than the value corresponding to the Mt. Gox failure, thus confirming the "two regimes" hypothesis. Moreover, as **Figure 2** shows, larger values of the moments (observed *after* the Mt. Gox failure) correspond to price drops, while temporal snapshots corresponding to smaller values of the same quantities seem to be characterized by price increases.

A multivariate Granger test [34] was also carried out to unveil possible lagged correlations hidden in the data (see Figure 5 in Bovet et al. [26]). For this purpose, the data were split into two sub-samples, corresponding to the time periods 2010–2013 and 2014–2017, and the number of nodes N, the number of links L, and the higher moments of the empirical out- and in-degrees distributions were related to the log-returns of the Bitcoin price (in USD) within each sub-sample. To sum up, when the BUN is considered on the weekly time scale, a positive feedback loop occurs between N and the price log-returns, whereas on the daily time scale a price increase predicts an increase in the number of nodes N but not vice versa. The causality structure is consistent within the two sub-samples.

3.1.4. Analysis of the BUN Mesoscale Structure

We now review the results concerning the mesoscale structure of the BUN. A recently proposed method [35] based on the *surprise* score function was adopted by Bovet et al. [27] to assess the statistical significance of a peculiar mesoscale organization known as the *core-periphery* structure. According to the interpretation proposed in de Jeude et al. [35], revealing the core-periphery structure by minimizing the surprise means distinguishing the partition that is least likely to be explained by the null model known as the *random graph model* (RGM) relative to the null model known as the *stochastic block model* (SBM); see also **Appendix A**. As **Figure 3** shows, a core-periphery structure is indeed present; more precisely, during 2014–2015 the core size amounts to $\simeq$ 30% of the total network size; but after 2016 it seems to shrink back to 2010–2013 levels. The presence of a core-periphery structure indicates that the BUN is characterized by subgraphs with very different link densities—evidence that cannot be accounted for by a model, such as the RGM, with just one global parameter.

A closer inspection of the BUN core-periphery structure reveals it to be even richer. In fact, the core portion of the BUN is the strongly connected component (SCC) of a *bow-tie* structure whose remaining portions (the IN and OUT components) make up the BUN periphery [27]. More specifically, while the SCC is the set of nodes that are mutually reachable (i.e., there exists a directed path from any node to any other node within the SCC), the IN and OUT components are defined, respectively, as the

