SCIENTIFIC REPORTS
natureresearch

OPEN

A percolation model for the emergence of the Bitcoin Lightning Network

Silvia Bartolucci[1,4], Fabio Caccioli[2,4,5] & Pierpaolo Vivo[3*]

The Lightning Network is a so-called second-layer technology built on top of the Bitcoin blockchain to provide "off-chain" fast payment channels between users, which means that not all transactions are settled and stored on the main blockchain. In this paper, we model the emergence of the Lightning Network as a (bond) percolation process and we explore how the distributional properties of the volume and size of transactions per user may impact its feasibility. The agents are all able to reciprocally transfer Bitcoins using the main blockchain and also – if economically convenient – to open a channel on the Lightning Network and transact "off chain". We base our approach on fitness-dependent network models: as in real life, a Lightning channel is opened with a probability that depends on the "fitness" of the concurring nodes, which in turn depends on wealth and volume of transactions. The emergence of a connected component is studied numerically and analytically as a function of the parameters, and the phase transition separating regions in the phase space where the Lightning Network is sustainable or not is elucidated. We characterize the phase diagram determining the minimal volume of transactions that would make the Lightning Network sustainable for a given level of fees or, alternatively, the maximal cost the Lightning ecosystem may impose for a given average volume of transactions. The model includes parameters that could be in principle estimated from publicly available data once the evolution of the Lighting Network will have reached a stationary operable state, and is fairly robust against different choices of the distributions of parameters and fitness kernels.

Bitcoin, the pioneering cryptocurrency, has brought about an unprecedented revolution in the payment industry[1]. Despite its traction and success over the last ten years, the original blockchain – the technological infrastructure underlying Bitcoin – suffers from some limitations that may hinder the future growth and adoption of the cryptocurrency. One of the major issues is the *scalability* of the system: the current number of transactions validated via this platform is between 3 and 7 transactions per second, compared for instance to thousands of transactions handled by the Visa circuit[2]. The lack of scalability is mainly caused by constraints on throughput of transactions, with the block size fixed at 1MB, and by the high latency – with a new block created on average only every ten minutes. Those limitations are imposed to safeguard the security of the platform against malicious attacks and are difficult to relax without major changes in the protocol.

The main solutions proposed to address the scalability issue include (i) changes to the main protocol (consensus algorithm, parameters) and (ii) *sidechains* and second-layer solutions (see[3] for a recent technical review). Notable examples of type-(i) solutions include new consensus protocols, which would allow a faster issuance of new blocks among other new features[4]. Sidechains are blockchains "connected" to the main Bitcoin blockchain such that Bitcoins can be transferred bidirectionally between the main and side blockchain[5]. At the same time, sidechains are completely separate ecosystems whose technical features or issues would not be shared with the main blockchain. The Lightning Network (LN), instead, is a so-called second-layer technology built on top of the Bitcoin blockchain to provide "off-chain" fast payment channels between users[6]. By off-chain we mean that not all transactions are settled and stored on the main blockchain. In a nutshell, the idea of a Lightning channel is the following: two parties lock the same amount of money as collateral and open a channel for a certain period of time. During this time, they can then exchange money back and forth through the channel, and only the netted

[1]Department of Finance, Imperial College London Business School South Kensington, SW7 2AZ, London, UK. [2]Department of Computer Science, University College London, 66-72 Gower Street, WC1E 6EA, London, UK. [3]Department of Mathematics, King's College London, Strand WC2R 2LS, London, UK. [4]Centre for Blockchain Technologies, University College London, London, UK. [5]Systemic Risk Centre, London School of Economics and Political Sciences, Houghton Street, WC2A 2AE, London, UK. *email: pierpaolo.vivo@kcl.ac.uk

transaction will be eventually validated and stored on the main blockchain. If one party is malicious and does not correctly update the balance, the other can keep the collateral posted by the malicious party, as a form of insurance. Any two users can open a channel and all other participants can use one or more existing channels to route transactions off-chain upon payment of a fee to channel "owners". The scalability problem could be solved if a sufficient number of channels were opened, implying that the Lightning Network spans across the whole pool of users of the main blockchain.

The Lightning Network topology is, indeed, relevant to understand the resilience of the system to attacks or random failures and its robustness. Measures of the network structure based on empirical data – such as degree distribution, assortativity, shortest paths length – provide an indication of the efficiency of payments' routing and features of the system (i.e. average number of channels per user, clusters and communities, etc.)[7]. Experiments on random or targeted nodes removal from the network give information on the system resilience by monitoring when the original network is broken into multiple isolated clusters[8–10]. In the Lightning Network case, it has been shown that some types of targeted attacks – aimed at consuming, for instance, the channels' liquidity of specific nodes – may yield severe consequences for the resilience of the network in terms of average payment flow and reachability[11].

The topology of the network is in turn driven by users' economic incentives to relay transactions "off-chain". Moreover, as the Lightning fees are set by channels' owners, an important question is how high such fees should be set in order to guarantee profits, while providing at the same time the right incentives for Bitcoin users to participate in the Lightning Network. In a recent work on simple network topologies (i.e. bidirectional channels and star graphs), the authors have estimated the demand for transactions on the main Bitcoin blockchain compared to the LN, the level of LN fees that would cover maintenance costs of the channel and their implication for the overall network security[12]. Indeed, transacting on the Lightning Network might impact the security of the main blockchain network by inducing a decrease in the amount of fees collected by the miners for the validation of blockchain transactions. Moreover, LN's transaction fees have been empirically studied using a traffic simulator[13].

Fees on the main blockchain are used as incentives to miners (i.e. nodes capable of validating transactions and generating new blocks) to contribute to the security of the platform. The blockchain security is associated with the platform's decentralization, hence to the miners' total computing power[14,15]. Normally, users "compete" to set up the minimal fees that would ensure their transaction to be validated within a given timeframe, as miners try to maximize the total amount of fees per block. A strand of the literature has been investigating the Bitcoin fee set-up mechanisms, the miners' incentives and their potential correlation with risks of attacks and manipulation of the transaction history. In[16] the authors use a game-theoretic model to investigate the factors influencing the value of Bitcoin fees, while in[17] they also examine the interplay between fees and security of the platform, theoretically showing that the current fee model may not be sustainable in the long run. Alternative fee mechanisms have also been proposed, for instance based on auction models[16], and compared with the existing one to highlight weaknesses and possible improvements.

The Bitcoin ecosystem has been already extensively investigated using approaches based on complex networks. The transactions network has been studied to understand latency issues and propagation mechanisms in peer-to-peer systems[18] and inefficiencies of the process of permanent inclusion of the transactions on the blockchain[19]. Global and local structural properties of the users' network in Bitcoin have also provided insights on boom-and-bust events[20] and Bitcoin price dynamics[21]. Moreover, data on users' behaviour and spending patterns have been used to understand the global state of the crypto-economy[22] and the drivers of the growth of the network[23]. More generally, our paper taps into the growing literature on quantitative investigations of the cryptocurrencies landscape, including models of pricing and adoption of tokens[24–27], analysis of the market structure[28–34], price prediction based on sentiment and social interactions[35–43], dynamical analysis of informational efficiency[44], and centralization of the Bitcoin economy[45].

In this paper, we investigate under which conditions in terms of blockchain and Lightning fees, average wealth and volume of transactions per user, a Lightning Network that spans a sizeable fraction of Bitcoin users – thus solving the scalability problem – emerges. We model the emergence of the Lightning Network as a (bond) percolation process on a graph, exploring how different conditions may impact its feasibility[46]. In particular, we consider fitness-dependent network models[47–50] where the probability of creating a new edge depends on intrinsic node features collectively denoted *node fitness*. In the LN case, the node fitness will be defined in terms of the node wealth and activity (i.e. volume of transactions). The viability of the Lightning Network will be characterized in terms of the presence (or not) of a giant connected cluster of nodes: a non-fragmented network would, indeed, guarantee a smooth relay of payments and information between users and will incentivize off-chain transactions. Our model depends on parameters that can be all obtained – or at least estimated – from publicly available data, and is fairly robust against different choices of distributions of parameters and fitness kernels.

