Payment-failure times for random Lightning paths

Taki E.M. Abedesselam
University of Rome "Tor Vergata" and University of Camerino
Rome, Italy
abedesselam@ing.uniroma2.it

Fabio Giacomelli
University of Rome "Tor Vergata" and University of Camerino
Rome, Italy
fabio.giacomelli@uniroma2.it

Francesco Pasquale
University of Rome "Tor Vergata"
Rome, Italy
pasquale@mat.uniroma2.it

Michele Salvi
University of Rome "Tor Vergata"
Rome, Italy
salvi@mat.uniroma2.it

***Abstract*—We study a random process over graphs inspired by the way payments are executed in the Lightning Network, the main layer-two solution on top of Bitcoin. We first prove almost tight upper and lower bounds on the time it takes for a payment failure to occur, as a function of the number of nodes and the edge capacities, when the underlying graph is complete. Then, we show how such a random process is related to the edge-betweenness centrality measure and we prove upper and lower bounds for arbitrary graphs as a function of edge-betweenness and capacity. Finally, we validate our theoretical results by running extensive simulations over some classes of graphs, including snapshots of the real Lightning Network.**

***Index Terms*—Lightning Network, Markov chains, Centrality measures, Bitcoin**

I. Introduction

The Lightning Network [1] is the main layer-two solution on top of Bitcoin [2] that promises to address the scalability issue (see, e.g., [3]) executing *off-chain* the vast majority of transactions, and using the blockchain layer as a notary service to resolve controversies. It consists of a series of cryptographic mechanisms that allow parties to build a *channel graph* and to execute *payments* over paths on such a graph in a trustless way, relying on the security of the underlying blockchain. Roughly speaking, the channel graph is built as follows: Two parties, u and v, can create a *channel* by locking bitcoins in a 2-of-2 multisignature address and recording it on the blockchain, the amount of locked funds is the channel *capacity* and it is known to all the parties in the system; at any point in time, each one of the two endpoints of a channel owns a share of such capacity, those shares are called the channel *balance*; the balance of a channel is known only by the two endpoints and it can be updated by exchanging private messages between them. For example, if u and v opened a channel of capacity 5, and the current balance of the channel is 4 for u and 1 for v, when u wants to pay 1 to v, u and v can agree to update the balance of their channel to 3 for u and 2 for v. The parties in the system and the channels between them constitute nodes and edges, respectively, of the *channel graph* of the Lightning Network. A node u that needs to pay a non-neighbor node w can look for a path $\mathcal{P}$ between u and w on the channel graph and request the execution of the payment on that path; in order for the payment to be successful, the balances of all the channels in path $\mathcal{P}$ have to be updated accordingly, hence a payment can *fail* if this is not possible; in this case, we say that a *payment failure* occurs, and node u has to try another path to pay w. We give a slightly more rigorous description of such a process in Section II and we refer the reader interested in the details of the cryptographic mechanisms that allow the routing of a payment to be *atomic* (either the balance of all the paths is updated or none is) to [4].

A. Our contribution

Motivated by the above scenario, in this paper we study the following random process: Given an undirected graph G, where each edge e has a capacity $c(e)$ and an initial balance equally divided between the two endpoints, at every discrete round a source node u and a destination node v are chosen uniformly at random (u.a.r.) among all the nodes, then a shortest path $\mathcal{P}$ is chosen u.a.r. among all shortest paths between u and v, and a payment of unit value is executed from u to v over path $\mathcal{P}$. Our goal is to investigate how long it takes for a payment failure to occur, depending on the topology of the graph and on the channel capacities.

We first analyze the case where the graph G is a clique where every edge has the same capacity $2k$. We prove that, in this case, the time until the first failure occurs is $\Omega(k^2n^2/\log n)$ and $\mathcal{O}(k^2n^2)$, with high probability (w.h.p.).

Then, we extend the analysis to general graphs and show that the time until the first failure occurs depends on the ratio between the squared capacity of the edges and their *edge-betweenness* [5]. More precisely, in this case we prove that the time until the first failure occurs is $\Omega\left(\xi \frac{n^2}{\log n}\right)$ and $\mathcal{O}\left(\xi\, n^2 \log n\right)$, where ξ is the minimum, over all edges e, of the ratio between the squared capacity k_e^2 and the edge-betweenness $g(e)$.

Finally, we validate our theoretical results through extensive simulations of the random process. We observe that the results of the simulations on the clique graph are consistent with the theoretical bounds, and that the lower bound is likely tight.

To highlight the impact of the correlation between edges, we simulate the process on ring graphs with n nodes and we compare the results with those that we get by considering n independent birth-and-death chains, each representing the evolution of the capacity of a single edge. Additionally, we run simulations on a snapshot of the channel graph of the Lightning Network. Specifically, we investigate how the overall payment failure rate in the network would improve, if it was possible to rearrange the capacity of the edges to maximize ξ.

B. Related work

The Lightning Network, first described in [1], is the source of several research problems related to distributed network formation, and the notion of edge-betweenness centrality [5] is recurrent in most of them. For example, in [6] the authors look at different attachment strategies and show how they affect the network in the long term. One of the strategies involves adding a node using a greedy algorithm that tries to maximize its betweenness centrality. This strategy was proposed in [7], where k edges are added to maximize the node's betweenness. Alternatively, a node can be connected to the nodes with the highest betweenness in the network. Both strategies turned out effective in reducing the network's diameter and increasing the success rate of payments in the Lightning Network. In [8], a greedy algorithm was designed to increase a node's profit by improving its betweenness centrality. In [9], a probabilistic model was proposed to account for uncertainty in channel balance and to support multi-part payments—splitting a payment into parts and sending them instead of sending it as a whole. While the proposed model does not consider the dynamics of channel balances, it allows for specifying service-level objectives and quantifying the amount of private information leaked to the sender as a side effect of payment attempts. A survey on networking problems related to blockchain and cryptocurrencies can be found in [10].

The model we study in this paper is inspired by the Lightning Network, however similar models have been previously studied in the context of *credit networks* [11], [12] as solution concepts for certain types of auctions. What distinguishes the design of the Lightning Network is that it relies on cryptography to establish channels between nodes and to allow parties to execute off-chain transactions in a trustless way. There have also been attempts to understand this type of networks mathematically, e.g., in [13] the authors study the long-term behavior of a similar random process for rings, stars, lines, cliques, Erdős–Rényi, and Barabási–Albert topologies. In [14] the authors provide tools for analyzing credit networks by reformulating transactions from discrete to continuous transaction models.

In our simulations we need to repeatedly sample shortest paths between nodes, u.a.r. among all shortest paths. In [15], several efficient algorithms for uniformly sampling shortest paths between fixed source and target nodes were proposed and analyzed.

C. Organization of the paper

In Section II, we provide the formal description of the random process we study in this paper. In Section III, we analyze the case where the graph is a clique with uniform capacities and derive high-probability bounds on the failure time. In Section IV we extend the analysis to general graphs and show that the failure time depends on the smallest ratio between squared capacity and edge betweenness. In Section V, we present the simulation results on cliques, rings, and a snapshot of the channel graph of the Lightning Network graph. Finally, in Section VI we draw some conclusions.