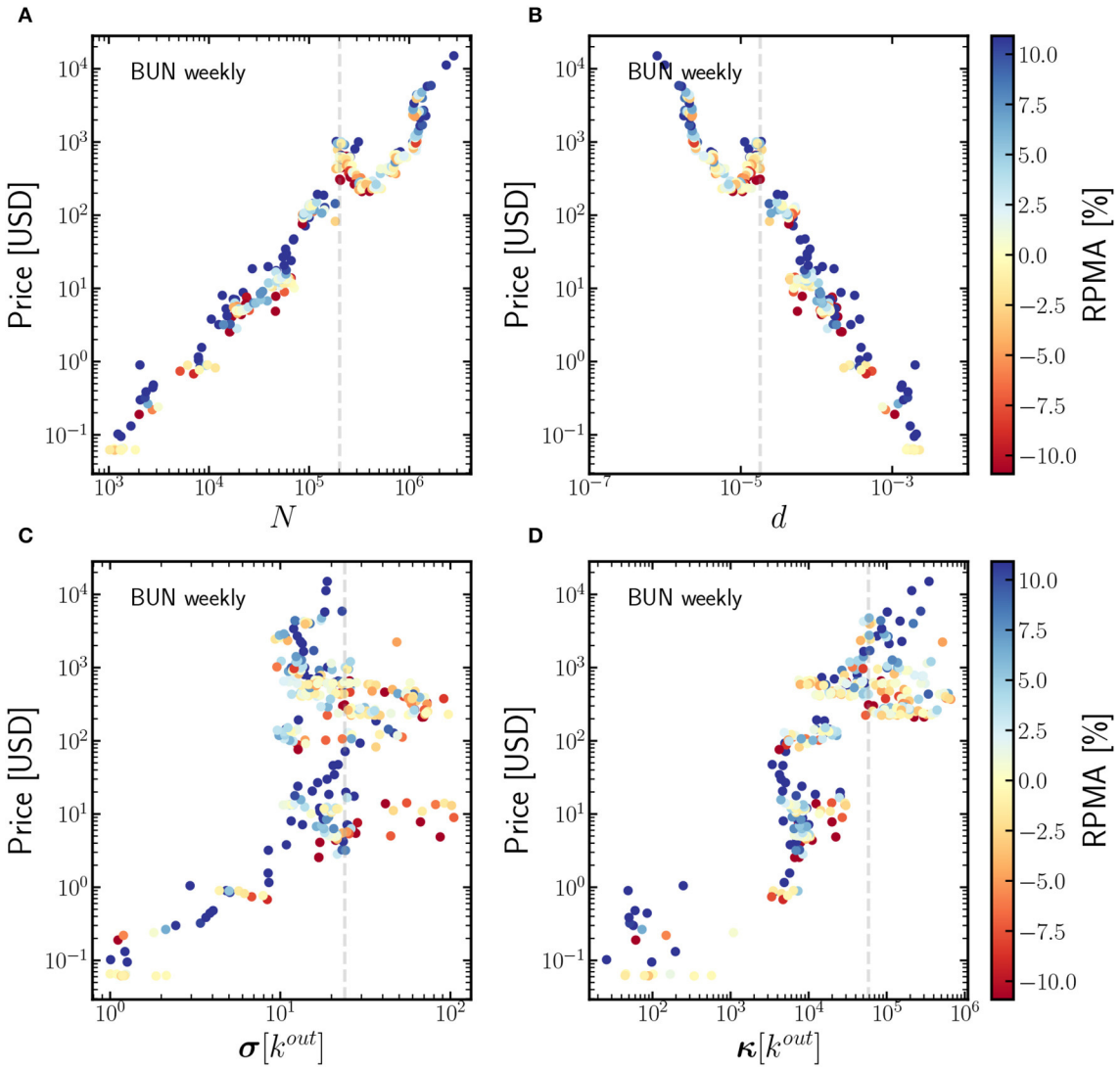

FIGURE 2 | Correlation between the Bitcoin price in USD, the basic statistics (number of nodes and link density -**A,B**) and the moments of the out-degrees distribution **(C,D)** for the BUN at the weekly time scale. Additionally, each dot representing an observation is colored according to the value of the Ratio between the current Price and its Moving Average (RPMA) indicator. The vertical, dashed line coincides with the bankruptcy of Mt. Gox. Purely structural quantities are correlated with exogenous quantities as the Bitcoin price; see, for example, the evolution of the out-degrees standard deviation whose larger values (observable after the Mt. Gox failure) correspond to price drops. See also [26].

set of nodes from which the SCC can be reached and the set of nodes that can be reached from the SCC. Hence, the picture provided by the evolution of the core-periphery structure can be further refined as follows: since 2016 both the SCC and the OUT component have shrunk while the IN component has become the dominant portion of the network [27]. Other SCCs are visible but are negligibly small relative to the largest one, which seems to indicate that they are, in fact, single nodes pointing to (or pointed to by) hubs.

An additional analysis aimed at better quantifying the extent to which a generic, purely topological quantity X and the Bitcoin price are related can be carried out by plotting the evolution of the temporal z-score

$$z_X^{(t)} = \frac{X^{(t)} - \overline{X}}{s_X}, \tag{2}$$

where $\overline{X} = \sum_t \frac{X^t}{T}$ is the mean over a sample of values covering the period T before time t (in our case, the year before t) and $s_X = \sqrt{\overline{X^2} - \overline{X}^2}$ is the corresponding standard deviation. For example, the choice $X = \sigma[k^{\text{out}}]$ allows price drawdowns to be revealed and, in some cases, anticipated [26]; in the 3-years period 2010–2012, as well as after 2017, the price rises as $z^{(t)}_{\sigma[k^{\text{out}}]}$ increases, whereas drawdowns occur in periods during which

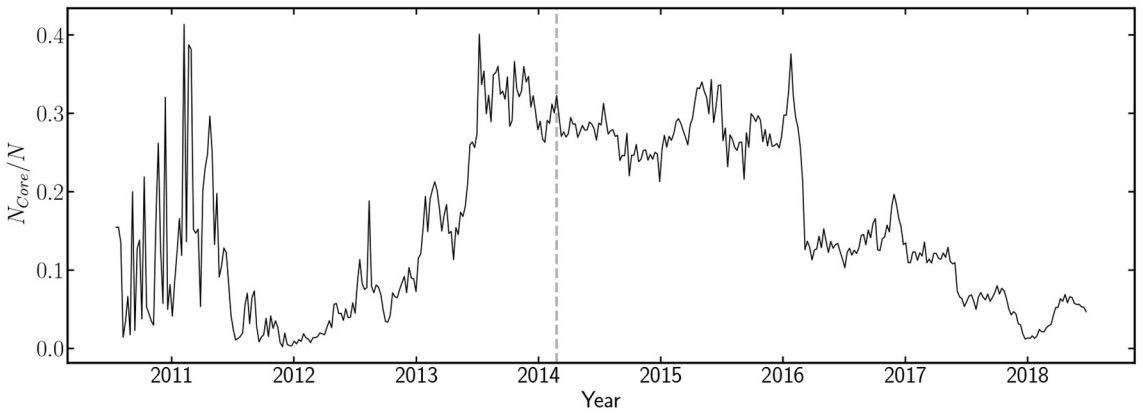

FIGURE 3 | Evolution of the percentage of nodes belonging to the core portion of the BUN on the weekly time scale. During 2012–2013 the core portion of the BUN steadily rises until it reaches $\simeq 30\%$ of the network; then, during 2014–2015, it remains fairly constant; during the last 2 years covered by our data set (2016–2018), the core portion of the BUN shrinks and the percentage of nodes belonging to it falls back to pre-2012 levels. The vertical dashed line coincides with the bankruptcy of Mt. Gox.

