Applied Probability Trust (7 November 2024)

A STOCHASTIC PROCESS ON A NETWORK WITH CONNECTIONS TO LAPLACIAN SYSTEMS OF EQUATIONS

IQRA ALTAF GILLANI,* *IIT Delhi*

AMITABHA BAGCHI,** *IIT Delhi*

POOJA VYAVAHARE,*** *IIT Tirupati*

Abstract

We study an open discrete-time queueing network. We assume data is generated at nodes of the network as a discrete-time Bernoulli process. All nodes in the network maintain a queue and relay data, which is to be finally collected by a designated sink. We prove that the resulting multi-dimensional Markov chain representing the queue size of nodes has two behavior regimes depending on the value of the rate of data generation. In particular, we show that there is a non-trivial critical value of the data rate below which the chain is ergodic and converges to a stationary distribution and above which it is non-ergodic, i.e., the queues at the nodes grow in an unbounded manner. We show that the rate of convergence to stationarity is geometric in the sub-critical regime.

Keywords: Queueing networks; Random Walks; Geometric Ergodicity

2020 Mathematics Subject Classification: Primary 60G10

Secondary 60G50;60K25

1. Introduction

We study an open discrete-time queueing network whose interconnections are described by an undirected simple graph. Some of the vertices of the graph produce "data packets" according to a discrete-time Bernoulli process. One vertex of the graph is designated as a "data sink" and packets disappear when they reach this vertex. Each node apart from the sink maintains a queue and relays at most one packet in a time slot in the manner of the "gossip" models widely studied in the networking and distributed computing literature [13][2][25]. The packet is relayed to a random neighbor after the fashion of a random walk on the graph. Viewing the sink as a vertex that "collects" the packets being generated and relayed through the graph, we call this process the *Data Collection Process.*

The Data Collection Process is defined on a graph $G = (V, E)$ equipped with a positive edge-weight function $w : E \to \mathbb{R}_+$. The edge weights determine the probability of a packet moving to a neighbour. The process takes two parameters, a *relative rate vector* $\boldsymbol{J} \in \mathbb{R}_+^{|V|}$ and a *rate* $\beta \in (0, 1)$; we assume that node $v \in V$ produces a packet with probability $\beta \boldsymbol{J}(v)$ in a given time slot. For a given relative rate vector, the process has two behavior regimes and undergoes a sharp transition between these two regimes, the controlling parameter being the rate β. Specifically, we will show that for a critical value β^* we have that when $\beta > \beta^*$ the process is non-ergodic, and the size of the queues grows to infinity, whereas, when $\beta < \beta^*$ process is ergodic such that all queues are almost surely finite and the system converges to a stationary distribution. For this latter regime, we also show that the rate of convergence is geometric, i.e., the Data Collection Process is geometrically ergodic whenever $\beta < \beta^*$.

For $\beta < \beta^*$ the process also has an unexpected connection with a subclass of systems of linear equations, which we refer to as "one-sink" Laplacian systems. This connection allows us to give a lower bound on the critical rate β^* in terms of the eigenvalues of the transition matrix of the natural random walk defined on G: $P_w = D^{-1}A$ where $A_{uv} = w_{uv}$ for all $(u, v) \in E$ and 0 otherwise, and D is a diagonal matrix with $D_{uu} = \sum_{v:(u,v)\in E} w_{uv}$. The parameters of the random walk also make an appearance in our geometric ergodicity result. We find that for $\beta = \beta^*(1 - \delta)$ the Data Collection Process converges exponentially at a rate proportional to the hitting time of the random walk with transition matrix P_w and inversely proportional to δ, the relative distance of the rate from the critical value. A feature of the proof of geometric convergence is an interesting use of the "backward analysis" argument of Leighton, Maggs and Ranade [17] that is a highlight of the theory of packet routing in networks.