The paper is organized as follows. In the following section we provide a quick overview of the main Bitcoin blockchain and the main ideas behind LN. In Section "Model setup" we describe our model and provide the relevant theory, which is then applied to two specific wealth distributions (uniform and exponential) in the subsections "Uniform wealth distribution" and "Exponential wealth distribution", respectively. In the "Results" section, we discuss the results of numerical simulations, and we provide some conclusions and outlook in Section "Discussion". The Appendices are devoted to technical aspects of percolation theory on networks and are included to make the paper self-contained.

The Bitcoin blockchain and Lightning Network. In this section, we summarize the main features of the Bitcoin main blockchain and Lightning Network payment layer. The Bitcoin blockchain is a distributed, shared ledger that immutably records transactions among peers in the network[1,14]. Transactions are bundled in blocks and chained together via cryptographic primitives to ensure that any change at any point in the transaction history would invalidate the full record. Transactions are validated for correctness, temporarily stored in memory

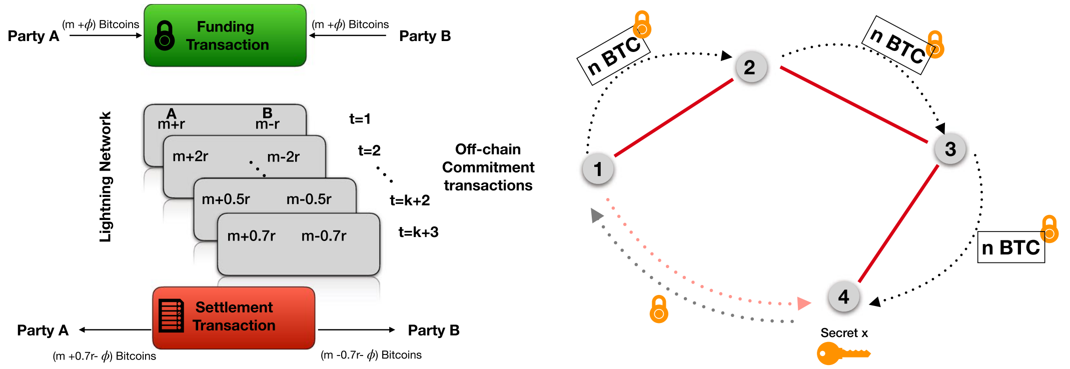

Figure 1. Panel (A): Scheme of a payment channel between party A and B, including opening and closing transactions settled on the main blockchain, and intermediate transfers handled off-chain on the Lightning Network. Initially, the two parties lock m bitcoins each, plus the fee ϕ that would cover broadcasting on the main blockchain. At each time step t, users exchange an amount $\alpha(t)r$ and update their balances accordingly, by sending each other redeemable receipts without committing them on the main blockchain. Only when the two parties agree that the channel is no longer needed, they settle the net balance of funds on their original Bitcoin addresses. Panel (B): Scheme of payments routing on a Lightning Network: even if party 1 and 4 are not directly connected via an existing LN channel, they can route their payments through other parties (upon payment of a fee) by choosing a suitable cryptographic lock for the Bitcoins.

pools and then arranged in the blocks data structure by *miners*: multiple miners compete using computational power to validate the next block of the chain – and therefore earn the associated reward for the service and transactions' fees–according to the Proof-of-Work consensus algorithm. Depending on the usage of the network and due to limitations in block size, waiting times can peak around 30 minutes (while the typical range is around 6–8 minutes), while blockchain fees per transaction exhibit a broad range of variability, from a few cents to 40–50 USD (data taken from https://www.blockchain.com/charts).

The idea behind the creation of the Lightning Network[6] is, therefore, to devise a network for frequent and fast micro-transactions that can be performed at low transactions fees. The basic components of the Lightning Network are *payment channels* (schematically shown in Fig. 1, panel A), enabling trustless transfers between users. In the typical payment channel implementation, a theoretically unlimited amount of payments can be made, with only two transactions broadcast on the blockchain. In addition to a reduction of the number of blockchain transactions and associated costs, payment channels also offer the advantage of speed and, importantly, the ability of users to recover their funds if one of the parties is malicious.

A channel is established between two parties by locking an initial amount of funds, for instance m Bitcoins for each user, on the main blockchain, which represents the maximum amount of Bitcoins that can be transferred over the channel. Funds are locked on so-called 2-of-2 multisignature addresses[14], which can be unlocked upon providing the signature of both interested parties. For instance, user A wishes to send r Bitcoins to user B: she signs a transaction, sends it to B, who will sign it and send it back to A. Only the first transaction is recorded on the main blockchain. At each time step in the lifetime of the channels, the users keep sending back and forth signed transactions that can be at any point consensually broadcast on the main blockchain to close the channel and redeem the net amount of funds. To prevent fraudulent behavior, for instance user B not acknowledging the receipt of a payment from A, a refund option is always included in any exchange. The refund option can be unilaterally unlocked and submitted to the blockchain after a certain amount of time $t^{\star}$ has elapsed from the moment the channel was first established. Every new refund option is indeed signed by both parties, signaling therefore that they are in agreement with the terms of the refund, which may be exercised unilaterally at a later time. In the worst-case scenario, one party would simply submit the original refund transaction created contextually with the opening of the channel.

Payments can be relayed via the Lightning Network also if two parties are not directly connected via a Lightning channel, if there exists a path indirectly linking them via existing channels owned by third parties. Exploiting an existing path to route the payments may often prove more convenient as the two interested parties need not open a new channel, therefore saving the associated costs in blockchain fees. Channels' owners are indeed owed "routing fees" to allow payments through their channel, but at the moment those fees are very competitive (~4 orders of magnitude less than the Bitcoin blockchain, data taken from https://1ml.com/statistics and https://bitcoinfees.info). In Fig. 1, Panel B we show an example of an indirect routing path between user 1 and 4. One of the biggest issues of the Lightning Network is the limitation in liquidity. Payments are made by effectively having intermediaries forwarding collateral across multiple channels: this means that if party 1 is transferring ℓ Bitcoins to party 4, each relaying channel needs to have at least ℓ Bitcoins available in the direction of the payment.

To prevent dishonest behavior in the transfer from party 1 to 4 via party 2 and 3 (see Fig. 1, Panel B), Party 1 will lock the Bitcoins with a secret key known only by the receiver: when party 4 receives the Bitcoins from party 3, the secret is revealed and every player can collect their coins and fees[14].

Model Setup

In this section, we model the emergence of the Lightning Network as a (bond) percolation process. We consider N agents, who are all able to reciprocally transfer Bitcoins using the main blockchain and – if economically convenient – to open a channel on the Lightning Network and transact "off chain". We introduce the node capacity (or *wealth*) w_i of node i, a random variable extracted from a pdf $\Pi(w)$, which is proportional to the maximum amount of Bitcoins that node i can lock in a Lightning channel it partakes in. We will consider two explicit examples for the wealth distribution (uniform and exponential) in the following, with qualitatively similar results.

Two nodes are more likely to open a Lightning channel if they expect to submit a large number of transactions over a given period of time. Therefore, we introduce for each node i a quantity ℓ_i that represents its "activity" in terms of the average number of transactions node i sends through each channel in the network. The average number of transactions is also a random variable extracted from the discrete distribution $\hat{\Pi}(\ell)$ over non-negative integers. We also include the costs associated with transacting over one of the two networks (main blockchain only or blockchain and Lightning). These costs can be fixed per transaction (*base fee*) or can be calculated as a percentage of the value transferred (*fee rate*).

- c, *Lightning channel maintenance/usage base fee:* Using the LN channel provided by an operator or other users to transfer coins carries an associated LN fee. Opening a channel has also maintenance costs (fee setup, market and nodes monitoring, connections) and costs related to locking Bitcoins and providing liquidity in the channel.
- ϕ, *main blockchain fee rate:* We assume that a fraction ϕ of the value transferred in each transaction needs to be paid by the sender to have it included in blocks and validated by miners.

The probability of opening a new LN channel between two nodes p_{ij}^{LN} can be modeled as a function of (i) the costs associated with opening the LN channel (if the costs are significantly smaller than using the Bitcoin blockchain, there is an incentive for the users towards opening the channel), (ii) users' affinity (the more likely are users to transact over a period of time τ, the higher the benefits of opening a channel), (iii) the wealth of the nodes (nodes wishing to open a LN channel have to lock a minimal amount of Bitcoins on the main blockchain as collateral).