$z^{(t)}_{\sigma[k^{\text{out}}]}$ decreases. Other possible choices are $X = N_{\text{core}}$, the number of core nodes, and $X = r$, the network reciprocity, defined as $r = \frac{\sum_i \sum_{j(\neq i)} a_{ij}a_{ji}}{\sum_i \sum_{j(\neq i)} a_{ij}}$, i.e., the percentage of links having a "partner" pointing in the opposite direction. **Figure 4** plots the evolution of the temporal z-scores for N_{core} and r. Overall, the two trends show some similarities, being characterized by peaks that correspond to so-called *bubbles*, i.e., periods of "unsustainable" price growth [36]; interestingly, such periods are characterized by values of the inspected topological quantities which are significant also in a statistical sense, as demonstrated by the values of the corresponding temporal z-scores (in fact, $z^{(t)} \geq 2$ in both cases). Moreover, peaks are also revealed in 2014–2016, thus signaling some kind of "activity" missed by purely financial indicators (e.g., the RPMA).

3.2. The Bitcoin Lightning Network

Let us now move on to results concerning the BLN. In what follows we will focus on the daily-block snapshot representation.

3.2.1. Basic Statistics

As observed for the BAN and the BUN, both the number of nodes N and the number of links $L = \sum_i \sum_{j(>i)} b_{ij}$ of the BLN increase steadily over time, while the network becomes sparser. Interestingly, however, the evolution of the BLN link density seems to point to the presence of two regimes. As **Figure 5** shows, during the first phase (i.e., $N \leq 10^3$), L increases linearly in N and the link density is well-described by the functional dependence $d \sim N^{-1}$; afterwards, the decrease in link density slows down and seems to indicate that L has started to grow in a super-linear fashion with respect to N.

3.2.2. Analysis of the BLN Mesoscale Structure

Although blockchain-based systems are designed to eliminate the need for a central authority to check the validity of exchanges between nodes (i.e., transactions in the case of cryptocurrencies) and authorize them, it is shown in Lin et al. [28] that centralization may still be recovered at a purely structural level. More precisely, the authors considered two sets of quantities. First, they computed the Gini coefficient

$$G_c = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} |c_i - c_j|}{2N \sum_{i=1}^{N} c_i} \tag{3}$$

for four centrality measures, the *degree*, *closeness*, *betweenness*, and *eigenvector centrality* (respectively denoted by the symbols $c_i = k^c_i$, c^c_i, b^c_i, and e^c_i; see also **Appendix B** and [37]), and plotted it against the number of nodes. As shown in Lin et al. [28], G_c increases for three out of the four measures, namely the degree, betweenness, and eigenvector centrality, but not for the closeness centrality, whose trend remains basically flat. Since the Gini coefficient quantifies the (un)evenness of a distribution, this result suggests that the centrality of the nodes is becoming more and more unevenly distributed. A concrete example is provided by the value G_{k^c} reaching $\simeq$ 0.8 in the last snapshot of our data set; this value is compatible with the picture of a network where 90% of connections are incident to 10% of the nodes. In other words, nodes exist that play the role of *hubs*, i.e., vertices with a large number of connections, which are crossed by a large percentage of paths and are connected to other well-connected nodes.

Additionally, Lin et al. [28] computed the so-called *centralization indices*, which encode comparisons between the structure of a given network and that of a reference network, i.e., the "most centralized" structure; see also **Appendix B**. For the degree, closeness, and betweenness centrality measures, it is the *star graph*; for the eigenvector index, the star graph does not represent the maximally centralized structure, although it is retained for consistency with the other quantities. The evolution of the centralization