1.1. Related work

The Data Collection Process is an open queueing network [14] that can be viewed as a multi-dimensional Markov Chain. The question of ergodicity of such chains was investigated by Tsybakov and Mikhailov in the context of computer networks in their work on Slotted ALOHA systems [32]. Subsequently, Georgiadis and Szpankowski proved a non-trivial stability regime for another networking-inspired model, the Token Passing ring [7] and, later, Szpankowski extended this result to general random access systems including Slotted ALOHA networks [30]. While the area of open queueing networks is vast we identify these models in particular since they are similar to our Data Collection Process and we will see in Section 3 that Szpankowski's program for establishing the ergodicity of Slotted ALOHA can be carried out for the Data Collection Process as well.

The Data Collection Process also shows connections to other fields of study. In the area of computer networking we note that such a process can be used as a traffic benchmark and, in fact, Kamra et. al. use exactly such a benchmark to test a coding scheme in [12]. In the area of distributed computing we have shown in another work that the Data Collection Process can be used to analyze an in-network function computation scenario [11]. In the area of information dissemination Markov chains similar to our Data Collection Process have been used to analyze the timeliness of information received by nodes in a network, c.f., e.g., the work of Tripathi, Talak and Modiano [31].

In a different setting the Data Collection Process also shows an interesting connection with Laplacian system of equations whose solutions find wide applicability in network analysis, computer vision, operations research, machine learning, and computational biology. After Spielman and Teng gave the first efficient methods for solving such system in [28], a number of efficient Laplacian solvers have been proposed over the years by Koutis, Miller and Peng [15], Cohen et al. [5], and Kyng and Sachdeva [16] among others. Most of these solvers are based on similar ideas; they use Chebyshev iteration or the conjugate gradient method with complex graph-theoretic constructions or sampling. However, Laplacian systems can also be related to random walks. In particular, the electrical flow in a network can be written as a Laplacian equation and the relations between electrical quantities and statistical properties of random walks have been known for a long time, and have been discussed at length by, e.g., Doyle and Snell [6], and Levin, Peres and Wilmer [19]. Becchetti, Bonifaci, and Natale [1] exploit this connection to solve such Laplacian equations using simple random walk primitives. In this work, we also show how the steady-state equation of the Data Collection Process maps to such Laplacian system of equations. This provides a framework that we were able to use in a subsequent work to design and analyze a simple and efficient Laplacian solver [8].

1.2. Organization

The rest of the paper is organized as follows. In Section 2, we discuss our main results. In Section 3, we prove the existence of a non-trivial critical data rate below which the Data Collection Process is ergodic and above which it is non-ergodic. Then, in Section 4 we characterize this rate in terms of the underlying graph's parameters. In Section 5, we prove that the process is not only ergodic but geometrically ergodic and find the rate at which the associated Markov chain converges to its stationary distribution. Finally, in Section 6 we conclude and give some directions for future work.

2. Main results and discussions

2.1. Our model: The Data Collection Process

We consider a stochastic process on a network modeled by an undirected graph $G = (V, E, w)$, where V is the set of n nodes, E is the set of edges such that $|E| = m$, and a positive weight function $w : E \to \mathbb{R}_+$. We say that $u \sim v$ if $(u, v) \in E$ and $\mathrm{Nbd}(u) := \{v \in V | (u, v) \in E\}$. We consider a diagonal matrix D such that $D_{uu} = \deg(u)$ where $\deg(u) := \sum_{v\in\mathrm{Nbd}(u)} w_{uv}$ is the generalized degree of node u. We denote the maximum and minimum generalized degree among all nodes in the network by $d_{\max}$ and $d_{\min}$ respectively.

We consider time to be discrete and define the process in terms of the generation, movement and disappearance of "packets" from the system. In order to do this we are given a *relative rate vector* $\boldsymbol{J} \in \mathbb{R}^n$ with the properties that (i) $\boldsymbol{J}(v) < 0$ for exactly one node and (ii) $\sum_{i=1}^n \boldsymbol{J}(i) = 0$. The node v for which $\boldsymbol{J}(v) < 0$ is called the *sink* and we will use u_s to denote it hereafter. We also define a set of *source nodes*: $V_s = \{v : \boldsymbol{J}(v) > 0\}$. We are also given a *rate parameter* $\beta \geq 0$ such that $\max_{i=1}^n \beta\boldsymbol{J}(i) \leq 1$. We assume that each node in $V \setminus \{u_\mathrm{s}\}$ is equipped with a queue. The number of packets in the queue at u at time t is denoted by $Q_t(u)$.