The growth of the Lightning Network can be modeled as a bond percolation process on a set of N nodes representing Bitcoin users. The edges then represent new Lightning channels being opened. In particular, we construct the bond percolation model considering fitness-dependent networks[47–50]. In fitness models, the network topology is determined by (i) an *attachment kernel* $f(x, y)$, describing the probability that a node with fitness x will connect to a node with fitness y, and (ii) the distribution of fitness $\rho(x)$ across nodes.

The network we consider has a fixed number of nodes N – corresponding to all Bitcoin users that may decide to switch to the LN – and is *sparse*, i.e. the number M of edges is $M \ll N^2$. If we consider node i and j having fitness x_i and x_j respectively, a LN channel, i.e. an edge between them, is added with probability

$$p_{ij}^{\mathrm{LN}} = f(x_i, x_j) \sim \mathcal{O}(1/N). \tag{1}$$

The resulting network is undirected if $f(x, y) = f(y, x)$, which is a sensible requirement: indeed, opening a LN channel between two nodes will require a "symmetric" commitment from both nodes to lock Bitcoins on the main blockchain. In our model, we will consider bond percolation only: number and "state" of the nodes (e.g. occupied/unoccupied or infected/susceptible) will not change.

In the context of the LN network, we define the fitness x_i of node i as the simplest increasing function of both capacity and volume of transactions, i.e.

$$x_i = w_i(\ell_i + 1), \tag{2}$$

where w_i represents the wealth of the node, and $\ell_i + 1 \geq 1$ its "activity" in terms of number of transactions expected to be sent through the channel. Note that we assume that all nodes are potentially active, $x_i > 0$ for all i. As in[48], we consider the fitness to be defined in the interval $[0, \infty)$.

Given this definition of the node fitness, the fitness distribution can be calculated from the wealth and activity distributions, $\Pi(w)$ and $\hat{\Pi}(\ell)$ respectively, as

$$\rho(x) = \sum_{\ell \geq 0} \hat{\Pi}(\ell) \int_0^\infty \mathrm{d}w \Pi(w) \delta(x - w(\ell + 1)). \tag{3}$$

If we imagine links are added one at a time at a given rate, from the kernel $f(x, y)$ we can derive the probability that a node with fitness x increases its degree by one as[47]

$$\lambda(x, N) = \frac{1}{N} \frac{\int_0^\infty \mathrm{d}y f(x, y) \rho(y)}{\kappa}, \tag{4}$$

where $N\kappa \sim \mathcal{O}(1)$ is the average degree of the network with N nodes, with

$$\kappa = \int_0^\infty \int_0^\infty \mathrm{d}x \mathrm{d}y \rho(x) \rho(y) f(x, y). \tag{5}$$

We also define $\lambda(x) = N\lambda(x, N)$ and rewrite it as

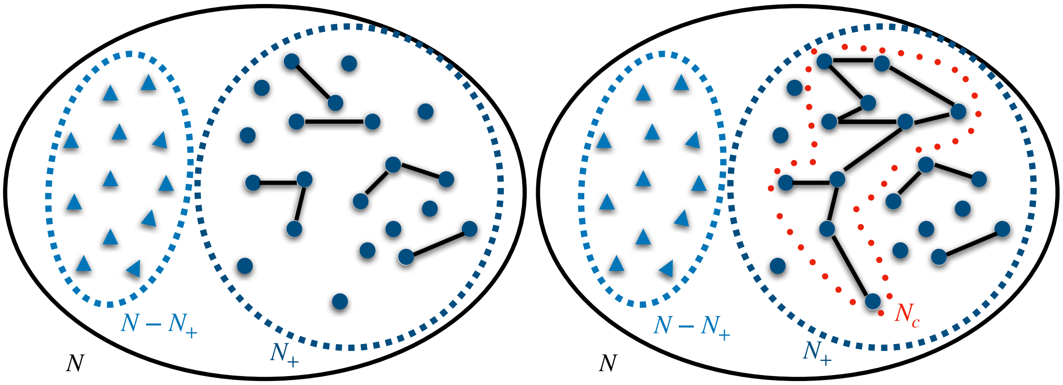

Figure 2. Schematic representation of the emergence of the connected component among fit nodes, below (left) and above (right) the percolation threshold. Light blue triangles represent nodes whose fitness is smaller than c/ϕ, while dark blue circles represent "high-fitness" nodes with $x > c/\phi$. N is the total number of Bitcoin users, N_+ is the fraction of "high-fitness nodes" (see Eq. (10)) and N_C is the fraction of high-fitness nodes belonging to the giant connected component.

$$\lambda(x) = \frac{1}{\kappa}\int_0^\infty \mathrm{d}y f(x, y)\rho(y). \tag{6}$$

Note that $\lambda(x)$ clearly satisfies the following normalization condition

$$\int_0^\infty \lambda(x)\rho(x)\mathrm{d}x = 1. \tag{7}$$

The degree distribution $P(k)$ for large N is given by

$$P(k) = \int_0^\infty \mathrm{d}x\rho(x)\frac{\mathrm{e}^{-N\kappa\lambda(x)}[N\kappa\lambda(x)]^k}{k!}, \tag{8}$$

whose average degree is $\langle k\rangle = N\kappa$, as shown in detail in Appendix A.

In the following, we will assume that the activity distribution is Poisson with average $\bar{n}$, $\hat{\Pi}(\ell) = \exp(-\bar{n})\bar{n}^{\ell}/\ell!$, and that the connectivity kernel models the effects of blockchain and LN fees as follows

$$f(x, y) = \frac{\mu}{N}\Theta(x\phi - c)\Theta(y\phi - c), \tag{9}$$

where $\Theta(z)$ is the Heaviside step function. The interpretation of this kernel is as follows: agent *i* expects to interact with μ other agents, which we assume for simplicity to be chosen randomly.

The probability of interacting with a given agent *j* is equal to μ/N for all *j*. Agent *i* wishes to transfer an amount $w_i(\ell_i + 1)$ (corresponding to $\ell_i + 1$ transactions of size w_i) to each of them, and is willing to open a Lightning channel if the cost of maintaining it (c) is lower than the cost of transferring the money through the blockchain ($w_i(\ell_i + 1)\phi$). The same considerations apply to its counterpart *j*.

We define

$$N_+ = \sum_{i=1}^{N}\Theta(x_i\phi - c) \tag{10}$$

the number of nodes with "high" fitness, for whom it is economically viable to engage in a LN. Note that N_+ is a random variable, which depends on the realization of the fitnesses. We define the average fraction $f_+ = \langle N_+\rangle/N$.

The network constructed via the sequential deposition of links (as described above) may undergo a *percolation transition*[46,47,51,52] as a function of f_+, such that – beyond a critical value of f_+ – a giant connected component of N_C nodes emerges, whose fractional average size $S = \langle N_C\rangle/N$ remains finite as $N \to \infty$. We stress that in any fixed instance $N_C \le N_+$, since some high-fitness nodes may still not engage in LN (see Fig. 2). In our language, this connected component represents the set of nodes that not only do exploit Lightning channels to exchange wealth off-chain between nearest neighbors, but may also transfer wealth to any "distant node", routing the transaction via connected paths. It is therefore of paramount importance to understand under which conditions on the average wealth, average volume of transactions, and routing fees, this transition may happen, and what finite fraction of nodes will it involve.

With the choice of the kernel in (9), the topology of the resulting Lightning Network of N_+ nodes is that of an Erdős-Rényi (E-R) graph with average degree equal to $\mu f_+ \sim \mathcal{O}(1)$. At odds with the standard model of E-R graphs, in our case the size of the graph N_+ is itself a random variable, which depends on the parameters of the model. In fact, once f_+ has been obtained, the model can be mapped onto a site percolation problem on random networks, where each node is occupied with probability f_+, and the emergence of a viable Lightning Network corresponds to the emergence of a giant component of occupied nodes[53].

The relevant percolation theory is summarized in Appendix B to make the paper self-contained.

Uniform wealth distribution.