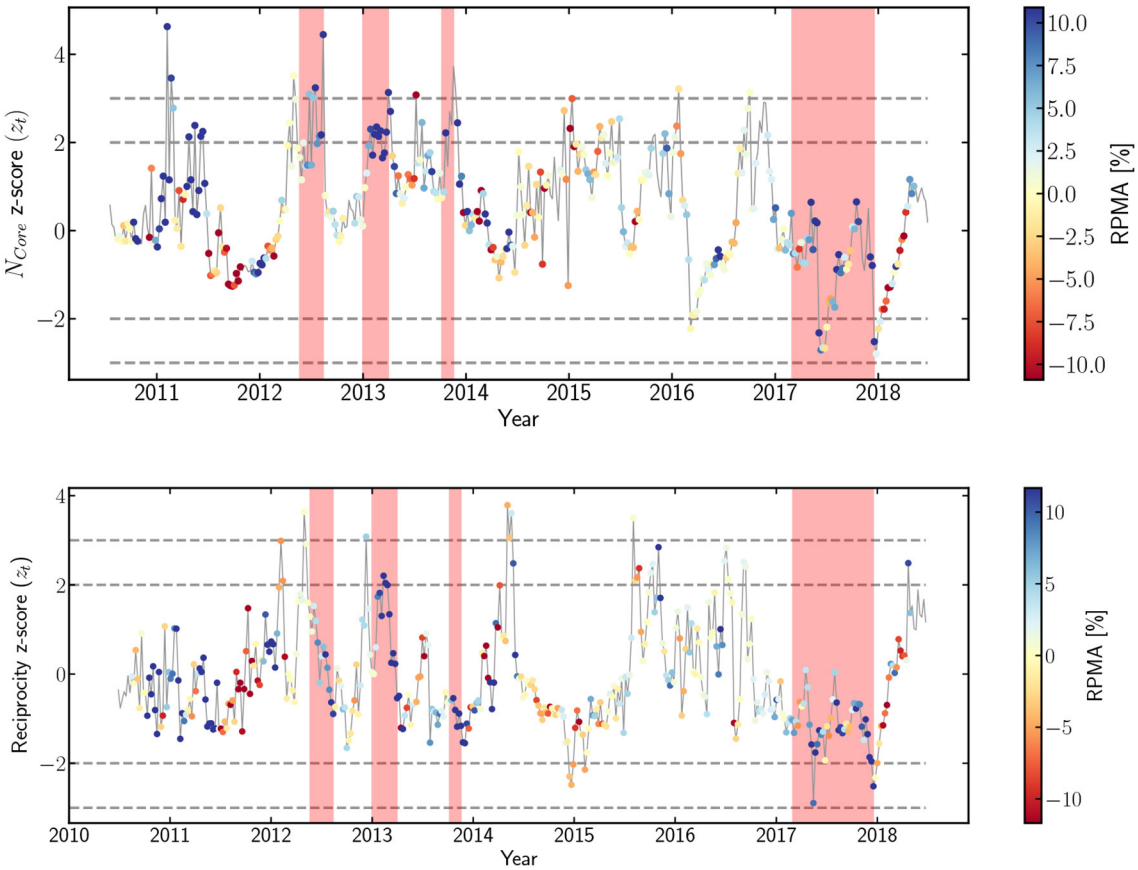

FIGURE 4 | Evolution of the temporal *z*-score for the number of core nodes (top panel) and the reciprocity (bottom panel) for the BUN weekly representation. Shaded areas indicate so-called *bubbles*, i.e., periods of price increase according to Wheatley et al. [36]. Additionally, each dot representing an observation is colored according to the value of the ratio between the current price and its moving average (RPMA) indicator. Overall, the two trends show some similarities, with peaks clearly visible in correspondence to so-called *bubbles*, identified by the shaded areas (see also Wheatley et al. [36]). Interestingly, these quantities are also significant in a statistical sense, as the temporal *z*-scores reach values of $z^{(t)} \geq 2$.

indices indicates that the BLN is evolving not toward a star graph (which is indeed a too-simplistic picture to faithfully describe the BLN topology) but toward a suitable generalization of it, i.e., the core-periphery structure [28] (see also later). Incidentally, the presence of a core-periphery structure is compatible with the aforementioned even distribution of the closeness centrality, since by definition the closeness of a core node does not differ much from the closeness of a periphery node.

In Lin et al. [28] the observations concerning the evolution of the centrality measures and the centralization indices were also benchmarked against the predictions for the same quantities output by the maximum-entropy null model known as the *undirected binary configuration model* (UBCM; see also **Appendix C**). To this end, the authors explicitly sampled the ensembles of networks induced by the UBCM [38, 39] and compared the ensemble average of each quantity of interest with the corresponding empirical value. For technical reasons, the authors adopted an iterative, reduced algorithm to solve the system of equations defining the UBCM,

$$k_i(\mathbf{A}) = \sum_{j(\neq i)=1}^{N} \frac{x_i x_j}{1 + x_i x_j}, \ \forall\, i \quad \Longrightarrow$$
$$x_k^{(n)} = \frac{k(\mathbf{A})}{\sum_{k'} [f(k') - \delta_{kk'}] \left(\frac{x_{k'}^{(n-1)}}{1 + x_k^{(n-1)} x_{k'}^{(n-1)}} \right)}, \ \forall\, k \quad (4)$$

which enabled them to solve the system of equations within tens of seconds even for configurations with millions of nodes [27]; see also **Appendix C**. As **Figure 6** shows, this comparison reveals that the UBCM tends to overestimate the values of the Gini index for the degree, the closeness, and the betweenness centrality measures and to underestimate the values for the eigenvector centrality. This seems to point to a non-trivial (i.e., not reproducible by just enforcing the degrees) tendency of well-connected nodes to establish

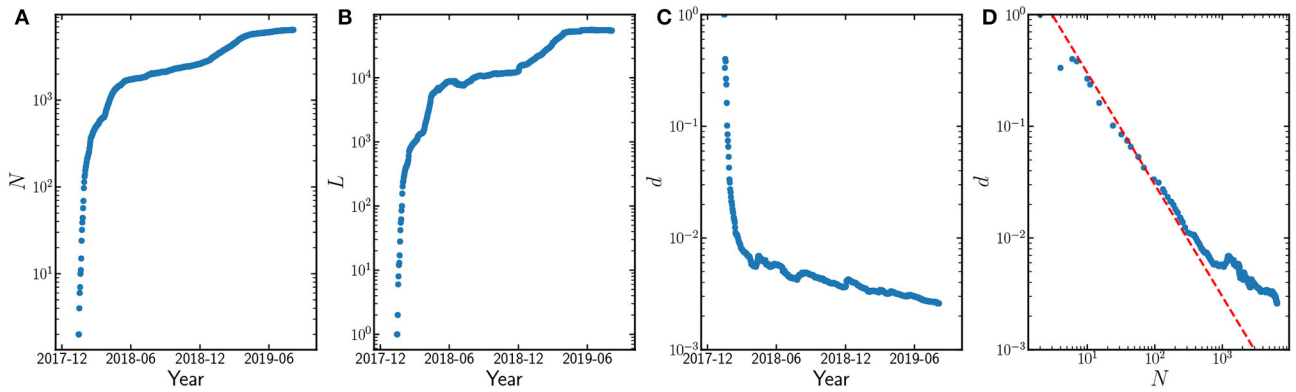

FIGURE 5 | Evolution of the total number of nodes N **(A)**, total number of links L **(B)** and link density $d = \frac{2L}{N(N-1)}$ **(C)** for the BLN daily-block snapshot representation. As for the BAN and the BUN, the position $d \sim N^{-1}$ well describes the link density dependence on N **(D)**, at least for the snapshots for which $N \leq 10^3$. See also Lin et al. [28].