Packets appear in the system at the source nodes $v \in V_s$ which receive external packet arrivals as an independent Bernoulli process with rate $\beta\boldsymbol{J}(v)$. The packet received externally is placed in the queue at v. Packet movement at time t takes place as follows: For each $u \in V \setminus \{u_{\rm s}\}$, if $Q_t^{\boldsymbol{J},\beta}(u) > 0$ a single data packet is picked at random from the queue and sent to v with probability $w_{uv}/\deg(u)$. So, each node sends at most one packet from its queue in one time step and may receive multiple packets, up to one from each neighbour. A packet is removed from the system when a neighbour of $u_{\rm s}$ decides to transmit that packet to the $u_{\rm s}$.

In the following we will refer to the $|V|-1$-dimensional Markov chain $\left\{Q_t^{\boldsymbol{J},\beta}\right\}_{t\geq 0}$ as the *Data Collection Process on G with relative rate vector $\boldsymbol{J}$ and rate parameter β.*

2.2. Ergodicity is a critical phenomenon for the Data Collection Process

For a given Data Collection Process on a network modeled by an undirected graph $G = (V, E, w)$, there is an associated $|V| - 1$-dimensional vector $Q_t^{\boldsymbol{J},\beta}$ where each $Q_t^{\boldsymbol{J},\beta}(u)$ represents the queue size at a given node $u \in V \setminus \{u_{\rm s}\}$ given a data rate β. Since the Data Collection Process is a queueing system, the question of stability arises, i.e., we need to understand whether the system is able to successfully transfer data at a given value β which is the controlling parameter for the rate at which packets appear in the system. For this, following Loynes [21] and Szpankowski [30], we formally define a notion of a *stable data rate* as follows.

Definition 1. (*Stable rate.*) Given a weighted undirected graph $G = (V, E, w)$ and a relative rate vector $\boldsymbol{J}$ with $\boldsymbol{J}(v) < 0$ for exactly one $v \in V$, the process $Q_t^{\boldsymbol{J},\beta}$ is said to be *stable* and a value $\beta \geq 0$ of the rate parameter is said to be a *stable rate* if

$$\lim_{t\to\infty} \mathbb{P}\left[||Q_t^{\boldsymbol{J},\beta}||_\infty < x\right] = F(x), \text{ and } \lim_{x\to\infty} F(x) = 1 \tag{1}$$

where $F(x)$ is the limiting distribution function.

However, if a weaker condition holds i.e.,

$$\lim_{x\to\infty} \liminf_{t\to\infty} \mathbb{P}\left[||Q_t^{\boldsymbol{J},\beta}||_\infty < x\right] = 1 \tag{2}$$

the process is said to be *substable* and otherwise *unstable.* So, a stable process is necessarily substable and for a substable process to be stable its distribution function should tend to a limit.

Thus, for the Data Collection Process by stability we mean the distribution of $Q_t^{\boldsymbol{J},\beta}$ as $t \to \infty$ exists. In particular, we show that the given process has two distinct regimes, one ergodic and one non-ergodic, as we vary β and there is a sharp transition between them. We find that there is a non-trivial $\beta^* > 0$ such that the chain $\left\{Q_t^{\boldsymbol{J},\beta}\right\}_{t\geq 0}$ is ergodic for β below this value and converges to a stationary distribution. Above β^* the queue sizes grow unbounded as $t \to \infty$. Specifically we show the following theorem:

Theorem 1. *Consider a weighted undirected graph $G = (V, E, w)$ and a relative rate vector $\boldsymbol{J}$ with $\boldsymbol{J}(v) < 0$ for exactly one $v \in V$. If the random walk on G with transition matrix P_w where $P_w[u, v] = w_{uv}/deg(u)$ is irreducible and aperiodic then there exists a $\beta^* > 0$ such that the resulting multi-dimensional Markov chain $\left\{Q_t^{\boldsymbol{J},\beta}\right\}_{t\geq 0}$ is ergodic for all $\beta < \beta^*$ and non-ergodic for all $\beta \geq \beta^*$.*

There are multiple ways of proving this theorem. In Section 3 we prove it by using the induction-based technique developed by Georgiadis and Szpankowski [7], and later summarized by Szpankowski in his study of slotted ALOHA [30]. This technique gives nice insights into the workings of the Data Collection Process.