II. The model and the problem

Consider an edge-weighted undirected graph $G = (V, E)$. In the context of the Lightning Network [1], [4], edges are called *channels* and edge weights $c : E \to \mathbb{N}$ are called *capacities*; throughout this paper we align with that terminology. The capacity $c(e)$ of a channel $e = \{u, v\}$ is "shared" between the two endpoints u, v: Such a share, $\{b_e(u), b_e(v)\}$ where $b_e(u), b_e(v) \in \{0, 1, \ldots, c(e)\}$ and $b_e(u) + b_e(v) = c(e)$, is called the channel *balance*. For convenience sake, for a channel $e = \{u, v\}$ we also define $b_{\min}(e) = \min\{b_e(u), b_e(v)\}$ and we say that a channel e is *empty in one direction* if $b_{\min}(e) = 0$.

A *payment* between two adjacent nodes u and v is an update of the balance of edge $e = \{u, v\}$, e.g., if the current balance between u and v is $\{b_e(u), b_e(v)\}$ and node u executes a payment $\texttt{pay}(u, v, x)$ to node v of an *amount* $x \in \{1, \ldots, b_e(u)\}$, then the new balance $\{b'_e(u), b'_e(v)\}$ of the edge will be $b'_e(u) = b_e(u) - x$ and $b'_e(v) = b_e(v) + x$.

Payments can be *routed* across the network: A node u willing to send an amount x to a non-neighbor node v can look for a path $\mathcal{P} = (u = u_0, u_1, \ldots, u_h = v)$ in the graph between nodes u and v, if it exists, and execute a payment $\texttt{pay}_{\mathcal{P}}(u, v, x)$ to node v over path $\mathcal{P}$ of an amount x, provided that $x \leqslant b_{e_i}(u_i)$ for every channel $e_i = \{u_i, u_{i+1}\}$ in the path. Such a payment will update the balances of all the channels in $\mathcal{P}$ accordingly, i.e., for every channel $e_i = \{u_i, u_{i+1}\}$ in the path the new balance $\left\{b'_{e_i}(u_i), b'_{e_i}(u_{i+1})\right\}$ after the payment will be[1]

$$\begin{cases} b'_{e_i}(u_i) &= b_{e_i}(u_i) - x \\ b'_{e_i}(u_{i+1}) &= b_{e_i}(u_{i+1}) + x \end{cases}$$

While the graph $G = (V, E)$ and the channel capacities $c : E \to \mathbb{N}$ are known to all the nodes in the network, the balance $\{b_e(u), b_e(v)\}$ of a channel $e = \{u, v\}$ is known only by the endpoints u and v. Hence, it is possible (and it is often the case in the real Lightning Network) that when a node u tries to pay an amount x to a non-neighbor node v using a

[1] In the real Lightning Network, for each intermediate channel there will be a small *fee* that the node forwarding the payment will subtract to the amount, hence if v needs to receive an amount x then node u has to send a larger amount, say $x + \varepsilon$, where ε is the total fee that will be subtracted by the intermediate nodes, $u_1, \ldots, u_{k-1}$ in the path. In order to keep our model simple and to focus it on the main graph theoretic problem that we want to address, we here ignore such a detail (as well as several other details) of the actual payment routing process in the Lightning network.

path $\mathcal{P} = (u = u_0, u_1, \dots, u_k = v)$, the payment cannot be executed, due to the fact that for some channel in the path the balance is not sufficient to accommodate the payment; this happens when there is a channel $e_i = \{u_i, u_{i+1}\}$ in the path with balance $\{b_{e_i}(u_i), b_{e_i}(u_{i+1})\}$ such that $b_{e_i}(u_i) < x$. In such a case we say that a *payment failure* occurs, and node u has to choose another path to pay node v.

a) The random process.: To formulate a concrete mathematical problem and theoretically analyze the impact of network topology and channel capacities on the rate of payment failures, we here consider the following discrete-time random process. Given an undirected connected graph $G = (V, E)$, where all channels have the same capacity that is initially perfectly balanced between the two endpoints, i.e., for each channel $e = \{u, v\} \in E$, we have $c(e) = 2k$ for some $k \in \mathbb{N}$ and we start with $b_e(u) = b_e(v) = k$. At every round $t = 1, 2, \dots$, we pick a source node $u \in V$ uniformly at random (u.a.r.), a destination node $v \in V$ u.a.r., and a shortest path $\mathcal{P}$ between u and v, u.a.r. among all the shortest paths between u and v, and we execute a payment of amount $x = 1$ over path $\mathcal{P}$ from u to v. We are interested in the expected number of rounds before we have that a channel $e = \{u, v\}$ exists such that $b_e(u) = 0$ or $b_e(v) = 0$.

Algorithm 1 The random process

Require: An undirected connected graph $G = (V, E)$;
Edge capacities $c(e) = 2k$ for every $e \in E$;
Initial balance $b_e(u) = b_e(v) = k$, for every edge $e = \{u, v\}$.

1: **while** $b_{\min}(e) > 0$ for every $e \in E$ **do**
2: Pick a source node u in V, u.a.r.
3: Pick a destination node v in V, u.a.r.
4: Pick a shortest path $\mathcal{P} = (u = u_0, u_1, \dots, u_h = v)$, u.a.r. among all shortest paths between u and v;
5: **for** $i = 0, \dots, h-1$ **do**
6: $b_{\{u_i, u_{i+1}\}}(u_i) = b_{\{u_i, u_{i+1}\}}(u_i) - 1$
7: $b_{\{u_i, u_{i+1}\}}(u_{i+1}) = b_{\{u_i, u_{i+1}\}}(u_{i+1}) + 1$
8: **return** The number of iterations of the while cycle

III. Complete graph with constant capacities

In this section we analyze the number of iterations of the while cycle in Algorithm 1 as a function of the number of nodes and of the initial capacity of the channels, when the underlying graph G is a clique. We first prove an equivalence between the random process in Algorithm 1 and the first hitting time of the boundary for a family of $m = |E|$ unbiased birth-and-death chains (see the random process in Algorithm 2).

Lemma III.1 (Equivalence lemma)**.** *Let τ_1 and τ_2 be the random variables indicating the number of iterations of the while cycles in Algorithms 1 and 2, respectively. Then, when m in Algorithm 2 equals $|E|$ in Algorithms 1, a coupling between the random processes exists such that $\tau_1 = \tau_2$.*

Proof. In a complete graph the edge connecting two nodes is the unique shortest path between them. Hence, when we pick a random source, a random destination, and a random shortest path at lines $2-4$ in Algorithm 1, we are picking a single edge u.a.r. and a random direction on that edge. Moreover, for an edge $e = \{u, v\}$, if we fix an arbitrary orientation of the edge and consider the quantity

$$X_t(u, v) = \frac{b_e^{(t)}(v) - b_e^{(t)}(u)}{2} \tag{1}$$

We have that, at each round t, such quantity changes in one of three ways: it increases by one, $X_{t+1}(u, v) = X_t(u, v) + 1$ (if u is picked at line 2 and v is picked at line 3 in Algorithm 1), it decreases by one, $X_{t+1}(u, v) = X_t(u, v) - 1$ (if v is picked at line 2 and u at line 3) or it remains the same $X_{t+1}(u, v) = X_t(u, v)$ (if the pair $\{u, v\}$ is not picked at lines $2, 3$). Hence, given the random process defined in Algorithm 1 we can define the random process in Algorithm 2, with $m = |E|$, as follows. Let $\hat{E}$ be an arbitrary orientation of the edges in E (i.e., for every $\{u, v\} \in E$ either $(u, v) \in \hat{E}$ or $(v, u) \in \hat{E}$) and let $f : E \to [m]$ be an arbitrary bijective map. When in the random process in Algorithm 1 we pick a source u and a destination v at lines 2 and 3, in the random process in Algorithm 2 we pick $f(\{u, v\})$ at line 4, and at line 5 we set $X(e) = X(e) + 1$ if $(u, v) \in \hat{E}$ and $X(e) = X(e) - 1$ if $(v, u) \in \hat{E}$. By construction, this defines a coupling of Algorithm 1 and Algorithm 2 with $\tau_1 = \tau_2$. □