connections among themselves—likely with nodes having a smaller degree associated with them. Such a disassortative structure could explain the less-than-expected level of unevenness characterizing the other centrality measures, as the nodes behaving as "leaves" of the hubs would basically have the same values of degree, closeness, and betweenness centrality.

For the analysis of the centralization indices, **Figure 6** shows that the UBCM underestimates both the betweenness and the eigenvector centralization indices; in other words, a tendency toward centralization "survives" even after the information encoded in the degrees is properly accounted for, letting the picture of a network characterized by some kind of more-than-expected "star-likeness" emerge. This observation can be better formalized by analyzing the BLN mesoscale structure via optimization of the surprise; as observed for the BUN, a core-periphery structural organization, whose statistical significance increases over time, indeed emerges [28] (see also **Figure 7**).

In Bovet et al. [27], the authors also adapted the iterative, reduced algorithm cited above for the resolution of the *directed binary configuration model* (DBCM; see also **Appendix C**).

3.2.3. A Quick Look at the Weighted Structure of the BLN

A quick look at the weighted structure of the BLN yields two notable observations: both the *total amount of exchanged bitcoins* and the *unevenness of their distribution* increase. This trend is confirmed by the evolution the Gini coefficient, whose value reaches 0.9 for the last snapshots of our data set. On average, over the entire period, about 10% (50%) of the nodes hold 80% (99%) of the bitcoins at stake in the network [28].

4. DISCUSSION

The public availability of the complete history of Bitcoin transactions allows researchers to analyze the *structure* characterizing different transaction networks, to inspect the inter-dependency between the network dynamics and the Bitcoin *price*, and to gain insight into the *behavior* of Bitcoin users. Still, understanding of the mechanisms underlying the joint evolution of these three entities remains far from complete.

This paper provides an overview of the most recent results on the topic. One of the main messages concerns the possibility of retrieving signals of exogenous events by analyzing the blockchain-induced transaction networks; the best example is provided by the failure of Mt. Gox in 2014, an event that strongly affected the structure of both the Bitcoin Address Network and the Bitcoin User Network. From this point of view, out-degrees have been found to represent particularly informative properties: higher moments of the out-degrees distribution (such as the standard deviation, skewness, and kurtosis) diverge as the network size becomes larger than the observed value corresponding to the Mt. Gox failure; moreover, the out-degrees heterogeneity rises during periods of price decline (and vice versa).

This result has been further refined by a Granger causality analysis, revealing that during the years 2010–2012 an increase in the out-degrees standard deviation *caused* a price decline [26]. This finding, in turn, suggests a sort of behavioral explanation for the price dynamics exhibited in the early stages of Bitcoin: during periods in which the price continuously increased, more traders were attracted to the system; the later-joining ones, likely performing only a few transactions, linked to the network hubs (usually exchange markets), which gained a large number of connections over these weeks, thus causing the price to rise even further.

Interestingly, analysis of the Bitcoin Lightning Network reveals the same trends as observed for the BAN and BUN, namely the emergence of an uneven distribution of the centrality and the wealth of nodes and of a statistically significant core-periphery structure. These results suggest a tendency of the Bitcoin "Layer 2" network to become less distributed, a process having the undesirable consequence of making this off-chain payment network less resilient to random failures, malicious attacks, etc. The emergence of hubs may be a consequence of