2.3. A lower bound on the critical rate

When $\beta < \beta^*$ the Data Collection Process is ergodic and has a stationary distribution so we can define $\boldsymbol{\eta}^\beta(v) = \lim_{t\to\infty} \mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(v) > 0\right]$ for all $v \in V \setminus \{u_{\mathrm{s}}\}$. We will show in Section 4 that at stationarity the vector $\boldsymbol{\eta}$ extended to u_{s} by setting $\boldsymbol{\eta}(u_{\mathrm{s}}) = 0$ is a solution to a linear system

$$\boldsymbol{\eta}^T(I - P_w) = \beta \boldsymbol{J}^T,$$

where P_w is the transition matrix of the random walk defined on G by the weight function w. In Section 4.1 we discuss the relationship of this system to the Laplacian of G and the implications of this relationship. For now, we state one important consequence of this relationship: a lower bound on β^*.

Theorem 2. *Suppose we have a Data Collection Process with relative rate vector $\boldsymbol{J}$ such that $\boldsymbol{J}(v) > 0$ for $v \in V_s$ and $\boldsymbol{J}(v) < 0$ only for $v = u_s$, defined on a graph $G = (V, E, w)$ that satisfies the conditions of Theorem 1 and has critical rate β^*. Then if P_w is the transition matrix of the random walk defined by w on G and λ_2^w is the* *second largest eigenvalue of P_w then*

$$\beta^* \geq \frac{(1-\lambda_2^w)}{\sum_{i \in V_s} \boldsymbol{J}(i)} \frac{\sqrt{d_{\min} deg(u_s)}}{(d_{\max} + deg(u_s))}. \tag{3}$$

We also prove an upper bound on the critical data rate for the special case where $V_s = V \setminus \{u_s\}$. In order to present this bound, we need to define some terms. For any vertex $u \in V$, we define its measure as, $\rho(u) := \sum_{v \in V} P_w[u,v]$. Similarly, for any $U \subset V$ we define the measure $\rho(U) = \sum_{u \in U} \rho(u)$. We also define the edge boundary as $\partial U := \{(u,v) : u \in U, v \notin U\}$, so, $\rho(\partial U) = \sum_{u \in U, v \notin U} P_w[u,v]$. We have the following upper bound result.

Proposition 1. *Given a graph $G = (V, E, w)$ with $|V| = n$ nodes out of which there is one sink u_s and set $V_s = V \setminus \{u_s\}$ of source nodes running a Data Collection Process having critical data rate β^* as defined by Theorem 1. To achieve stable queues β^* must satisfy*

$$\beta^* \leq \min \left\{ \hat{h}(G), \sum_{u: u \sim u_s} \frac{P_w[u, u_s]}{n-1} \right\} \tag{4}$$

where P_w is the transition matrix of random walk defined by w, $\hat{h}(G) = \min_{U \subset V, u_s \notin U} \frac{\rho(\partial U)}{\rho(U)}$ is a constant and $\hat{h}(G)$ is at most $h(G)$, the edge expansion of graph G.

Table 1: Rate lower bounds for various graphs with $w : E \to \mathbf{1}$ and $|V_s| = S$

Graph	$\beta \geq \frac{(1-\lambda_2^w)}{\sum_{i \in V_s} \boldsymbol{J}(i)} \frac{\sqrt{d_{\min} \deg(u_s)}}{(d_{\max} + \deg(u_s))}$	**Exact rate**
Star Graph with sink at centre and ϵ as self loop probability at each node	$\frac{1}{2S\sqrt{n-1}}$	$1-\epsilon$
Star Graph with sink and source at outer nodes	$\frac{1}{Sn}$	$\frac{1}{S(n-1)}$
Complete graph	$\frac{n}{2S(n-1)}$	$\frac{n}{(S+1)(n-1)}$
Random Geometric Graph	$\frac{\log n}{2Sn}$	-