Algorithm 2 Multiple Birth-and-Death chains process

Require: Two integers $m, k \in \mathbb{N}$

1: **for all** $e = 1, \dots, m$ **do**
2: Set $X(e) = 0$
3: **while** for every e, $X(e) \neq \pm k$ **do**
4: Pick e, u.a.r.
5: $X(e) = \begin{cases} X(e) + 1 & \text{with probability } 1/2 \\ X(e) - 1 & \text{with probability } 1/2 \end{cases}$
6: **return** The number of iterations of the while cycle

Theorem III.2. *Let $G = (V, E)$ be a complete graph with n nodes. For every channel $e \in E$, let $c(e) = 2k$ be its capacity, with $k \in \mathbb{N}$ and $k > \sqrt{4\alpha \log n}$ for some constant $\alpha > 1$, and let $b_e^{(0)}(u) = b_e^{(0)}(v) = k$ be its initial balance. Let τ be the random variable indicating the first time one of the channels is empty in one direction, i.e., $\tau = \inf\{t \in \mathbb{N} : \exists e \in E \text{ with } b_{\min}(e) = 0\}$. Then,*

1) $\tau = \mathcal{O}\left(k^2 n^2\right)$, *w.h.p.*[2]
2) $\tau = \Omega\left(k^2 n^2 / \log n\right)$, *w.h.p.*

Proof. Due to the equivalence proved in Lemma III.1 we can study the distribution of the number of iterations of the while cycle in Algorithm 2: let τ be the random variable indicating such number of iterations. Observe that $\tau = \min\{\tau^{(e)} : e = 1, \dots, m\}$ where $\tau^{(e)} = \inf\{t \in \mathbb{N} : X_t^{(e)} = \pm k\}$ and here

[2]We say that an event $\mathcal{E}_n$ depending on a parameter n, that here indicates the number of nodes in the graph, holds *with high probability (w.h.p.)* if a constant $c > 0$ exists such that $\mathbf{P}(\mathcal{E}_n) \geqslant 1 - n^{-c}$, for every sufficiently large n.

we indicate with $X_t^{(e)}$ the value that variable $X(e)$ at line 5 of Algorithm 2 has at the t-th iteration of the while cycle.

For $e = 1, \dots, m$ and for $t = 1, 2, \dots$, let $Y_t(e)$ be the indicator random variable of the event "Edge e is chosen at round t" in Algorithm2, and let

$$\overline{Y}_t(e) = \sum_{i=1}^{t} Y_i(e)$$

Since $\mathbf{P}\left(Y_i(e)\right) = 1/m$ for every e and every i, it holds that $\mathbf{E}\left[\overline{Y}_t(e)\right] = t/m$, and observe that for every fixed e random variables $\{Y_i(e) : i = 1, 2, \dots\}$ are independent.

As for the upper bound, from Chernoff Bound (see Lemma A.1 in the Appendix A) it follows that, for every $e = 1, \dots, m$

$$\mathbf{P}\left(\overline{Y}_t(e) \leqslant \frac{t}{2m}\right) \leqslant e^{-\frac{t}{8m}}.$$

Hence, if we chose $t = 4mk^2$, we have that for every $e = 1, \dots, m$, the probability that e has been chosen less than $2k^2$ times is at most

$$\mathbf{P}\left(\overline{Y}_t(e) \leqslant 2k^2\right) \leqslant e^{-k^2/2}.$$

and by using the union bound, the probability that an e exists that has been chosen less than $2k^2$ times is

$$\mathbf{P}\left(\exists e : \overline{Y}_t(e) \leqslant 2k^2\right) \leqslant me^{-k^2/2}. \tag{2}$$

Thus, if $k > \sqrt{2\alpha \log n}$ for some $\alpha > 1$, we have that with probability at least $1 - m^{-\alpha+1}$ at round $t = 4mk^2$ all $e \in \{1, \dots, m\}$ have been chosen at least $2k^2$ times.

For one single unbiased birth-and-death chain it is well-known that (see Lemma B.1 in Appendix B), if the chain starts at 0 then it will hit the boundary after k^2 steps, in expectation. From Markov inequality it thus follows that the probability that a single chain does not hit the boundary within $2k^2$ steps is at most $1/2$.

Let $\mathcal{E}$ be the event $\mathcal{E}$ = "At round $t = 4mk^2$ rounds all $e \in [m]$ have been chosen at least $2k^2$ times". From (2) we have that $\mathbf{P}\left(\mathcal{E}\right) = 1 - m^{-\alpha+1}$. Moreover, observe that conditionally on event $\mathcal{E}$, events $\{\tau(e) > 4mk^2 : e \in [m]\}$ are independent and each one of them has probability at most $1/2$ (since it is the probability that a chain e that has been chosen at least $2k^2$ times has not reached the boundary yet). Hence,

$$\begin{aligned}
&\mathbf{P}\left(\tau > 4mk^2\right) = \\
&= \mathbf{P}\left(\tau > 4mk^2 \mid \mathcal{E}\right)\mathbf{P}\left(\mathcal{E}\right) + \mathbf{P}\left(\tau > 4mk^2 \mid \overline{\mathcal{E}}\right)\mathbf{P}\left(\overline{\mathcal{E}}\right) \\
&\leqslant \mathbf{P}\left(\tau > 4mk^2 \mid \mathcal{E}\right) + \mathbf{P}\left(\overline{\mathcal{E}}\right) \\
&\leqslant \mathbf{P}\left(\text{for all } e \in [m],\ \tau(e) > 4mk^2 \mid \mathcal{E}\right) + m^{-\alpha+1} \\
&\leqslant 2^{-m} + m^{-\alpha+1}
\end{aligned}$$

As for the lower bound, from Chernoff bound (see Lemma A.1 in the Appendix) it follows that, for every $e = 1, \dots, m$

$$\mathbf{P}\left(\overline{Y}_t(e) \geqslant \frac{3t}{2m}\right) \leqslant e^{-\frac{9t}{4m}}.$$

Hence, if we chose $t = mk^2/(27 \log n)$, we have that for any $e = 1, \dots, m$, the probability that e has been chosen at least $k^2/(18 \log n)$ times is at most

$$\mathbf{P}\left(\overline{Y}_t(e) \geqslant k^2/(18 \log n)\right) \leqslant e^{-k^2/12}.$$

and by using the union bound, the probability that an e exists that has been chosen at least $k^2/(18 \log n)$ times is

$$\mathbf{P}\left(\exists e : \overline{Y}_t(e) \geqslant k^2/(18 \log n)\right) \leqslant me^{-k^2/12}. \tag{3}$$

Thus, if $k > \sqrt{\alpha \log n}$ for some large enough constant α, we have that with probability at least $1 - n^{-1}$ at round $t = mk^2/(27 \log n)$ all $e \in \{1, \dots, m\}$ have been chosen at most $k^2/(18 \log n)$ times.