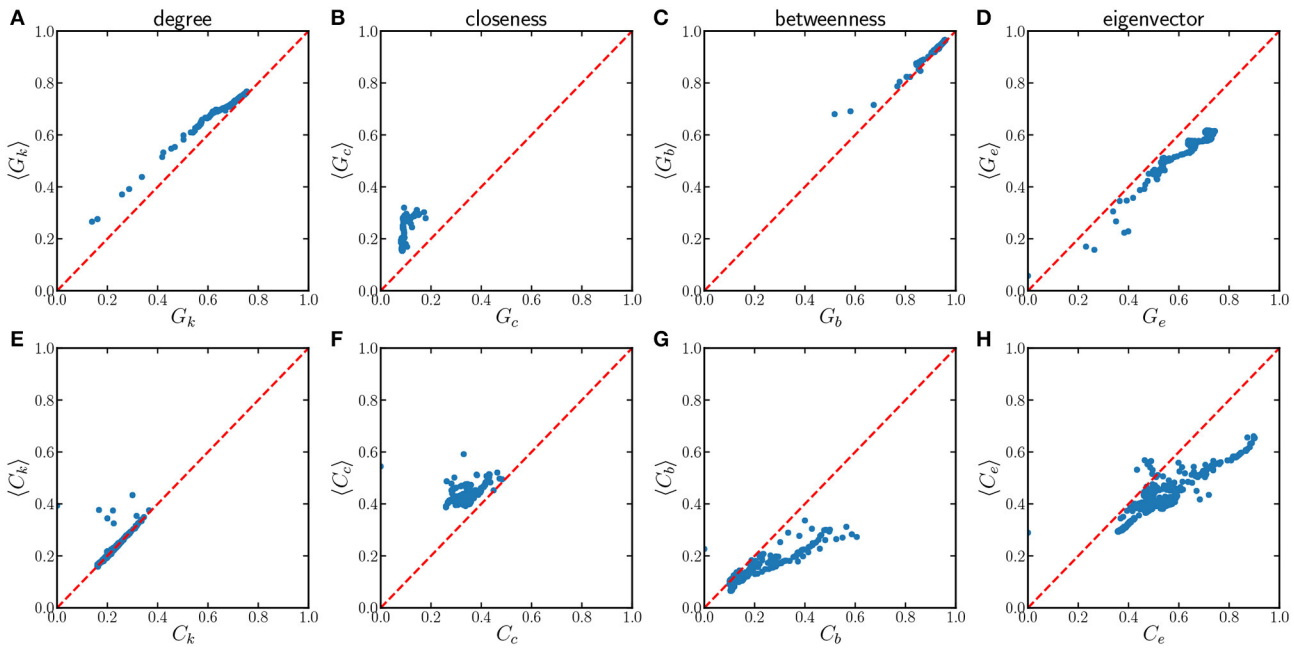

FIGURE 6 | Top panels: comparison between the observed Gini index for the degree **(A)**, closeness **(B)**, betweenness **(C)** and eigenvector centrality **(D)** on the x-axis and their expected value, computed under the UBCM (y-axis) for the BLN daily-block snapshot representation. Bottom panels: comparison between the observed degree- **(E)**, closeness- **(F)**, betweenness- **(G)** and eigenvector- **(H)** centralisation measures and their expected value computed under the UBCM. Once the information contained into the degree sequence is properly accounted for, a (residual) tendency to centralisation is still visible, letting the picture of a network characterised by some kind of more-than-expected 'star-likeness' emerge. See also [28].

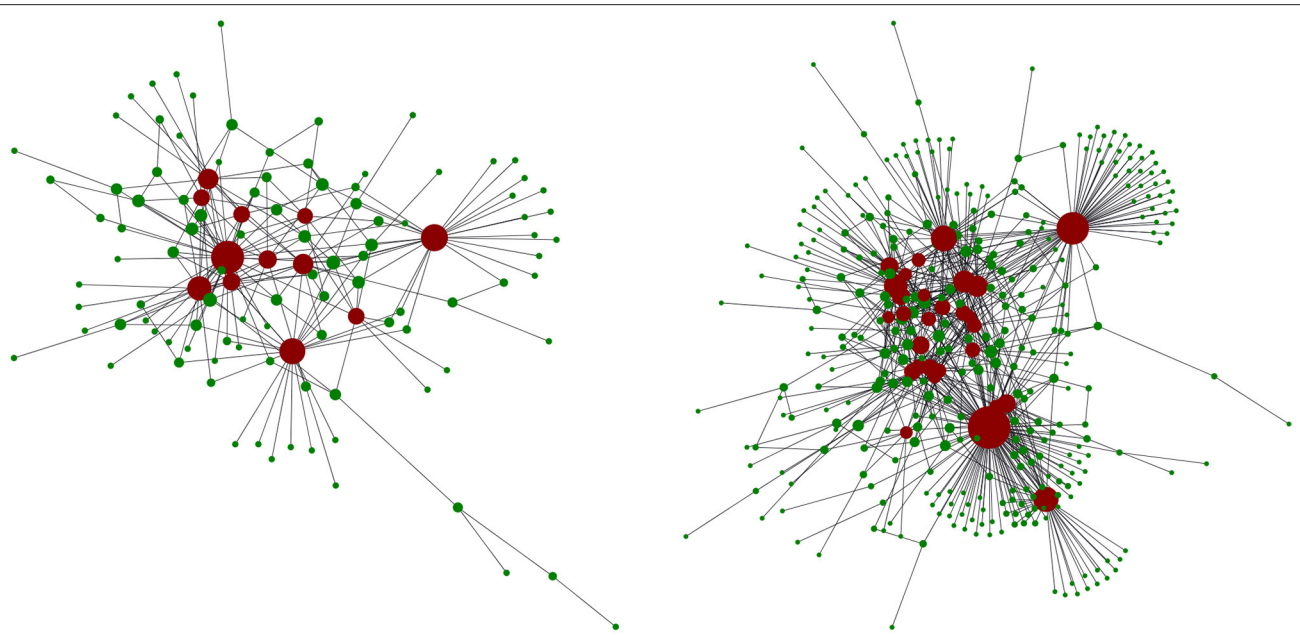

FIGURE 7 | Core-periphery structure of the BLN daily-block representation on day 17 (left panel) and on day 35 (right panel), with core nodes shown in red and periphery nodes in green. Adapted from Lin et al. [28].