Table 2: Rate lower bounds for various graphs with $w : E \to \mathbf{1}$ and $|V_s| = 1$

Graph	$\beta \geq \frac{(1-\lambda_2^w)}{\sum_{i \in V_s} \boldsymbol{J}(i)} \frac{\sqrt{d_{\min} \deg(u_S)}}{(d_{\max}+\deg(u_S))}$	**Exact rate**
Cycle	$\frac{1}{2n^2}$	$\frac{2}{n}$
Wheel Graph W_{n+1} with sink at centre and source at one of the cycle vertices	$\frac{\log n \sqrt{3n}}{2n^2}$	$\frac{1}{3}$
Wheel Graph W_{n+1} with source at centre and sink at one of the cycle vertices	$\frac{3 \log n}{n(n+1)}$	$\frac{5}{3n}$
Complete Binary tree with both source and sink at leaves	$\frac{1}{4n}$	$\frac{1}{6 \log n - 3}$
k-times star of star graph with both source and sink at leaves	$\frac{1}{n^2 + n^{\frac{2k-1}{k}}}$	$\frac{1}{1 + (2k-1)n^{1/k}}$
k-times star of star graph with source at center and sink at leaf	$\frac{1}{2n^{\frac{4k-1}{2k}}}$	$\frac{1}{1 + (k-1)n^{1/k}}$

Table 3: Rate upper bounds for various graphs with $w : E \to \mathbf{1}$ and $V_s = V \setminus \{u_S\}$

Graph	$\beta \leq \min \left\{ \hat{h}(G), \sum_{u: u \sim u_S} \frac{P_w[u, u_S]}{n-1} \right\}$	**Exact rate**
Cycle	$\frac{1}{n-1}$	$\frac{4}{n^2}$
Star Graph with sink at centre and ϵ as self loop probability at each node	$1 - \epsilon$	$1 - \epsilon$
Star Graph with sink and source at outer nodes	$\frac{1}{(n-1)^2}$	$\frac{1}{(n-1)^2}$
Complete graph	$\frac{1}{n-1}$	$\frac{1}{n-1}$

In Table 1 and 2, we present lower bound on the critical data rate for the stochastic Data Collection Process on various graphs for the case where the source set size is $|V_s| = S$ and $|V_s| = 1$ respectively. We also present the exact values of data rate which are easy to calculate using elementary algebra for these topologies. In all these cases, we assume that all edges have unit weight $w : E \to \mathbf{1}$ i.e., random walk defined by P_w is simple random walk.

If we consider the complete graph topology it is easy to see that the exact rate is $n/(S+1)(n-1)$. As, the spectral gap of the simple random walk on the complete graph of n nodes is $n/n-1$, we note that for this case our lower bound is tight i.e., both the exact value and the lower bound have order $\Theta(1/S)$ where $|V_s| = S$. Similarly, for the star graph with sink at outer node, our lower bound is tight and is of order $\Theta(1/Sn)$. Hence it is clear that our lower bound cannot admit any asymptotic improvement in general. On the other hand, consider cycle topology which shows that for specific cases a better lower bound may be possible. We note that our spectral gap-based lower bound is a $\Theta(1/n)$ lower than the exact value for this case. Similarly, for other topologies like wheel graph, complete binary tree and k-times star of star graph ($n^{1/k}$-regular tree defined on k levels) a better lower bound is possible.

Regarding the upper bound, we achieve tight upper bound in case of the complete graph and the star graph with both sink at centre and at outer node. However, for the cycle graph our upper bound is $\Theta(1/n)$ higher than the exact rate for this setup.