For one single unbiased birth-and-death chain, if the chain starts at 0 then the probability that it hits the boundary at $\pm k$ within $k^2/(18 \log n)$ steps is smaller than $1/n^3$ (see Lemma B.3 in Appendix B). Hence, if we define event $\mathcal{E}$ = "At round $t = mk^2/(27 \log n)$ all $e \in [m]$ have been chosen at most $k^2/(18 \log n)$ times", then conditional on event $\mathcal{E}$ we have that

$$\begin{aligned}
&\mathbf{P}\left(\tau \leqslant \frac{mk^2}{27 \log n}\right) \leqslant \mathbf{P}\left(\tau \leqslant \frac{mk^2}{27 \log n} \;\middle|\; \mathcal{E}\right) + \mathbf{P}\left(\overline{\mathcal{E}}\right) \\
&= \mathbf{P}\left(\exists e \in \{1, \dots, m\} : \tau(e) \leqslant \frac{mk^2}{27 \log n} \;\middle|\; \mathcal{E}\right) + \mathbf{P}\left(\overline{\mathcal{E}}\right) \\
&\leqslant m\mathbf{P}\left(\tau(e) \leqslant \frac{mk^2}{27 \log n} \;\middle|\; \mathcal{E}\right) + 1/n \\
&\leqslant m/n^3 + 1/n \leqslant 2/n.
\end{aligned}$$

□

IV. Edge-betweenness and failure time

In Section III we have seen that, when the underlying graph is complete, the process described in Algorithm 1 can be coupled on the same probability space with a process involving the hitting time of the boundary for a family of independent birth-and-death chains (Algorithm 2). That is possible since any shortest-path in a clique consists in a single edge, hence at each round the balance of one single edge is updated and the updates of the balances of different edges in different rounds are independent. In order to generalize the results obtained for complete graph and constant capacities to the case of different graph topologies and to the case in which edges can have different initial capacities, the main technical difficulty stands in the fact that the updates of the balances of edges that belong to the path chosen in a specific round (line 4 in Algorithm 1) are not independent. However, we can still give upper and lower bounds on the time it takes to have the first payment failure, by losing only an extra $\mathcal{O}(\log n)$ factor in the upper bound due to the dependence of the edges, and by using the *edge-betweenness centrality* as a parameter, in Theorem IV.1 we indeed prove that the critical quantity in the estimation of the first failure time is the minimum of the ratios $k_e^2/g(e)$, where k_e is the capacity of edge e and $g(e)$ is its betweenness centrality in graph G.

Edge betweenness is a commonly used centrality measure in network analysis [5]. It quantifies how often an edge lies in shortest paths between pairs of nodes in a graph. Formally, the edge betweenness centrality $g(e)$ of an edge e is defined as

$$g(e) = \sum_{s,t \in V} \frac{\sigma(s,t \,|\, e)}{\sigma(s,t)} \tag{4}$$

where V is the set of nodes, $\sigma(s,t)$ is the number of shortest (s,t)-paths, and $\sigma(s,t \,|\, e)$ is the number of those paths containing edge e.

In our random process (Algorithm 1), the probability that an edge e is contained in the shortest path selected at any specific iteration is, by the law of total probability,

$$\mathbf{P}\left(e\right) = \sum_{s,t \in V} \mathbf{P}\left(e \,|\, s,t\right) \cdot \mathbf{P}\left(s,t\right) \tag{5}$$

where $\mathbf{P}\left(e \,|\, s,t\right)$ is the probability of picking a shortest path containing e at line 4 in Algorithm 1 conditional on the fact that s and t were selected as source and destination at lines 2 and 3, and $\mathbf{P}\left(s,t\right)$ is the probability of selecting s and t as source and destination of the path. Notice that the probability of picking e given s and t is exactly $\frac{\sigma(s,t \,|\, e)}{\sigma(s,t)}$ and $\mathbf{P}\left(s,t\right) = \frac{2}{n(n-1)}$, since we select the pair of nodes uniformly at random. Thus, $\mathbf{P}\left(e\right)$ is the normalized edge-betweenness of edge e

$$\mathbf{P}\left(e\right) = \frac{2}{n(n-1)} g(e) \tag{6}$$

Theorem IV.1. *Let $G = (V, E)$ be a graph with n nodes. For every channel $e \in E$, let $c(e) = 2k_e$ be its capacity, with $k_e \in \mathbb{N}$ and $k_e > \alpha\sqrt{\log n}$ for a sufficiently large constant α, and let $b_e^{(0)}(u) = b_e^{(0)}(v) = k_e$ be its initial balance. Let τ be the random variable indicating the first time one of the channels is empty in one direction, i.e., $\tau = \inf\{t \in \mathbb{N} : \exists e \in E \text{ with } b_{\min}(e) = 0\}$. Then,*

1) $\tau = \mathcal{O}\left(\xi\, n^2 \log n\right)$, *w.h.p.*
2) $\tau = \Omega\left(\xi \frac{n^2}{\log n}\right)$, *w.h.p.*

where $\xi = \min\left\{\frac{k_e^2}{g(e)} : e \in E\right\}$ and $g(e)$ is the betweenness centrality of edge e.

Proof. Let e be an edge, let $Y_t(e)$ be the indicator random variable of the event "Edge e is included in the path chosen at round t", and let

$$\overline{Y}_t(e) = \sum_{i=1}^{t} Y_i(e)$$

From (6) we have that $\mathbf{P}\left(Y_i(e) = 1\right) = \frac{2}{n(n-1)} g(e)$ for every i and thus the expected number of rounds edge e has been included in payment paths up to round t is

$$\mathbf{E}\left[\overline{Y}_t(e)\right] = t \cdot \frac{2}{n(n-1)} g(e)\,. \tag{7}$$

Let τ_e be the random variable indicating the first time edge e is empty in one direction

$$\tau_e = \inf\left\{t \in \mathbb{N} : b_{\min}\left(e\right) = 0\right\}$$

Notice that, restricted only to the rounds in which edge e is included in a path, the balance of edge e behaves according to an unbiased birth-and-death chain.

As for the upper bound, for an edge e with capacity k_e, the expected time to become empty in one direction is at most k_e^2 rounds in which edge e is included in the path chosen at that round, regardless of the initial balance of the edge (see Lemma B.1 in Appendix B). From Markov inequality it thus follows that, if at round t an edge e updated its balance more than $4k_e^2$ rounds, then the probability that it is not yet empty in one direction is at most $1/4$,

$$\mathbf{P}\left(\tau_e > t \mid \overline{Y}_t(e) > 4k_e^2\right) \leqslant \frac{1}{4}$$

Hence, if we take

$$\bar{t}_e = 4n(n-1)\frac{k_e^2}{g(e)}$$

from (7) we have that $\mathbf{E}\left[\overline{Y}_{\bar{t}_e}(e)\right] = 8k_e^2$ and, since $\{Y_i(e) : i = 1, \ldots, \bar{t}_e\}$ are independent, from Chernoff bound it follows that

$$\mathbf{P}\left(\overline{Y}_{\bar{t}_e}(e) \leqslant 4k_e^2\right) \leqslant e^{-(1/2)n(n-1)\left(k_e^2/g(e)\right)} \leqslant \frac{1}{4}$$