We now take $\Pi(w)$ – the pdf of wealth across nodes – as uniform in the interval $[0, w_0]$. Hence, we have

$$\rho^{(u)}(x) = \sum_{\ell \geq 0} \frac{e^{-\bar{n}}\bar{n}^{\ell}}{\ell!} \int_0^{w_0} \frac{dw}{w_0} \delta(x - w(\ell + 1)) \, , \tag{11}$$

where the superscript $^{(u)}$ refers to uniform wealth distribution. Simplifying we obtain

$$\rho^{(u)}(x) = \frac{1}{w_0} \sum_{\ell \geq \left\lceil \frac{x}{w_0} - 1 \right\rceil} \frac{e^{-\bar{n}}\bar{n}^{\ell}}{\ell!(\ell + 1)} = \frac{1}{w_0 \bar{n}} \left(1 - \frac{\Gamma\left(\left\lceil \frac{x}{w_0} \right\rceil, \bar{n} \right)}{\Gamma\left(\left\lceil \frac{x}{w_0} \right\rceil \right)} \right), \tag{12}$$

where $\Gamma(a, x) = \int_x^{\infty} t^{a-1} e^{-t} dt$, and $\lceil z \rceil$ denotes the smallest integer larger than z. In this case, it follows from (6) and (9) that

$$\lambda^{(u)}(x) = \frac{1}{f_+^{(u)}} \Theta(x\phi - c), \tag{13}$$

where $f_+^{(u)}$ is the average fraction of high-fitness nodes and is given by

$$f_+^{(u)} = \int_{c/\phi}^{\infty} dx \rho^{(u)}(x) = \frac{1}{w_0} \sum_{\ell \geq 0} \frac{e^{-\bar{n}}\bar{n}^{\ell}}{\ell!} \int_0^{w_0} dw \Theta(w(\ell + 1) - c/\phi) \, , \tag{14}$$

which requires $w > c/[(\ell + 1)\phi]$, in turn constraining $c/[(\ell + 1)\phi] \leq w_0 \to \ell \geq \left\lceil \frac{c}{w_0 \phi} - 1 \right\rceil$ (which may also be negative). Therefore

$$f_+^{(u)} = \frac{1}{w_0} \sum_{\ell = \max\left(0, \left\lceil \frac{c}{w_0\phi} - 1 \right\rceil\right)} \frac{e^{-\bar{n}}\bar{n}^{\ell}}{\ell!} \int_{\frac{c}{(\ell+1)\phi}}^{w_0} dw = \Psi(0) - \frac{c}{\phi w_0} \Psi(-1) \, , \tag{15}$$

where

$$\Psi(t) = \sum_{\ell = \max\left(0, \left\lceil \frac{c}{w_0\phi} - 1 \right\rceil\right)} \frac{e^{-\bar{n}}\bar{n}^{\ell}}{\ell!} (\ell + 1)^t . \tag{16}$$

The evaluation of $P^{(u)}(k)$ from (8) requires some care, as $\lambda^{(u)}(x)$ is zero if $x < c/\phi$. Splitting the integration region, we get

$$\begin{aligned} P^{(u)}(k) &= \delta_{k,0} \int_0^{c/\phi} dx \rho^{(u)}(x) + \frac{e^{-\mu f_+^{(u)}} \left(\mu f_+^{(u)} \right)^k}{k!} \int_{c/\phi}^{\infty} dx \rho^{(u)}(x) \quad . \\ &= (1 - f_+^{(u)}) \delta_{k,0} + f_+^{(u)} \frac{e^{-\mu f_+^{(u)}} \left(\mu f_+^{(u)} \right)^k}{k!} \end{aligned} \tag{17}$$

The interpretation of (17) is quite neat: on average, the network contains a fraction $1 - f_+$ of isolated (low-fitness) nodes, and a fraction f_+ of high-fitness nodes that may (or may not) partake in the LN, establishing sparse random connections with an average of μf_+ other high-fitness nodes. Computing now the generating function (35)

$$G_0^{(u)}(s) = 1 - f_+^{(u)} + f_+^{(u)} \exp(\mu f_+^{(u)} (s - 1)), \tag{18}$$

it follows from Eq. (36) that

$$G_1^{(u)}(s) = \frac{G_0^{(u)'}(s)}{G_0^{(u)'}(1)} = \exp(\mu f_+^{(u)} (s - 1)) \, . \tag{19}$$

The general theory (see Appendix B, in particular Eq. (54)) then implies that the equation determining $0 < \xi^\star \leq 1$ is

$$\xi^\star = \exp[\mu f_+^{(u)} (\xi^\star - 1)]. \tag{20}$$

The average size of the giant component thus reads from Eq. (53)

$$S^{(u)} = (1 - \xi^\star) f_+^{(u)} \,, \tag{21}$$

and the condition in Eq. (51) for the giant component to appear is

$$\mu f_+^{(u)} > 1. \tag{22}$$

The interpretation of this condition is fairly obvious: the giant connected component can only arise if "fit" nodes open on average more than one channel with other fit nodes (see Figs. 3 and 4).

Exponential wealth distribution. We now take $\Pi(w)$ to be the exponential pdf with mean w_0. The fitness distribution now becomes

$$\rho^{(e)}(x) = \sum_{\ell \geq 0} \frac{\mathrm{e}^{-\bar{n}} \bar{n}^\ell}{\ell!} \int_0^\infty \frac{\mathrm{d}w}{w_0} \mathrm{e}^{-\frac{w}{w_0}} \delta(x - w(\ell+1)) = \frac{1}{w_0} \sum_{\ell \geq 0} \frac{\mathrm{e}^{-\bar{n}} \bar{n}^\ell}{\ell!(\ell+1)} \mathrm{e}^{-\frac{x}{w_0(\ell+1)}}.$$

As in the uniform wealth case

$$\lambda^{(e)}(x) = \frac{1}{f_+^{(e)}} \Theta(x\phi - c), \tag{23}$$

where this time $f_+^{(e)}$ reads

$$f_+^{(e)} = \int_{c/\phi}^\infty \mathrm{d}x \rho^{(e)}(x) = \frac{1}{w_0} \sum_{\ell \geq 0} \frac{\mathrm{e}^{-\bar{n}} \bar{n}^\ell}{\ell!} \int_0^\infty \mathrm{d}w \mathrm{e}^{-w/w_0} \Theta(w(\ell+1) - c/\phi) \,, \tag{24}$$

which requires $\ell \geq \left\lceil \frac{c}{\phi w} - 1 \right\rceil$. Therefore,

$$f_+^{(e)} = 1 - \frac{1}{w_0} \int_0^\infty \mathrm{d}w \mathrm{e}^{-w/w_0} \frac{\Gamma\left(\left\lfloor \frac{c}{\phi w} \right\rfloor, \bar{n}\right)}{\Gamma\left(\left\lfloor \frac{c}{\phi w} \right\rfloor\right)} \,, \tag{25}$$

where $\lceil z \rceil$ denotes the largest integer smaller than z. As in the uniform-wealth case

$$P^{(e)}(k) = (1 - f_+^{(e)}) \delta_{k,0} + f_+^{(e)} \frac{\mathrm{e}^{-\mu f_+^{(e)}} \left(\mu f_+^{(e)}\right)^k}{k!}. \tag{26}$$

Now, consider the solution $0 < \eta^\star \leq 1$ of

$$\eta^\star = \exp[\mu f_+^{(e)} (\eta^\star - 1)] \,. \tag{27}$$

Then, the average size of the giant component reads

$$S^{(e)} = (1 - \eta^\star) f_+^{(e)} \,, \tag{28}$$

and the condition for the giant component to appear reads $\mu f_+^{(e)} > 1$ (see Fig. 5).

Results

We present numerical simulations on networks of $N = 5 \cdot 10^4$ nodes, generated by sequential deposition of links with probability as in Eq. (1), using the kernel in Eq. (9). In Fig. 3, where we use a uniform distribution of wealth with average $w_0 = 1, 2$, we plot the average size $S^{(u)}$ of the connected component as a function of $\bar{n}$, the average volume of transactions to be deployed on the LN, for varying values of the fees ratio c/ϕ. Fixing a certain average fraction $S^{(u)}$ of nodes – which can reach each other via a connected LN path – and increasing the ratio c/ϕ between the LN and main-blockchain fees, we observe that a larger average volume of LN transactions is required to make the off-chain network financially sustainable. Increasing the average wealth w_0 would push the curves upwards: as more liquidity becomes available across nodes, more and more players may get involved in the LN for the same level of routing fees. In Fig. 4, the average size $S^{(u)}$ of the connected component is plotted instead as a function of $f_+^{(u)}$, the fraction of high-fitness nodes, for different values of μ, showing that the transition value between $S = 0$ and $S > 0$ happens at $1/\mu$ as predicted by the condition in Eq. (22). In Fig. 5, we observe qualitatively the same phenomenon, this time for an exponential distribution of wealth.