2.4. Geometric Ergodicity

We show that when $\beta < \beta^*$ the Data Collection Process converges to its stationary distribution at a geometric rate, i.e., the process is geometrically ergodic. Following Meyn and Tweedie [24], we define geometric ergodicity formally:

Definition 2. (*Geometric ergodicity.*) Given an irreducible and aperiodic Markov chain Φ defined on state space $\mathcal{X}$ with transition probability $\mathcal{P}[\cdot,\cdot]$ and stationary distribution π, the chain is said to be *geometrically ergodic* if there exist constants $\rho < 1$, $R > 0$, and, for every state $\boldsymbol{x} \in \mathcal{X}$ there exists a $C_{\boldsymbol{x}} < \infty$, such that for all $t > 0$,

$$\|\mathcal{P}^t[\boldsymbol{x},\cdot] - \pi\| \leq R C_{\boldsymbol{x}} \rho^t. \tag{5}$$

We use the coupling method to prove that convergence happens at a geometric rate. The convergence rate is in terms of the *hitting time*, t_{hit}, of the random walk P_w defined on G which has been widely studied since long time, see [3, 26, 33]. In particular, it is defined as follows. If $\{X_t\}_{t\geq 0}$ is a random walk on G and $\tau_v = \min\{t : X_t = v\}$, then $t_{\text{hit}} = \max_{u,v\in V} \mathbb{E}[\tau_v \mid X_0 = u]$, i.e., the maximum over all pairs (u,v) of vertices of the expected time taken for a random walk begun at u to first reach the vertex v. We show the following convergence theorem.

Theorem 3. *Consider* $\left\{Q_t^{\boldsymbol{J},\beta}\right\}_{t\geq 0}$ *defined on* $G = (V,E,w)$ *such that there is a critical* β^* *as described in Theorem 1. Let* $\beta = \beta^*(1-\delta)$ *for* $\delta \in (0,1)$ *and denote by* $\mathcal{P}$ *the transition matrix for the resulting multi-dimensional Markov Chain. Suppose we have* $\boldsymbol{x}, \boldsymbol{y} \in (\mathbb{N}\cup\{0\})^{|V|-1}$ *with* $\sum_{i=1}^{|V|-1} \boldsymbol{x}_i = N^{(\boldsymbol{x})}, \sum_{i=1}^{|V|-1} \boldsymbol{y}_i = N^{(\boldsymbol{y})}$*. Then*

$$||\mathcal{P}^t[\boldsymbol{x},\cdot] - \mathcal{P}^t[\boldsymbol{y},\cdot]||_{TV} \leq 2^{\left(1+\frac{\left(\max\{N^{(\boldsymbol{x})},N^{(\boldsymbol{y})}\}-1\right)\delta}{2t_{hit}}\right)} \cdot \left(\frac{1}{2}\right)^{\frac{\delta}{2t_{hit}}\cdot t}. \tag{6}$$

Convergence to stationarity can be derived as a special case of Theorem 3 by choosing $\boldsymbol{y} \in (\mathbb{N}\cup\{0\})^{|V|-1}$ according to the π, the stationary distribution of chain $\left\{Q_t^{\boldsymbol{J},\beta}\right\}_{t\geq 0}$. This establishes the geometric ergodicity of the Data Collection Process in the subcritical regime.

Corollary 1. *Consider the multi-dimensional Markov chain* $\left\{Q_t^{\boldsymbol{J},\beta}\right\}_{t\geq 0}$ *with* $\beta = \beta^*(1-\delta)$ *for* $\delta \in (0,1)$ *as defined in Theorem 3 and denote its stationary distribution by* π*. For* $\boldsymbol{x} \in (\mathbb{N}\cup\{0\})^{|V|-1}$ *such that* $\sum_{i=1}^{|V|-1} \boldsymbol{x}_i = N^{(\boldsymbol{x})}$,

$$||\mathcal{P}^t[\boldsymbol{x},\cdot] - \pi||_{TV} \leq 2^{\left(1+\frac{N^{(\boldsymbol{x})}\delta}{2t_{hit}}\right)} \cdot \left(\frac{1}{2}\right)^{\frac{\delta}{t_{hit}((1-\delta)\beta^*+2)}\cdot t}. \tag{7}$$

Moreover, for the special case that $\boldsymbol{x} = \boldsymbol{0}$*, i.e., the system begins with empty queues, the Markov chain mixes to within* $1/M$ *of its stationary distribution in terms of total variation distance for any parameter* $M > 0$ *in time* t *that is* $\Theta\left(\frac{t_{hit}\log M}{\delta}\right)$.