Hence

$$\begin{aligned}
&\mathbf{P}\left(\tau_e > \bar{t}_e\right) = \\
&= \mathbf{P}\left(\tau_e > \bar{t}_e \mid \overline{Y}_{\bar{t}_e}(e) > 4k_e^2\right) \mathbf{P}\left(\overline{Y}_{\bar{t}_e}(e) > 4k_e^2\right) \\
&\quad + \mathbf{P}\left(\tau_e > \bar{t}_e \mid \overline{Y}_{\bar{t}_e}(e) \leqslant 4k_e^2\right) \mathbf{P}\left(\overline{Y}_{\bar{t}_e}(e) \leqslant 4k_e^2\right) \\
&\leqslant \mathbf{P}\left(\tau_e > \bar{t}_e \mid \overline{Y}_{\bar{t}_e}(e) > 4k_e^2\right) + \mathbf{P}\left(\overline{Y}_{\bar{t}_e}(e) \leqslant 4k_e^2\right) \leqslant \frac{1}{2}
\end{aligned}$$

Since the above bound holds regardless of the initial balance of edge e, it follows that

$$\mathbf{P}\left(\tau_e > \bar{t}_e \log n\right) \leqslant \left(\frac{1}{2}\right)^{\log n} = \frac{1}{n}$$

Hence, for every edge $e \in E$, w.h.p. it holds that

$$\tau_e \leqslant 4n(n-1)\frac{k_e^2}{g(e)} \log n = \mathcal{O}\left(n^2 \log n \frac{k_e^2}{g(e)}\right)$$

The upper bound follows considering the edge(s) e with the smallest ratio $k_e^2/g(e)$.

As for the lower bound, we can proceed as in the proof of Theorem III.2. For an edge e with capacity $2k_e$, if e starts from a perfectly balanced configuration, from Lemma B.3 in Appendix B it follows that, at any time t,

$$\mathbf{P}\left(\tau_e \leqslant t \;\middle|\; \overline{Y}_t(e) \leqslant \frac{k_e^2}{18 \log n}\right) \leqslant 4e^{\frac{-k_e^2}{k_e^2/(3\log n)}} = \frac{4}{n^3} \tag{8}$$

If we take

$$\bar{t}_e = \frac{n(n-1)}{54 \log n} \cdot \frac{k_e^2}{g(e)}$$

from (7) we have that

$$\mathbf{E}\left[\overline{Y}_{\bar{t}_e}(e)\right] = \frac{k_e^2}{27 \log n}$$

and from Chernoff bound (Lemma A.1 in Appendix A with $\delta = 1/2$)

$$\mathbf{P}\left(\overline{Y}_{\bar{t}_e}(e) \geqslant \frac{k_e^2}{18\log n}\right) \leqslant e^{-\frac{1}{12}\frac{k_e^2}{27\log n}} \leqslant \frac{1}{n^2} \tag{9}$$

where in the last inequality we used the hypothesis $k_e > \alpha\sqrt{\log n}$ for a sufficiently large constant α.

Now let $\bar{t}$ be

$$\bar{t} = \frac{n(n-1)}{54\log n} \cdot \min\left\{\frac{k_e^2}{g(e)} \,:\, e \in E\right\} \tag{10}$$

and let $\mathcal{E}$ be the event $\mathcal{E}$ = "At round $\bar{t}$ for each edge $e \in E$ it holds that $\overline{Y}_{\bar{t}}(e) < \frac{k_e^2}{18\log n}$". Since $\bar{t} \leqslant \bar{t}_e$ for every $e \in E$, from (9) it follows that $\mathbf{P}\left(\overline{Y}_{\bar{t}}(e) \geqslant \frac{k_e^2}{18\log n}\right) \leqslant 1/n^2$. Hence,

$$\mathbf{P}\left(\overline{\mathcal{E}}\right) = \mathbf{P}\left(\exists e \in E \,:\, \overline{Y}_{\bar{t}}(e) \geqslant \frac{k_e^2}{18\log n}\right) \leqslant \frac{1}{n} \tag{11}$$

The first time an edge is empty in one direction is $\tau = \min\{\tau_e \,:\, e \in E\}$. For $\bar{t}$ defined as in (10) we thus have

$$\begin{aligned}
\mathbf{P}\left(\tau < \bar{t}\right) &\leqslant \mathbf{P}\left(\tau < \bar{t} \mid \mathcal{E}\right) + \mathbf{P}\left(\overline{\mathcal{E}}\right) \\
&\leqslant \mathbf{P}\left(\exists e \in E \,:\, \tau_e < \bar{t} \mid \mathcal{E}\right) + \mathbf{P}\left(\overline{\mathcal{E}}\right) \\
&\leqslant \mathbf{P}\left(\exists e \in E \,:\, \tau_e < \bar{t}_e \mid \mathcal{E}\right) + \mathbf{P}\left(\overline{\mathcal{E}}\right) \\
&\leqslant \frac{4}{n} + \frac{1}{n}
\end{aligned}$$

where in the last inequality we used the bounds in (8) and in (11). □

Theorem IV.1 generalizes the result on the complete graph of Theorem III.2 with respect to both graph topology and channel capacity. Notice that, when all edge capacities are equal, $k_e = k$ for all $e \in E$, then upper and lower bounds in Theorem IV.1 turn out to be, w.h.p.,

$$\tau = \begin{cases} \Omega\left(\frac{k^2}{g_{\max}} n^2 \log n\right) \\ \mathcal{O}\left(\frac{k^2}{g_{\max}} \cdot \frac{n^2}{\log n}\right) \end{cases}$$

where with $g_{\max} = \max\{g(e) \,:\, e \in E\}$ is the largest edge betweenness centrality. The above observation indicates that, despite the fact that the updates of the edges in general graphs are not independent, the impact of such correlation can only be limited to a $\log n$ factor, as a function of n. For fixed n and as a function of the edge capacity k the impact of the correlation is limited to a constant factor. In the next section we will show the results obtained with some simulations that highlight such constant-factor impact.

V. Simulations

In this section, we present the results of the simulations of the random process in Algorithm 1 on different classes of graphs.

We begin with the case of the complete graph (see Subsection V-A), where the goal is to compare the results of the simulations with the theoretical upper and lower bounds in Section III. As we will see in Figure 1, the results of the simulations give some evidence that the lower bound in Theorem III.2 could be tight.

To highlight the impact of the correlation between the updates of the balances of different edges, we then proceed with a comparison of the results of the simulations obtained in a graph with n nodes and a ring topology with the results that we get on an unstructured set of n edges (see Subsection V-B): In a graph with a ring topology, every time we pick a source and a destination uniformly at random and we pick the shortest path between them to route the payment, for any fixed edge e the probability that e is included in the path is about $1/4$, but edges are correlated (two edges that are close in the ring topology have larger probability to be updated together than two edges that are far apart); for comparison, we use an unstructured set of n edges, where at each round we pick each edge with the same probability, about $1/4$, independently of the other edges, and we independently update the balances of all the picked edges.

Finally, to highlight the impact of the distribution of the channel capacities in the Lightning Network, we run the simulation of our random process on a snapshot of the channel graph of the real Lightning Network (see Subsection V-C): We first consider each edge having its real channel capacity, as observed in the Lightning Network; then, we run the simulations on a graph whose topological structure is the same but the capacities of the channels are suitably redistributed, in order to quantify the improvement that we would get, with respect to the time to have the first payment failure in our model, if it was possible to modify the channel capacities.

The code for all simulations have been written in C++. The simulations were conducted on a system equipped with an AMD Ryzen 5 CPU (3.9 GHz) and 32 GB of RAM. The simulations in Subsection V-C refer to a snapshot of the Lightning Network collected on 2025/04/23 using our Lightning Network node, that runs an instance of LND [16].