To find the size of the largest connected component, we use a breadth-first search algorithm[54]: starting from a source node s, we label it as belonging to cluster #1. We then explore its neighborhood and assign all nodes reachable from s to cluster #1 as well. The algorithm proceeds recursively until either the whole network has been labelled, or no unlabelled nodes can be further reached. In the latter case, we select another random source among the unlabelled nodes, assign it the label #2, and restart the procedure to find another cluster. At the end, all disjoint clusters have been identified, and their size recorded. In our plots, we monitor the size of the largest cluster.

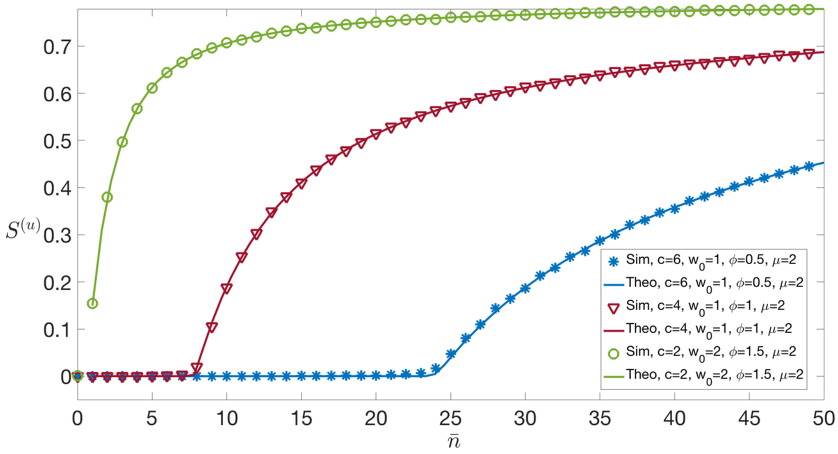

Figure 3. Size of the giant component as a function of $\bar{n}$ for different combinations of the parameters ϕ, c, w_0, μ. Simulations with kernel $f(x, y)$ in Eq. (9) and uniform wealth distribution in the interval $[0, w_0]$. Numerical results (showed with symbols) have been obtained for a network of $N = 5 \cdot 10^4$ nodes averaging over 5 independent instances, whereas the theoretical solid line corresponds to Eq. (21). Fixing a certain fraction S of nodes – connected via LN – and increasing the ratio c/ϕ between the Lightning Network and main blockchain fees, we observe that a larger average volume of LN transactions is required to make the system financially sustainable. The transition between $S = 0$ and $S > 0$ happens at a value of $\bar{n}$ that we denote $\bar{n}^\star$.

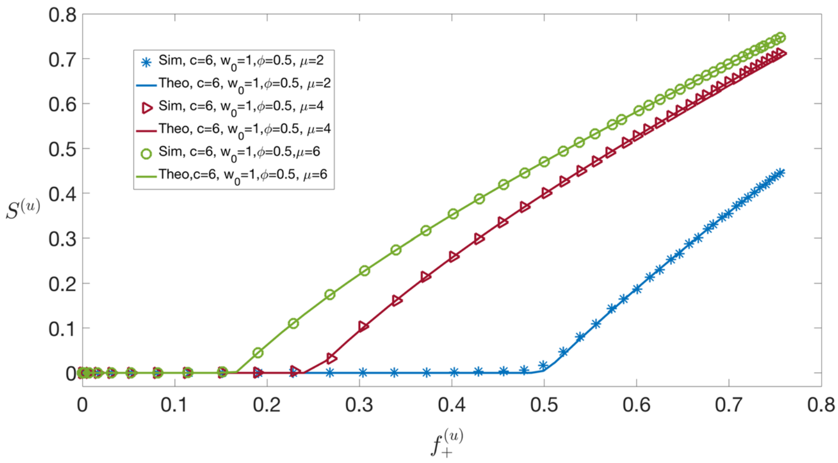

Figure 4. Size of the giant component as a function of $f_+^{(u)}$ (the fraction of high-fitness nodes) varying $\mu = 2, 4, 6$ and fixing $w_0 = 1$, $\phi = 0.5$, $c = 6$. Simulations with kernel $f(x, y)$ in Eq. (9) and uniform wealth distribution with parameter w_0. Numerical results (showed with symbols) have been obtained for a network of $N = 5 \cdot 10^4$ nodes averaging over 5 independent instances, whereas the theoretical solid line corresponds to Eq. (21). The transition between $S = 0$ and $S > 0$ happens at $1/\mu$ as correctly predicted by the condition in Eq. (22).

In Fig. 6, we plot the phase diagram in the (ϕ, c) plane for the uniform wealth distribution model (very similar results are obtained for the exponential wealth distribution, not shown). The colors from blue to yellow represent (from low to high) the values of $\bar{n}^\star$, the minimal average volume of transactions that need to be deployed to make a LN financially viable for a given value of LN and main-blockchain fees, c and ϕ, respectively. We observe a transition between two regimes, signalled by the red line: one (region **1**) where the LN fees are sufficiently low (compared to main-blockchain fees) that *any* volume of transactions (however low, $\bar{n}^\star = 0$ strictly) can be transferred off-chain and still be financially viable, the other (region **2**) where the LN fees are sufficiently high that agents may be discouraged from opening channels and transferring wealth off-chain *unless* there is a minimal volume of transactions to be deployed ($\bar{n}^\star > 0$ strictly). The higher the ratio c/ϕ, the less convenient it is to open LN channels for a fixed value of transactional activity.

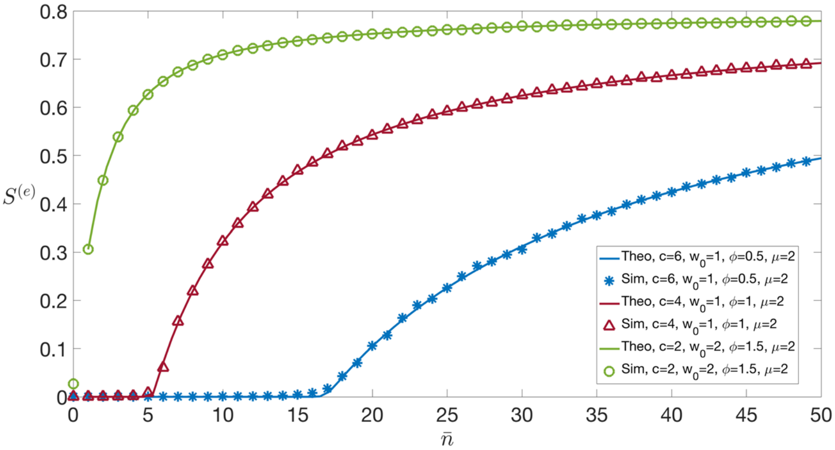

Figure 5. Size of the giant component as a function of $\bar{n}$ for different combinations of the parameters w_0, ϕ, c. Simulations with kernel $f(x, y)$ in Eq. (9) and exponential wealth distribution with average w_0. Numerical results (showed with symbols) have been obtained for a network of $N = 5 \cdot 10^4$ nodes averaging over 5 independent instances, whereas the theoretical solid line corresponds to Eq. (28). Increasing w_0 for similar values of the ratio c/ϕ makes the nodes wealthier on average, and therefore more likely to engage in a LN: as a consequence, the average size of the LN-connected component increases for a given value of average volume of transactions. The transition between $S = 0$ and $S > 0$ happens at a value of $\bar{n}$ that we denote $\bar{n}^\star$.