3. Ergodicity as a critical phenomenon

In this section, we prove Theorem 1, i.e., we show the existence of a non-trivial critical data rate β^* for the multi-dimensional Markov chain $\left\{Q_t^{\boldsymbol{J},\beta}\right\}_{t\geq 0}$ associated with the Data Collection Process such that the chain is ergodic for all values below β^* and non-ergodic above it.

We first state two lemmas that we will need: Szpankowski's "isolation lemma" (Lemma 1) and Loynes' scheme [21] as adapted to our situation (Lemma 2).

Lemma 1. (Szpankowski [29].) *Given* $N_t = (N_t^1, N_t^2, \cdots, N_t^M)$*, an* M*-dimensional Markov chain.*

1. *If it is defined on a countable state space, then the stability of* N_t^j *for all* $j \in M$ *implies the stability of the multi-dimensional Markov chain* N_t*.*

2. *If for some* j*, say* j^**,* $N_t^{j^*}$ *is unstable, then* N_t *is also unstable.*

Lemma 2. (Loynes [21].) *Given a pair* (X_t^j, Y_t^j) *of a strictly stationary and ergodic process, let* $U_t^j = X_t^j - Y_t^j$*. Then, the following holds:*

1. *If* $\mathbb{E}\left[U_t^j\right] < 0$*, then* N_t^j *is stable.*

2. *If* $\mathbb{E}\left[U_t^j\right] > 0$*, then* N_t^j *is unstable and* $\lim_{t\to\infty} N_t^j = \infty$ *(a.s.).*

We will also need the following property of the Markov chain $Q_t^{\boldsymbol{J},\beta}$ associated with the Data Collection Process: The queue occupancy probability of a node $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right]$ is an increasing function of β for all $u \in V \setminus \{u_{\mathrm{s}}\}$ and it is continuous for all $\beta < \beta^*$ where β^* is the critical rate above which the queues are unstable and below which they are stable.

Lemma 3. *Given an undirected graph* $G = (V, E, w)$ *running a Data Collection Process. Let* $Q_t^{\boldsymbol{J},\beta}$ *represent the queues at time* t *for all nodes* $u \in V \setminus \{u_s\}$*. Then, for all such nodes* $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right]$ *is*

1. *an increasing function of* β*, and*

2. *continuous for all* $\beta < \beta^*$ *where* β^* *is the critical data rate such that all data rates* $\beta < \beta^*$ *are stable and* $\beta \geq \beta^*$ *are unstable.*

Proof. (1). To prove this property, we will first establish that the multi-dimensional Markov chain $Q_t^{\boldsymbol{J},\beta}$ is stochastically ordered i.e., stochastically larger initial states will produce stochastically larger chains at all times. For this, let us consider a coupling as used by Szpankowski of two trajectories of this chain $\{Q_t^{\boldsymbol{J},\beta}\}$ and $\{\bar{Q}_t^{\boldsymbol{J},\beta}\}$ such that $\bar{Q}_0^{\boldsymbol{J},\beta} \preceq_{\mathrm{SD}} Q_0^{\boldsymbol{J},\beta}$. Now, assume the stochastic dominance relation between the two holds at time t i.e., $\bar{Q}_t^{\boldsymbol{J},\beta} \preceq_{\text{SD}} Q_t^{\boldsymbol{J},\beta}$. Then, at time step $t+1$ for both $Q_t^{\boldsymbol{J},\beta}$ and $\bar{Q}_t^{\boldsymbol{J},\beta}$ from the one-step basic queue evolution equation at all nodes $u \in V \setminus \{u_{\text{s}}\}$ we have

$$\mathbb{E}\left[Q_{t+1}^{\boldsymbol{J},\beta}(u) \mid Q_t^{\boldsymbol{J},\beta}(u)\right] = Q_t^{\boldsymbol{J},\beta}(u) - \mathbf{1}_{\{Q_t^{\boldsymbol{J},\beta}(u)>0\}} \sum_{v:v\sim u} P_w[u,v] + \sum_{v:v\sim u} P_w[v,u]\mathbf{1}_{\{Q_t^{\boldsymbol{J},\beta}(v)>0\}} + A_t(u) \tag{8}$$