the way the BLN is designed: as a route through the network must be found and longer routes are more expensive (fees are charged for the gateway service provided by intermediate nodes), any two BLN users will search for a short(est) path; at the same time, nodes have the incentive to become as central as possible, in order to maximize the transaction fees they can earn. Hubs

may thus have emerged as a consequence of the collective action of users exhibiting one of the two aforementioned behaviors—not surprisingly, since the very beginning of the BLN's history. Regarding the interconnectedness of hubs, previous results have shown that mechanisms aimed at maximizing the centrality of agents give rise to a core-periphery structure (regardless of the notion of centrality the agents attempt to maximize) [40, 41]. As a final remark, we also note that the presence of "centrality hubs" seems to be the source of another peculiarity of the BLN structure, i.e., its small-worldness (a feature already revealed by previous studies [42]).

The results reviewed in this article ultimately—and consistently—point to a tendency toward centralization, which has been observed in the Bitcoin network structure at different levels [28, 43] and is evidence that deserves to be investigated in greater detail. A natural extension of the present work is to analyze the *weighted* counterparts of the three constructs considered here. Of particular interest would be analysis of the weighted centrality measures and centralization indices considered in Lin et al. [28], the outcome of which would help clarify to what extent the observation that binary and weighted quantities are usually correlated in financial systems holds true for cryptocurrencies as well. Other promising avenues of research concern the analysis of different cryptocurrencies and other blockchain-based systems, to understand whether the mechanisms shaping the Bitcoin structure are also at work elsewhere.

AUTHOR CONTRIBUTIONS

CJT and TS designed the review structure. NV performed the additional analyses. NV and TS wrote the manuscript. All authors reviewed and approved the manuscript.

ACKNOWLEDGMENTS

CJT acknowledges financial support from the University of Zurich through the University Research Priority Program on Social Networks. TS acknowledges support from the EU project SoBigData-PlusPlus (grant no. 871042). The authors thank A. Bovet, C. Campajola, F. Mottes, V. Restocchi, and J.-H. Lin for useful discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2020.00286/full#supplementary-material

REFERENCES

1. Antonopoulos M. *Mastering Bitcoin: Programming the Open Blockchain*. Sebastopol, CA: O'Reilly Media, Inc. (2017).
2. Nakamoto S. *Bitcoin: A Peer-to-Peer Electronic Cash System*. (2008).
3. Halaburda H, Sarvary M. *Beyond Bitcoin: The Economics of Digital Currencies*. Basingstoke: Palgrave Macmillan (2016).
4. Glaser F. In: *Proceedings of the Hawaii International Conference on System Sciences 2017 (HICSS-50)*. (2017).
5. Decker C, Wattenhofer R. Decision analytics, mobile services, and service science In: *IEEE P2P 2013 Proceedings*. Waikoloa Village, HI: IEEE (2013). p. 1–10.
6. Poon J, Dryja T. *The Bitcoin Lightning Network: Scalable Off-Chain Instant Payments*. Trento. (2006).
7. Hencic A, Gourieroux C. *Econometrics of Risk*. Cham: Springer (2014).
8. Chu J, Nadarajah S, Chan S. Statistical analysis of the exchange rate of bitcoin. *PLoS ONE*. (2015) 10:e0133678. doi: 10.1371/journal.pone.0133678
9. Chan S, Chu J, Nadarajah S, Osterrieder J. A statistical analysis of cryptocurrencies. *J Risk Financ Manag*. (2017) 10:12. doi: 10.3390/jrfm10020012
10. Sapuric S, Kokkinaki A. DC workshop. In: *Business Information Systems Workshops, Lecture Notes in Business Information Processing*. Larnaca: Springer (2014).
11. Briere M, Oosterlinck K, Szafarz A. Virtual currency, tangible return: portfolio diversification with bitcoin. *J Asset Manag*. (2015) 16:365–73. doi: 10.1057/jam.2015.5
12. Kristoufek A. What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis. *PLoS ONE*. (2015) 10:e0123923. doi: 10.1371/journal.pone.0123923
13. Yiying W. Cryptocurrency price analysis with artificial intelligence. In: *Proceedings of the 5th International Conference on Information Management Cryptocurrency Price Analysis With Artificial Intelligence*. Cambridge: IEEE (2019). doi: 10.1109/INFOMAN.2019.8714700
14. Garcia D, Tessone CJ, Mavrodiev P, Perony N. The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. *J R Soc Interface*. (2014) 11:99. doi: 10.1098/rsif.2014.0623
15. El Bahrawy A, Alessandretti L, Baronchelli A. Wikipedia and digital currencies: interplay between collective attention and market performance. *Front Blockchain*. (2019) 2:1–13. doi: 10.2139/ssrn.3346632
16. Androulaki E, Karame GO, Roeschlin M, Scherer T, Capkun S. Electronic payment (Bitcoin). In: *International Conference on Financial Cryptography and Data Security*. Okinawa: Springer (2013). p. 34–51. doi: 10.1007/978-3-642-39884-1_4
17. Harrigan M, Fretter C. Cryptocurrencies and blockchain based systems. In: *Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), 2016 Intl IEEE Conferences*. Toulouse: IEEE (2016). p. 368–73.
18. Meiklejohn S, Pomarole M, Jordan G, Levchenko K, McCoy D, Voelker GM, et al. In: *Proceedings of the 2013 Conference on Internet Measurement Conference*. ACM (2013). p. 127–40.
19. Ober M, Katzenbeisser S, Hamacher K. Structure and anonymity of the Bitcoin transaction graph. *Fut Internet*. (2013) 5:237–50. doi: 10.3390/fi5020237
20. Reid F, Harrigan M. *Security and Privacy in Social Networks*. Springer (2013).
21. Kondor D, Csabai I, Vattay G. Do the rich get richer? An empirical analysis of the Bitcoin transaction network. *PLoS ONE*. (2014) 9:e0086197. doi: 10.1371/journal.pone.0086197
22. Javarone MA, Wright CS. From Bitcoin to Bitcoin cash: a network analysis. *arXiv*. (2018) 1804.02350v2. doi: 10.1145/3211933.3211947
23. Parino F, Beiró MG, Gauvin L. Analysis of the Bitcoin blockchain: socio-economic factors behind the adoption. *Eur Phys J Data Sci*. (2018) 7:38. doi: 10.1140/epjds/s13688-018-0170-8
24. Motamed AP, Bahrak B. Quantitative analysis of cryptocurrencies transaction graph. *Appl Netw Sci*. (2019) 4:131. doi: 10.1007/s41109-019-0249-6
25. Liang J, Li L, Zeng D. Evolutionary dynamics of cryptocurrency transaction networks: an empirical study. *PLoS ONE*. (2018) 13:e0202202. doi: 10.1371/journal.pone.0202202