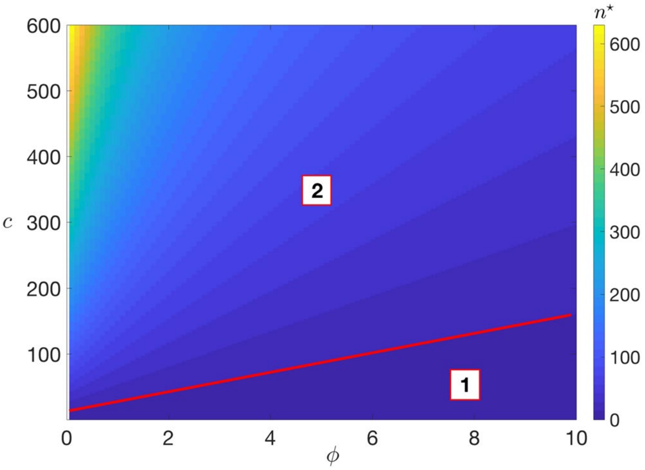

Figure 6. Phase diagram in the (ϕ, c) plane for uniform wealth distribution in the interval $[0, w_0 = 1]$, $\mu = 20$ and kernel $f(x, y)$ (Eq. (9)). The granularity of the axes is as follows: for the ϕ-axis, 100 evenly spaced values, while for the c-axis, 1200 evenly spaced values. The colors from blue to yellow represent (from low to high) the values of the minimal average volume of transactions $\bar{n}^\star$ that need to be deployed to make a LN financially viable. In region (**1**) (low ratio between LN and main-blockchain fees), $\bar{n}^\star = 0$ *strictly*, i.e. for any average volume of transactions, a LN connected component is sustainable. In region (**2**) (larger ratio between Lightning and main blockchain fees), $\bar{n}^\star > 0$ *strictly*, implying that a LN is not financially viable unless the volume of transactions is sufficiently high. The sharp transition between the two regimes (i.e. the line across which $\bar{n}^\star$ jumps from zero to nonzero values) is highlighted in red in the figure. Since only the ratio c/ϕ matters in our setting, the coloring of sub-regions in (**2**) is uniform along straight lines. Note also that the different scale between the axes is due to the different nature of the blockchain fees: c is a *base fee* ("absolute" value), while ϕ is a *fee rate* ("relative" value), as explained in section Model Setup.

Discussion

In summary, we have presented a simple fitness-based network model for the emergence of a connected set of nodes exchanging wealth off-chain, whose average fractional size S remains finite as $N \to \infty$. The percolation transition resulting from sequential deployment of edges is studied numerically and analytically as a function of a limited set of parameters that we predict will be in principle possible to infer from empirical or synthetic[13] data: w_0 (related to the average wealth jointly owned by the agents), $\bar{n}$ (the average volume of transactions that can be handled off-chain), c, ϕ (the fees associated with off-chain and on-chain transactions) and μ (the average number of channels per node). As a matter of fact, different platforms are currently being offered – but only at a test stage – where users can experience the Lighting Network services in a simulated environment. Already at this early stage in the development of a fully operational payment system, some useful data can be gathered: for instance, the platform '1ML' (https://1ml.com/statistics) currently aggregates information about ~10000 nodes sharing ~30000 channels, with an average capacity per node of ~1000 USD, and a base fee per transaction of around 0.000072 USD. Similarly, for the Bitcoin blockchain we can gather an estimate of ~0.52 USD as base fee per transaction, as well as more accurate figures about number of transactions per day and average transaction values (data available at https://www.blockchain.com/en/charts).

The function $f(x, y)$ in Eq. (9) has been selected as the simplest but nontrivial attachment kernel that favors a link (i.e. the opening of a Lightning channel) whenever the fitness of both concurring nodes (in terms of exchangeable wealth and volume of predicted activity) exceeds a financially viable threshold. We have checked that "smoothing" the 0/1- kernel in Eq. (9), e.g. by multiplying the thetas by $xy/(1+xy)$ or $1-\exp(-(x+y))$, has negligible effects on the results, while making the analytical treatment unnecessarily more complicated. Similarly, the model is fairly insensitive to the details of the full probability distribution of wealth that is used (see however[23] for a data-driven analysis of Bitcoin wealth distribution), while being flexible enough to generate a desired degree distribution $P(k)$ via a different choice of the attachment kernel $f(x, y)$[55]. A percolation transition separates a phase where no sustainable LN can be formed, from a phase where the fees being charged, the total available wealth and the average activity conspire to make off-chain payments a viable option for a finite fraction of the network in the limit $N\to\infty$. The transition is elucidated analytically and numerically, with excellent agreement.

In the future, this investigation can be extended in the following ways:

- A mechanism for the dynamical update of wealth as more channels are opened and funds are locked may be introduced to investigate the liquidity constraints of the network in more detail. Dynamically generated wealth inequalities and concentration may be detected by means of centrality measures.
- The resilience of the network can be studied under different types of attacks and compared with available empirical results[7,11].
- Different choices of the kernel $f(x, y)$ (e.g. non-factorized) may be also explored. This could lead to networks with heterogeneous (heavy-tailed) degree distribution, which seems to be in line with recent empirical studies[7].

Once the development of the Lightning Network technology and implementation will have reached maturity, it will be possible to gather data to calibrate our model, which can serve as a driver for policy changes and as guidance for incentive mechanisms design.

Appendix A: Degree distribution $P(k)$

Following[47], the probability $p_{M,N}(k|x)$ that a node in a large undirected graph with N nodes and $M\ll N^2$ edges has degree k given that its fitness is x follows the recursion

$$p_{M+1,N}(k+1|x)=p_{M,N}(k+1|x)[1-2\lambda(x,N)]+2p_{M,N}(k|x)\lambda(x,N). \tag{29}$$

The interpretation is easy: the probability of having a node with degree $k+1$ after an edge addition $(M+1)$ is equal to the probability that the node already had degree $k+1$ times the probability that the new edge does not have any of its two terminal points attached to it ($[1-2\lambda(x, N)]$), plus the probability that the node had degree k times the probability that the new edge has one of its two terminal points connected to it ($2\lambda(x, N)$).

Multiplying both sides of Eq. (29) by s^k and summing over $k\geq 0$, we obtain the following equation for $F_{M,N}(s|x)=\sum_{k\geq 0}s^k p_{M,N}(k|x)$

$$\begin{aligned} F_{M+1,N}(s|x)-F_{M,N}(s|x) &= 2F_{M,N}(s|x)(s-1)\lambda(x,N)+F_{M+1,N}(0|x)\\ &\quad -F_{M,N}(0|x)(1-2\lambda(x,N)). \end{aligned} \tag{30}$$

For large M, Eq. (30) can be rewritten as an ordinary differential equation of the form $\frac{\partial F}{\partial M}=2(s-1)\lambda(x,N)F+\frac{\partial F}{\partial M}+2\lambda(x,N)F\Big|_{s=0}$, with solution

$$F_{M,N}(s|x)=\exp\left[\frac{2M}{N}\lambda(x)(s-1)\right], \tag{31}$$

where we recall that we defined $N\lambda(x)=\lambda(x, N)$ and we used the initial condition $F_{0,N}(s|x)=1$ that follows from the fact that in a network with zero edges, $p_{0,N}(k|x)=\delta_{k,0}$.

Taylor-expanding around $s=0$ and noting that $2M/N=N\kappa$, we obtain the degree distribution conditional on the fitness of the node x

$$p_{M,N}(k|x)=\frac{\mathrm{e}^{-N\kappa\lambda(x)}(N\kappa\lambda(x))^k}{k!}. \tag{32}$$

Marginalizing with respect to x, we eventually obtain the probability that a node has degree k (irrespective of its fitness) as

$$P(k)=\int_0^\infty \mathrm{d}x\, p_{M,N}(k|x)\rho(x)=\int_0^\infty \mathrm{d}x\,\frac{\mathrm{e}^{-N\kappa\lambda(x)}(N\kappa\lambda(x))^k}{k!}\rho(x), \tag{33}$$

which correctly implies

$$\langle k\rangle = \sum_{k\,\geq\,0} kP(k) = N\kappa, \tag{34}$$

using (7).