A. Complete graph

In this subsection we present the results of the simulations of the random process on the complete graph. In Figure 1 we show the results that we get when we fix the capacity of each edge at 512 and we increase the number of nodes n up to 2800; for each n, we run the simulation 10 times. The plots in the figure show average, maximum and minimum failure time, for each value of n, over the 10 runs. To compare the results of the simulations with the theoretical results in Theorem III.2 we also plot the lines that we get for the theoretical upper and lower bounds, where the constants hidden in the asymptotic notation is chosen using nonlinear least squares fitting [17], i.e., minimizing the squared error with the simulated average. Namely, for the upper bound we plot the function $\tau = p \cdot (512)^2 n^2$, with $p = 0.0044$, and for the lower bound we plot $\tau = p \cdot \frac{(512)^2 n^2}{\log(N)}$, with $p = 29.0202$. The growth rate of the average of the simulations seems much closer to the growth rate of the theoretical lower bound than that of the upper bound.

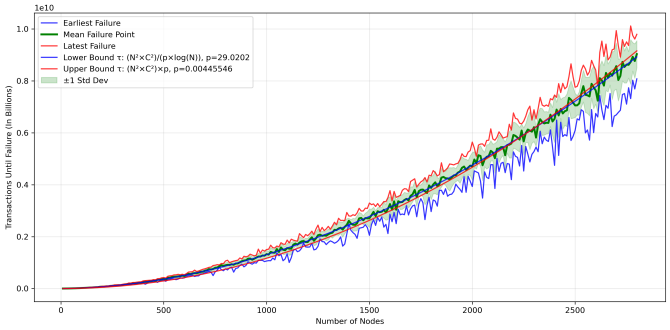


Fig. 1. The payment throughput in a clique with fixed channel capacity of 512. Lines depict earliest (blue), average (green), and latest (red) failure points, with theoretical bounds (dashed).The green band indicates standard deviation, demonstrating how payment capacity scales non-linearly with network size.

B. Ring topology & Independent edges

In this subsection we consider a ring with n nodes and a system of n independent chains. On a graph with a ring topology, when we sample source and destination of the payment u.a.r. among all nodes, it is easy to see that for each $i = 1, 2, \ldots, \lceil \frac{n}{2} \rceil - 1$, the probability that source u and destination v are at distance i is $\mathbf{P}\left(d(u,v) = i\right) = n/\binom{n}{2} = \frac{2}{n-1}$. For a given distance i and a given edge e, the probability that e is included in the chosen path conditional on the fact that source and destination are at distance i is $\mathbf{P}\left(e \mid d(u,v) = i\right) = \frac{i}{n}$.

$$\mathbf{P}\left(e\right) = \sum_{i=1}^{\lceil \frac{n}{2} \rceil - 1} \frac{2}{n-1} \cdot \frac{i}{n} = \frac{(\lceil \frac{n}{2} \rceil - 1)\lceil \frac{n}{2} \rceil}{n(n-1)} \simeq \frac{1}{4} \qquad (12)$$

On a ring topology (as well as in almost any other graph topologies) the update of the balances of edges involved in the same path in a given round of the random process are not independent, since the balances of all edges in the path are updated in the same *direction*: The balance of every edge e in the path decreases of one unit for the endpoint closer to the source and it increases of one unit for the endpoint closer to the destination. On the other hand, on a system of n independent chains, where at every round every chain is selected with probability given by (12) and each selected chain updates its balance independently of the other chains, the expected number of chains that update their balance equals the expected number of edges that update their balance at each round in the ring, but the updates of the balances in this system are completely independent.

Figure 2 shows the results that we get when we fix the number of nodes at 4096, all edges have the same capacity, and the capacity of each edge increases up to 3040. For each value of the capacity we run 10 simulations and we plot maximum, minimum and average first failure time over the runs, for the case in which edges are disposed according to a ring topology and for the case in which edges are independent. The plots show that in the case of the ring topology the growth rate of the average is faster than the growth rate of the independent case. It is also interesting to note that the results for the ring topology exhibit a larger variance than the results for the independent chains.

TX amount	min	max	mean	std
1k	1	108	13	12
10k	1	30	4	3.63
100k	1	12	1	0.95

TABLE I
TIME FAILURE COUNTS FOR THE UNMODIFIED GRAPH OF THE LIGHTNING NETWORK WITH DIFFERENT PAYMENT UNITS: $1k$, $10k$ AND $100k$ SATOSHIS, WHERE EACH CAPACITY SIMULATION WAS REPEATED 1000 TIMES.

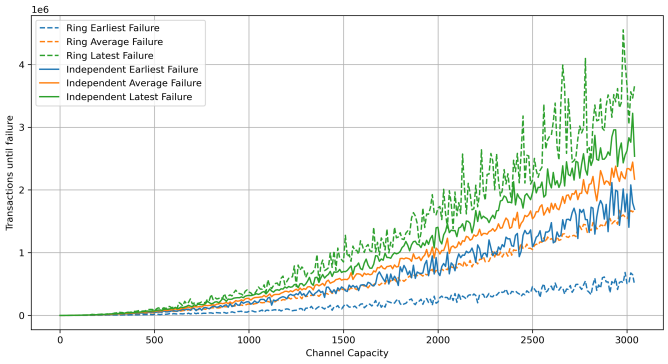


Fig. 2. Simulated failure probabilities for ring and independent chain systems with $n = 4096$ nodes and maximum capacity 3040.

C. Lightning Network

In this subsection we present the results of the simulations of our random process on a snapshot of the channel graph of the real Lightning Network, the snapshot was taken on April 23rd, 2025 and it consists of 16426 nodes and 51063 edges. Since the whole graph is not connected, in the simulation we only consider the giant component, that consists in 14900 nodes and 47087 edges. Furthermore, as the average channel capacity of the network is about 10 million *satoshi*[3], we use payment units of $1k$, $10k$, and $100k$ *satoshi*. For each unit, we measure the number of payments between randomly chosen source-destination pairs before the first payment failure occurs. Table I summarizes the results of the simulations, recording maximum, minimum, average, and standard deviation over 10 runs.

In a second set of simulations, we consider the same Lightning Network graph structure, but we redistribute the overall network capacity evenly over the edges, so that every edge has the same capacity and the sum of the capacities is preserved. Table II summarizes the results of this set of simulations.

Finally, in a third set of simulations, we consider the same Lightning Network graph structure, but we redistribute the overall network capacity in a way that is *optimized* with respect to our random process where we take source and

[3]In the Lightning Network, channel capacities and payment values are usually expressed in *satoshi* and *millisatoshi*; recall that 1 *bitcoin* corresponds to 100million *satoshi*

TX amount	min	max	mean	std
1k	74833576	968962509	439701719	204199218
10k	765772	9801570	3862534	1617777
100k	8084	115597	40089	17233

TABLE II

TIME FAILURE COUNTS FOR THE UNMODIFIED GRAPH WITH UNIFORM REDISTRIBUTION CAPACITY OF EDGES WITH PAYMENT UNITS: $1k$, $10k$, AND $100k$ SATOSHIS, WHERE EACH CAPACITY SIMULATION WAS REPEATED 1000 TIMES.

TX amount	min	max	mean	std
100k	841	2144747	782110	521689

TABLE III

TIME FAILURE COUNTS FOR THE UNMODIFIED GRAPH WITH OPTIMIZATION ON REDISTRIBUTION CAPACITY OF EDGES WITH PAYMENT UNIT: $100k$ SATOSHIS, WHERE THE SIMULATION WAS REPEATED 1000 TIMES.