26. Bovet A, Campajola C, Mottes F, Restocchi V, Vallarano N, Squartini T, et al. The evolving liaisons between the transaction networks of Bitcoin and its price dynamics. *arXiv*. (2019) 1907.03577.
27. Bovet A, Campajola C, Lazo JF, Mottes F, Pozzana I, Restocchi V, et al. Network-based indicators of Bitcoin bubbles. *arXiv*. (2020) 1805.04460.
28. Lin JH, Primicerio K, Squartini T, Decker C, Tessone CJ. Lightning Network: a second path towards centralization of the Bitcoin economy. *arXiv*. (2020) 2002.02819. doi: 10.1088/1367-2630/aba062
29. Tasca P, Hayes A, Liu S. The evolution of the Bitcoin economy: extracting and analyzing the network of payment relationships. *J Risk Finance*. (2018) 19:94–126. doi: 10.1108/JRF-03-2017-0059
30. Ron D, Shamir A. Electronic payment (Bitcoin). In: *International Conference on Financial Cryptography and Data Security*. Okinawa: Springer (2013). p. 6–24. doi: 10.1007/978-3-642-39884-1_2
31. Baumann A, Fabian B, Lischke M. *WEBIST*. Barcelona: SciTePresss (2014).
32. Bauke H. Parameter estimation for power-law distributions by maximum likelihood methods. *Eur Phys J B*. (2007) 58:167–73. doi: 10.1140/epjb/e2007-00219-y
33. Bianconi G, Barabasi S. Bose-einstein condensation in complex networks. *Phys Rev Lett*. (2001) 86:5632. doi: 10.1103/PhysRevLett.86.5632
34. Granger W. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*. (1969) 37:424–38. doi: 10.2307/1912791
35. de Jeude JvL, Caldarelli G, Squartini T. Detecting core-periphery structures by surprise. *Europhys Lett*. (2019) 125:68001. doi: 10.1209/0295-5075/125/68001
36. Wheatley S, Sornette D, Huber T, Reppen M, Gantner RN. *Are Bitcoin Bubbles Predictable? Combining a Generalized Metcalfe's Law and the LPPLS Model*. Zurich: Swiss Finance Institute Research Paper (2018).
37. Wang J, Li C, Xia C. Improved centrality indicators to characterize the nodal spreading capability in complex networks. *Appl Math Comput*. (2018) 334:388–400. doi: 10.1016/j.amc.2018.04.028
38. Park J, Newman ME. Statistical mechanics of networks. *Phys Rev E*. (2004) 70:066117. doi: 10.1103/PhysRevE.70.066117
39. Squartini T, Garlaschelli D. Analytical maximum-likelihood method to detect patterns in real networks. *New J Phys*. (2011) 13:083001. doi: 10.1088/1367-2630/13/8/083001
40. König MD, Tessone CJ, Zenou Y. Nestedness in networks: a theoretical model and some applications. *CEPR Discuss Pap*. (2014) 9:695–752. doi: 10.3982/TE1348
41. König MD, Tessone CJ, Zenou Y. From assortative to dissortative networks: the role of capacity constraints. *Adv Complex Syst*. (2010) 13:483. doi: 10.1142/S0219525910002700
42. Rohrer E, Malliaris J, Tschorsch F. Discharged payment channels: quantifying the lightning network's resilience to topology-based attacks. *arXiv*. (2019) 1904.10253. doi: 10.1109/EuroSPW.2019.00045
43. Gervais A, Karame G, Capkun S, Capkun V. Is Bitcoin a decentralized currency? *IEEE Sec Priv*. (2014) 12:54–60. doi: 10.1109/MSP.2014.49

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.