Appendix B: Giant component

The generating function of the probability that a node has degree k is denoted by

$$G_0(s) = \sum_{k\,\geq\,0} P(k)s^k. \tag{35}$$

We introduce the generating function $G_1(s)$ of the (normalized) probability that by following a randomly chosen edge we reach a node with degree k

$$G_1(s) = \frac{\sum_{k\,\geq\,1} kP(k)s^{k-1}}{\sum_{k\,\geq\,0} kP(k)} = \frac{G_0{}'(s)}{G_0{}'(1)} = \frac{G_0{}'(s)}{N\kappa}. \tag{36}$$

This is because the node we reach by following a randomly chosen edge has degree distribution $kP(k)/N\kappa$ rather than just $P(k)$ – since a randomly chosen edge is more likely to lead to a node of higher degree.

We also define the generating function of the number of nodes that can be reached following a randomly chosen edge and that belong to a connected component of size t with size distribution $\psi(t)$

$$H_1(x) = \sum_{t\,\geq\,1} \psi(t)x^t. \tag{37}$$

Moreover, we indicate with $H_0(x)$ the generating function of the probability that a randomly chosen node belongs to a connected component of size t

$$H_0(x) = \sum_{t\geq 1} Q(t)x^t. \tag{38}$$

More precisely Eq. (37) must be interpreted as

$$H_1(x) = \lim{}_{N\to\infty} \sum_{t=1}^{N} \psi(t,\, N)x^t\ , \tag{39}$$

where $\psi(t, N)$ is the probability that – in a network with N nodes – by following a randomly chosen link, we reach a component of size $t \leq N$, and similarly for $H_0(x)$ in Eq. (38). Assuming that the typical component sizes are finite and that the chances of a component containing a closed loop of edges are negligible for sufficiently large N, the distribution of components generated by $H_1(x)$ can be obtained as follows[47,51,52]. Let us denote by $\zeta(t|k)$ the probability that a node with degree k belongs to a component of size t

$$\zeta(t|k) = \sum_{t_1\,\geq\,1} \cdots \sum_{t_k\,\geq\,1} \delta\left(t - 1,\, \sum_{m=1}^{k} t_m\right) \prod_{m=1}^{k} \psi(t_m)\ , \tag{40}$$

where $\delta(a,\, b)$ is the Kronecker delta. Indeed, the sum of the sizes of the components that can be reached by following the k edges departing from the node must be equal to $t - 1$, and each of these sizes is drawn from the distribution $\psi(t)$.

Marginalizing over the degree distribution, we obtain the probability Q(t) that a randomly chosen node belongs to a component of size t as

$$Q(t) = \sum_{k\geq 0} P(k)\zeta(t|k). \tag{41}$$

Computing $H_0(x)$ from (38)

$$\begin{aligned} H_0(x) &= \sum_{t\geq 1} Q(t)x^t = \sum_{t\geq 1} x^t \sum_{k\geq 0} P(k)\zeta(t|k) \\ &= \sum_{k\geq 0} P(k) \sum_{t\geq 1} x^t \sum_{t_1\geq 1} \cdots \sum_{t_k\geq 1} \delta\left(t - 1,\, \sum_{m=1}^{k} t_m\right) \prod_{m=1}^{k} \psi(t_m) \qquad , \\ &= x \sum_{k\geq 0} P(k) \sum_{t_1\geq 1} \cdots \sum_{t_k\geq 1} x^{\sum_m t_m} \prod_{m=1}^{k} \psi(t_m) = x \sum_{k\geq 0} P(k) \left[\sum_{t\geq 1} \psi(t)x^t\right]^k = xG_0(H_1(x)) \end{aligned} \tag{42}$$

where we have used (35) and (37). The calculation for $H_1(x)$ is analogous, with the replacement $P(k) \to \frac{kP(k)}{\sum_{k'} k'P(k')}$. Summarizing, the two equations to be solved together are

$$H_0(x) = xG_0(H_1(x)), \tag{43}$$

$$H_1(x) = xG_1(H_1(x)). \tag{44}$$

The average size $\langle t\rangle$ of the connected components is given from (38) as

$$\langle t\rangle = \sum_{t\geq 1} tQ(t) = {H_0}'(1). \tag{45}$$

${H_0}'(1)$ can be obtained from (43) as

$${H_0}'(1) = G_0(H_1(1)) + {G_0}'(H_1(1)){H_1}'(1)\ . \tag{46}$$

Note that from (37) it follows that $H_1(1) = 1$ (by normalization of $\psi(t)$). Similarly, from (35), we have that $G_0(1) = 1$ (by normalization of $P(k)$). Eq. (46) can be therefore simplified as follows

$${H_0}'(1) = 1 + {G_0}'(1){H_1}'(1). \tag{47}$$

We can then compute ${H_1}'(1)$ using (44)

$${H_1}'(1) = G_1(H_1(1)) + {G_1}'(H_1(1)){H_1}'(1). \tag{48}$$

As before, we can simplify it using the fact that $H_1(1) = 1$ and that $G_1(1) = \frac{\sum_k kP(k)s^{k-1}}{\sum_k kP(k)}\ |_{s=1} = 1$ (see (36)), obtaining:

$${H_1}'(1) = 1 + G_1^{'}(1){H_1}'(1) \Rightarrow {H_1}'(1) = \frac{1}{1 - {G_1}'(1)}. \tag{49}$$

Substituting (49) in (47) yields

$$\langle t\rangle = {H_0}'(1) = 1 + \frac{{G_0}'(1)}{1 - {G_1}'(1)}, \tag{50}$$

which diverges when

$$1 - {G_1}'(1) = 0 \tag{51}$$

or equivalently (using (36)) when $G_0^{''}(1) = N\kappa$, signalling the emergence of the giant component.

When the giant component has formed, $H_0(x)$ and $H_1(x)$ (see Eq. (37), (38), (39)) become the sum of two contributions: one where the sum is restricted to components of size $t \sim o(N)$, and the other restricted to (giant) components of size $t \sim \mathcal{O}(N)$. Assuming that there is only one such giant component, Eq. (38) for $x = 1$ can then be written as

$$1 = H_0^{(f)}(1) + S, \tag{52}$$

where $H_0^{(f)}(1)$ (and similarly $H_1^{(f)}(1)$) satisfy the equations (43) and (44), as they include $\sim o(N)$ contributions for $N \to \infty$ coming from components other than the giant one, whereas $S = N_C/N$ is the fraction of nodes that belong to the giant component.

Therefore (from (43) and (44))

$$S = 1 - G_0(\xi^\star), \tag{53}$$

where $\xi^\star$ satisfies

$$\xi^\star = G_1(\xi^\star). \tag{54}$$

Data availability

The code used for the simulations and to generate the synthetic data is publicly available and can be found at https://bitbucket.org/sixbit26/lightningnetwork/src/master/.

Received: 7 January 2020; Accepted: 17 February 2020;
Published online: 11 March 2020

References

1. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. (2008).
2. Croman, K. *et al*. On scaling decentralized blockchains. In *International Conference on Financial Cryptography and Data Security*, 106–125 (Springer, 2016).
3. Gudgeon, L., Moreno-Sanchez, P., Roos, S., McCorry, P. & Gervais, A. Sok: Off the chain transactions. *IACR Cryptology ePrint Archive* **2019**, 360 (2019).
4. Mingxiao, D., Xiaofeng, M., Zhe, Z., Xiangwei, W. & Qijun, C. A review on consensus algorithm of blockchain. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2567–2572 (IEEE, 2017).