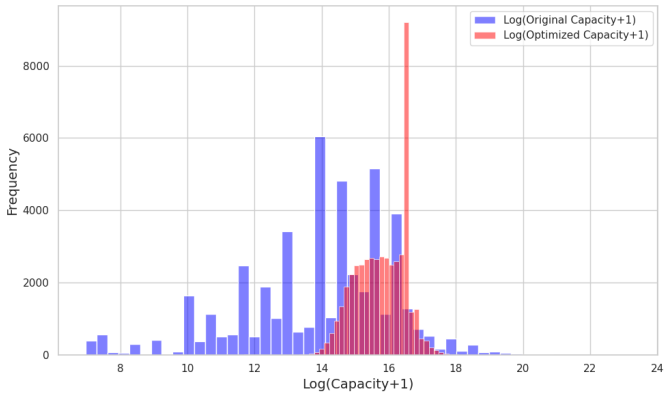


Fig. 3. Logarithmic distribution of channel capacities before and after optimization based on edge betweenness and initial capacities.

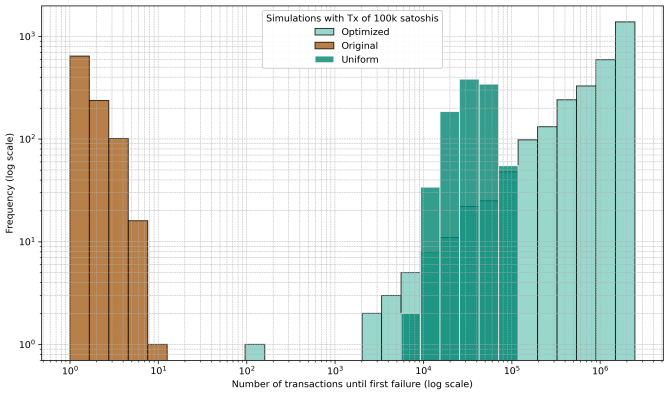


Fig. 4. Log-log histogram of simulations for sending payments of value $100k$ satoshis with different distribution each simulation was repeated 1000 times

destination nodes uniformly at random over all the nodes: More precisely, the overall network capacity is redistributed over the edges so that the ratio $k_e^2/g(e)$, the square of the capacity and the betweenness centrality, is the same for every edge. Figure 3 shows how the capacity was redistributed over the edges: On the x-axis there is the log of the capacity and on the y-axis the number of edges with that capacity, in the original Lightning Network and after the optimized redistribution. We can observe that the optimizing process concentrates the capacity distribution, and that most channels converge to similar capacity values. Table III presents the results for the simulations with payments of $100k$ satoshis. Running the simulations with payments of $10k$ and $1k$ satoshis would take too long with such redistribution of the channel capacity. A comparison of the results for the case of payments of $100k$ satoshis in the three sets of simulations is summarized in Figure 4: We can observe that the network essentially cannot handle transactions of $100k$ satoshis with its original capacity distribution; if we redistribute the overall capacity evenly over all edges, then the network becomes more resistant with respect to failure; finally, if we redistribute the overall capacity so the ratio $k_e^2/g(e)$ is equal for all edges, then there is a difference of almost 20 times on average, based on the results we showed in Table III and Table II. Clearly, since it is not possible to organize and modify the capacity of the channels of the Lightning Network in a centralized way, all the above redistributions are to be intended as thought experiments.

VI. CONCLUSIONS

In this paper we studied a random process defined over undirected graphs inspired by the problem of payment failures on the channel graph of the Lightning Network.

We first analyzed upper and lower bounds on the time it takes to have the first payment failure when the underlying graph is complete and all edges have the same capacity, as a function of the number of nodes in the graph and of the capacity of the edges. Our upper and lower bounds are almost tight, except for a gap of $\log n$, that would be interesting to close. Our conjecture is that the upper bound can be improved but the lower bound is tight.

In our random process, we pick source and destination of a payment uniformly at random among all nodes in the graph, and we pick a path between source and destination uniformly at random among all shortest paths between them. It implies that, in general graphs, the larger the betweenness centrality of an edge, the more often the edge appears in the selected paths. Since the number of times an edge has to be selected before a payment failure occurs is proportional to the square of the capacity, the critical quantity for the random process is thus the ratio $k_e^2/g(e)$ of the square of the capacity of an edge and its betweenness centrality. For arbitrary graphs we thus give upper and lower bounds on the time it takes to have a payment failure as functions of the number of nodes and of the minimum over all the edges of such ratios.

Our simulations on the complete graph validate the theoretical results and give some empirical evidence that the theoretical lower bound might be tight. The simulations on the ring give hints about the impact that the dependency among the balance updates of the edges belonging to the same path has on the time it takes for the first payment failure to occur. Finally, our simulations on the channel graph of the Lightning Network quantify the impact of the distribution of the channel capacity over the edges on the transaction failure rate.

References

[1] J. Poon and T. Dryja, "The bitcoin lightning network: Scalable off-chain instant payments," 2016, white paper. [Online]. Available: https://lightning.network/lightning-network-paper.pdf

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," https://bitcoin.org/bitcoin.pdf, 2008.

[3] K. Croman, C. Decker, I. Eyal, A. E. Gencer, A. Juels, A. Kosba, A. Miller, P. Saxena, E. Shi, E. Gün Sirer *et al.*, "On scaling decentralized blockchains: (a position paper)," in *International conference on financial cryptography and data security*. Springer, 2016, pp. 106–125.

[4] A. Antonopoulos, O. Osuntokun, and R. Pickhardt, *Mastering the Lightning Network: A Second Layer Blockchain Protocol for Instant Bitcoin Payments*. O'Reilly, 2021. [Online]. Available: https://github.com/lnbook/lnbook

[5] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[6] K. Lange, E. Rohrer, and F. Tschorsch, "On the impact of attachment strategies for payment channel networks," in *2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, 2021, pp. 1–9.

[7] E. Bergamini, P. Crescenzi, G. D'Angelo, H. Meyerhenke, L. Severini, and Y. Velaj, "Improving the betweenness centrality of a node by adding links," *ACM J. Exp. Algorithmics*, vol. 23, Aug. 2018. [Online]. Available: https://doi.org/10.1145/3166071

[8] O. Ersoy, S. Roos, and Z. Erkin, "How to profit from payments channels," in *Financial Cryptography and Data Security: 24th International Conference, FC 2020, Kota Kinabalu, Malaysia, February 10–14, 2020, Revised Selected Papers*. Berlin, Heidelberg: Springer-Verlag, 2020, p. 284–303.

[9] R. Pickhardt, S. Tikhomirov, A. Biryukov, and M. Nowostawski, "Security and privacy of lightning network payments with uncertain channel balances," *CoRR*, vol. abs/2103.08576, 2021. [Online]. Available: https://arxiv.org/abs/2103.08576

[10] M. Dotan, Y.-A. Pignolet, S. Schmid, S. Tochner, and A. Zohar, "Survey on blockchain networking: Context, state-of-the-art, challenges," *ACM Computing Surveys (CSUR)*, vol. 54, no. 5, pp. 1–34, 2021.

[11] A. Ghosh, M. Mahdian, D. M. Reeves, D. M. Pennock, and R. Fugger, "Mechanism design on trust networks," ser. WINE'07. Berlin, Heidelberg: Springer-Verlag, 2007, p. 257–268.

[12] D. DeFigueiredo and E. Barr, "Trustdavis: a non-exploitable online reputation system," in *Seventh IEEE International Conference on E-Commerce Technology (CEC'05)*, 2005, pp. 274–283.