5. Franco, P. *Understanding Bitcoin* (Wiley Online Library, 2014).
6. Poon, J. & Dryja, T. The Bitcoin lightning network: Scalable off-chain instant payments, https://lightning.network/lightning-network-paper.pdf (2016).
7. Seres, I. A., Gulyás, L., Nagy, D. A. & Burcsi, P. Topological analysis of Bitcoin's lightning network. *arXiv preprint arXiv:1901.04972* (2019).
8. Barrat, A., Barthelemy, M. & Vespignani, A. *Dynamical processes on complex networks* (Cambridge University Press, 2008).
9. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *Nature* **406**, 378 (2000).
10. Cohen, R., Erez, K., Ben-Avraham, D. & Havlin, S. Resilience of the Internet to random breakdowns. *Phys. Rev. Lett.* **85**, 4626 (2000).
11. Rohrer, E., Malliaris, J. & Tschorsch, F. Discharged payment channels: Quantifying the lightning network's resilience to topology-based attacks. *arXiv preprint arXiv:1904.10253* (2019).
12. Brânzei, S., Segal-Halevi, E. & Zohar, A. How to charge lightning. *arXiv preprint arXiv:1712.10222* (2017).
13. Béres, F., Seres, I. A. & Benczúr, A. A. A cryptoeconomic traffic analysis of Bitcoins lightning network. *arXiv preprint arXiv:1911.09432* (2019).
14. Antonopoulos, A. M. *Mastering Bitcoin: unlocking digital cryptocurrencies* (O'Reilly Media, 2014).
15. Orcutt, M. How secure is blockchain really. *MIT Technology Review* (2018).
16. Easley, D., O'Hara, M. & Basu, S. From mining to markets: The evolution of Bitcoin transaction fees. *J. Financial Econ.* **134**, 91–109, https://doi.org/10.1016/j.jfineco.2019.03.004 (2019).
17. Houy, N. The economics of Bitcoin transaction fees. GATE WP 1407, https://doi.org/10.2139/ssrn.2400519 (2014).
18. Decker, C. & Wattenhofer, R.Information propagation in the Bitcoin network. In *IEEE P2P 2013 Proceedings*, 1–10 (IEEE, 2013).
19. Pappalardo, G., Di Matteo, T., Caldarelli, G. & Aste, T. Blockchain inefficiency in the Bitcoin peers network. *EPJ Data Science* **7**, 30 (2018).
20. Bovet, A. *et al.* Network-based indicators of Bitcoin bubbles. *arXiv preprint arXiv:1805.04460* (2018).
21. Bovet, A. *et al.* The evolving liaisons between the transaction networks of Bitcoin and its price dynamics. *arXiv preprint arXiv:1907.03577* (2019).
22. Lischke, M. & Fabian, B. Analyzing the Bitcoin network: The first four years. *Futur. Internet* **8**, 7 (2016).
23. Kondor, D., Pósfai, M., Csabai, I. & Vattay, G. Do the rich get richer? an empirical analysis of the Bitcoin transaction network. *Plos One* **9**, e86197 (2014).
24. Ciaian, P., Rajcaniova, M. & Kancs, d. The economics of Bitcoin price formation. *Appl. Econ.* **48**, 1799–1815 (2016).
25. Cong, L. W., Li, Y. & Wang, N. Tokenomics: Dynamic adoption and valuation. *Columbia Business School Research Paper* (2019).
26. Bartolucci, S. & Kirilenko, A. A model of the optimal selection of crypto assets. *arXiv preprint arXiv:1906.09632* (2019).
27. Alessandretti, L., ElBahrawy, A., Aiello, L. M. & Baronchelli, A. Anticipating cryptocurrency prices using machine learning. *Complexity* **Article ID 8983590** (2018).
28. Drożdż, S., Gębarowski, R., Minati, L., Oświęcimka, P. & Wa˛torek, M. Bitcoin market route to maturity? Evidence from return fluctuations, temporal correlations and multiscaling effects. *Chaos: An Interdiscip. J. Nonlinear Sci.* **28**, 071101, https://doi.org/10.1063/1.5036517 (2018).
29. Drożdż, S., Minati, L., Oświęcimka, P., Stanuszek, M. & Wa˛torek, M. Signatures of crypto-currency market decoupling from the forex. *arXiv preprint arXiv:1906.07834* (2019).
30. Sigaki, H. Y., Perc, M. & Ribeiro, H. V. Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market. *Scientific Reports* **9**, 1440, https://doi.org/10.1038/s41598-018-37773-3 (2019).
31. Urquhart, A. The inefficiency of Bitcoin. *Econ. Lett.* **148**, 80–82 (2016).
32. Alessandretti, L., ElBahrawy, A., Aiello, L. M. & Baronchelli, A. Machine learning the cryptocurrency market. *Available at SSRN 3183792* (2018).
33. ElBahrawy, A., Alessandretti, L., Kandler, A., Pastor-Satorras, R. & Baronchelli, A. Evolutionary dynamics of the cryptocurrency market. *Royal Soc. Open Sci.* **4**, 170623 (2017).
34. Cocco, L., Concas, G. & Marchesi, M. Using an artificial financial market for studying a cryptocurrency market. *J. Econ. Interact. Coord.* **12**, 345–365 (2017).
35. Aste, T. Cryptocurrency market structure: connecting emotions and economics. *Digit. Finance* **1**, 5–21 (2018).
36. Abraham, J., Higdon, D., Nelson, J. & Ibarra, J. Cryptocurrency price prediction using tweet volumes and sentiment analysis. *SMU Data Sci. Rev.* **1**, 1 (2018).
37. Kim, Y. B. *et al.* Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *PloS One* **11**, e0161197, https://doi.org/10.1371/journal.pone.0161197 (2016).
38. Li, T. R., Chamrajnagar, A., Fong, X., Rizik, N. & Fu, F. Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model. *Front. Phys.* **7**, 98, https://doi.org/10.3389/fphy.2019.00098 (2019).
39. Bartolucci, S. *et al.* The butterfly "affect": Impact of development practices on cryptocurrency prices. https://www.researchgate.net/publication/335543115_The_Butterfly_Affect_Impact_of_Development_Practices_on_Cryptocurrency_Prices (2019).
40. Kristoufek, L. Bitcoin meets Google trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Sci. Reports* **3**, 3415, https://doi.org/10.1038/srep03415 (2013).
41. Garcia, D., Tessone, C. J., Mavrodiev, P. & Perony, N. The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. *J. Royal Soc. Interface* **11**, 2014.0623, https://doi.org/10.1098/rsif.2014.0623 (2014).
42. Chen, C. Y.-H. & Hafner, C. M. Sentiment-induced bubbles in the cryptocurrency market. *J. Risk Financial Manag.* **12**, 53 (2019).
43. Yelowitz, A. & Wilson, M. Characteristics of Bitcoin users: an analysis of Google search data. *Appl. Econ. Lett.* **22**, 1030–1036 (2015).
44. Sigaki, H. Y., Perc, M. & Ribeiro, H. V. Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market. *Sci. reports* **9**, 1–9 (2019).
45. Lin, J.-H., Primicerio, K., Squartini, T., Decker, C. & Tessone, C. J. Lightning network: a second path towards centralisation of the bitcoin economy. *ArXiv:2002.02819* (2019).
46. Callaway, D. S., Newman, M. E., Strogatz, S. H. & Watts, D. J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.* **85**, 5468 (2000).
47. Hoppe, K. & Rodgers, G. J. Percolation on fitness-dependent networks with heterogeneous resilience. *Phys. Rev. E* **90**, 012815 (2014).
48. Caldarelli, G., Capocci, A., De Los Rios, P. & Muñoz, M. A. Scale-free networks from varying vertex intrinsic fitness. *Phys. Rev. Lett.* **89**, 258702 (2002).
49. Servedio, V. D., Caldarelli, G. & Butta, P. Vertex intrinsic fitness: How to produce arbitrary scale-free networks. *Phys. Rev. E* **70**, 056126 (2004).
50. Bianconi, G. & Barabási, A.-L. Competition and multiscaling in evolving networks. *EPL (Europhysics Lett.* **54**, 436 (2001).
51. Newman, M. E., Strogatz, S. H. & Watts, D. J. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**, 026118 (2001).
52. Newman, M. E. Component sizes in networks with arbitrary degree distributions. *Phys. Rev. E* **76**, 045101 (2007).
53. Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. Critical phenomena in complex networks. *Rev. Mod. Phys.* **80**, 1275 (2008).
54. Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. *Introduction to algorithms* (MIT press, 2009).
55. Smolyarenko, I. E., Hoppe, K. & Rodgers, G. J. Network growth model with intrinsic vertex fitness. *Phys. Rev. E* **88**, 012805 (2013).

Acknowledgements

S.B. and F.C. acknowledge funding by UCL Centre for Blockchain Technologies as part of the *1st Internal Call for Project Proposals on Distributed Ledger Technologies*. P.V. acknowledges support from the UKRI Future Leaders Fellowship grant MR/S03174X/1. We are grateful to István András Seres for many interesting remarks and for generously sharing data.

Author contributions

S.B., F.C. and P.V. conceived the model, performed numerical simulations and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.