[13] P. Dandekar, A. Goel, R. Govindan, and I. Post, "Liquidity in credit networks: a little trust goes a long way," ser. EC '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 147–156. [Online]. Available: https://doi.org/10.1145/1993574.1993597

[14] A. Goel and G. Ramseyer, "Continuous credit networks and layer 2 blockchains: Monotonicity and sampling," *Proceedings of the 21st ACM Conference on Economics and Computation*, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:215754249

[15] S. Dreyer, A. Genitrini, and M. Naima, "Optimal Uniform Shortest Path Sampling," in *WALCOM: Algorithms and Computation*, ser. Lecture Notes in Computer Science, vol. 15411. Chengdu, China: Springer Nature Singapore, Feb. 2025, pp. 160–179. [Online]. Available: https://hal.science/hal-04669060

[16] L. Labs, "lnd: Lightning network daemon," https://github.com/lightningnetwork/lnd, 2025.

[17] T. L. Lai, H. Robbins, and C. Z. Wei, "Strong consistency of least squares estimates in multiple regression," *Proceedings of the National Academy of Sciences*, vol. 75, no. 7, pp. 3034–3036, 1978. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.75.7.3034

[18] D. P. Dubhashi and A. Panconesi, *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

[19] M. Mitzenmacher and E. Upfal, *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press, 2017.

[20] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, pp. 493–507, 1952.

[21] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times*. American Mathematical Society, Providence, 2009.

Appendix A
Concentration inequalities

In this section, we state a fundamental concentration bound that we used in the proofs in Sections III and IV. We refer the reader intrested in a simple proof to [18] or [19].

Lemma A.1 (Chernoff bound [20])**.** *Let* $\{X_i : i = 1, \dots, n\}$ *be a family of independent Bernoulli random variables with* $\mathbf{P}(X_i = 1) = p_i$ *and let* $X = \sum_{i=1}^{n} X_i$*. Then, for every* $0 < \delta < 1$ *it holds that*

- $\mathbf{P}(X \leqslant (1-\delta)\mu) \leqslant e^{-\frac{\delta^2}{2}\mu}$
- $\mathbf{P}(X \geqslant (1+\delta)\mu) \leqslant e^{-\frac{\delta^2}{3}\mu}$

where $\mu = \mathbf{E}[X] = \sum_{i=1}^{n} p_i$.

Appendix B
Preliminaries on birth-and-death chains

In this section we briefly recall some terminology and fundamental bounds about *birth-and-death* Markov chains that we used in Sections III and IV.

A finite birth-and-death chain with absorbing boundary is a Markov chain $\{X_t : t \in \mathbb{N}\}$ with state space

$$\Omega = \{-k, -k+1, \dots, -1, 0, 1, \dots, k-1, k\}$$

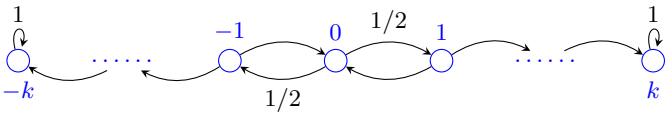
for some $k \in \mathbb{N}$, where

$$\mathbf{P}(X_{t+1} = -k \,|\, X_t = -k) = \mathbf{P}(X_{t+1} = k \,|\, X_t = k) = 1$$

and for each $j = -k+1, \dots, -1, 0, 1, \dots, k-1$,

$$\begin{aligned} \mathbf{P}(X_{t+1} = j+1 \,|\, X_t = j) &= p \\ \mathbf{P}(X_{t+1} = j-1 \,|\, X_t = j) &= q \\ \mathbf{P}(X_{t+1} = j \,|\, X_t = j) &= 1-(p+q) \end{aligned}$$

for some non-negative p and q with $p+q<1$. When $p = q = 1/2$ the chain is said to be *unbiased*.



For unbiased birth-and-death chains it is well known that the expected time before the chain hits the boundary starting from an arbitrary state j is $k^2 - j^2$ (see, e.g., Chapter 2 in [21]).

Lemma B.1. *Let* $\tau = \inf\{t \in \mathbb{N} : X_t = \pm k\}$ *be the random variable indicating the first time in which the chain hits state* k*. The expectation of* τ *for the chain starting at an arbitrary state* $j \in \Omega$ *is*

$$\mathbf{E}[\tau \,|\, X_0 = j] = k^2 - j^2$$

To bound the probability that a finite chain hits the boundary before or after a certain value, it is sometimes convenient to consider infinite birth-and-death chains. The following Lemma B.2 is a well-known bound (see, e.g., Exercise 2.10 in [21]) for infinite chains that can be used to give an upper bound on the probability that a chain hits the boundary within a certain number of time steps, as shown in Lemma B.3.

For an event $\mathcal{E}$ and a state $j \in \Omega$ we use the standard notation $\mathbf{P}_j(\mathcal{E})$ for the probability of $\mathcal{E}$ conditional on the event that the Markov chains $\{X_t\}$ starts at state j,

$$\mathbf{P}_j(\mathcal{E}) = \mathbf{P}(\mathcal{E} \mid X_0 = j) .$$

Lemma B.2. *Let $\{X_t : t \in \mathbb{N}\}$ be an unbiased birth-and-death with state space $\Omega = \mathbb{Z}$. For every time $t \in \mathbb{N}$ and for every target state $k \in \mathbb{N}$ it holds that*

$$\mathbf{P}_0(|X_t| \geqslant k) \leqslant \mathbf{P}_0\left(\max_{i=1,\dots,t} |X_i| \geqslant k\right) \leqslant 2\mathbf{P}_0(|X_t| \geqslant k)$$

The following lemma is used in the proofs of the lower bounds in Theorems III.2 and IV.1.

Lemma B.3. *Let $\{X_t : t \in \mathbb{N}\}$ be an unbiased birth-and-death chain with state space $\Omega = \mathbb{Z}$, let $k \in \mathbb{N}$ be an integer and let τ be the random variable indicating the first time the chain hits state $\pm k$,*

$$\tau = \inf\{t \in \mathbb{N} : |X_t| = k\} .$$

Then, for every $t \in \mathbb{N}$ it holds that

$$\mathbf{P}_0(\tau \leqslant t) \leqslant 4e^{-k^2/(6t)}$$

Proof. Observe that, the chain hits states $\pm k$ within time t if and only if the maximum over all times i up to t of $|X_i|$ is at least k,

$$\begin{aligned}\mathbf{P}_0(\tau \leqslant t) &= \mathbf{P}_0\left(\max_{i=1,\dots,t} |X_i| \geqslant k\right) \leqslant 2\mathbf{P}_0(|X_t| \geqslant k) \\ &= 4\mathbf{P}_0(X_t \geqslant k)\end{aligned}$$

where in the second inequality we used Lemma B.2 and in the third equality we used the symmetry of the process, i.e., $\mathbf{P}_0(X_t \geqslant +k) = \mathbf{P}_0(X_t \leqslant -k)$.

Now observe that, since the state of the chain at each round increases of one unit with probability $1/2$ and decreases of one unit with probabiity $1/2$, then the probability that the state of the chain at time t is at least k equals the probability that in a sequence of t unbiased coin tosses at least $(t+k)/2$ ended up head,

$$\mathbf{P}_0(X_t \geqslant k) = \mathbf{P}\left(B(t, 1/2) \geqslant \frac{k+t}{2}\right)$$

where with $B(t, 1/2)$ we indicated a binomial random variable with parameters t and $1/2$. The thesis then follows from the Chernoff bound (see Lemma A.1). □