where $A_t(u)$ is the number of packets generated at u, which is 0 if $u \notin V_s$ and is 1 with probability $\beta\mathbf{J}(v)$ if $v \in V_s$, so, $\mathbb{E}\left[A_t(u)\right] = \beta\mathbf{J}(u)$. Now consider any node u at time $t+1$, from the induction hypothesis queues at node u as well as its neighbours in $Q_t^{\boldsymbol{J},\beta}$ will dominate over the ones in $\bar{Q}_t^{\boldsymbol{J},\beta}$, so the first three terms on the right of Eq. (8) in $Q_t^{\boldsymbol{J},\beta}(u)$ will dominate the ones for $\bar{Q}_t^{\boldsymbol{J},\beta}(u)$ and since β is same, the last term is same for both cases. So, we have $\bar{Q}_{t+1}^{\boldsymbol{J},\beta}(u) \preceq_{\text{SD}} Q_{t+1}^{\boldsymbol{J},\beta}(u)$. This is true for all nodes $u \in V \setminus \{u_{\text{s}}\}$, so we have at time $t+1$, $\bar{Q}_{t+1}^{\boldsymbol{J},\beta} \preceq_{\text{SD}} Q_{t+1}^{\boldsymbol{J},\beta}$. Hence, by induction the Markov chain $Q_t^{\boldsymbol{J},\beta}$ is stochastically ordered.

Now to prove monotonicity, for $\beta < \beta'$ let us consider a coupling similar to the one used before of two stochastically ordered Markov chains $Q_t^{\boldsymbol{J},\beta}$ and $Q_t^{\boldsymbol{J},\beta'}$ such that $Q_0^{\boldsymbol{J},\beta} \preceq_{\text{SD}} Q_0^{\boldsymbol{J},\beta'}$. Then, as we know for all $u \in V \setminus \{u_{\text{s}}\}$, $\beta\boldsymbol{J}(u) < \boldsymbol{J}, \beta'\boldsymbol{J}(u)$, so by using induction and evolving queues using one-step queue evolution equation (Eq. (8)), we can show that $Q_t^{\boldsymbol{J},\beta} \preceq_{\text{SD}} Q_t^{\boldsymbol{J},\beta'}$ for all t. Hence, by induction we have $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right]$ is an increasing function of β for all $u \in V \setminus \{u_{\text{s}}\}$.

(2). To prove the continuity of the given function for $\beta < \beta^*$, we will again consider a similar coupling, however between two stochastically ordered Markov chains $Q_t^{\boldsymbol{J},\beta}$ and $Q_t^{\boldsymbol{J},\beta-d\beta}$ with infinitesimal $d\beta$. For the data generation rule in the two chains, we have whenever new data packet is generated at any node in $Q_t^{\boldsymbol{J},\beta-d\beta}$ chain then, it is definitely generated at the corresponding node in $Q_t^{\boldsymbol{J},\beta}$ chain but not vice-versa. To understand the difference in the two chains, let $N_t^{\beta-d\beta}$ and N_t^{β} denote the total number of packets in the respective chains till time t and $\Lambda_t^{\beta} = N_t^{\beta} - N_t^{\beta-d\beta}$. Now, consider $g : [0,1] \to \mathbb{R}$ to be a function dependent on β such that $g(\beta) = \mathbb{E}\left[Q_{t+1}^{\boldsymbol{J},\beta}(u) - Q_t^{\boldsymbol{J},\beta}(u)\right]$ which is bounded by definition. So, if we look at the derivative of this function, the term where $\Lambda_t^{\beta} = 0$ will be zero by definition of coupling, as the two chains behave differently only when there is an extra generated packet. Similarly, terms with $\Lambda_t^{\beta} \geq 2$ will have higher powers of $d\beta$ which will become zero as $d\beta \to 0$. Hence, the derivative

$g'(\beta)$ only depends on $\Lambda_t^\beta = 1$ term i.e.,

$$g'(\beta) = \lim_{d\beta \to 0} \left(g(\beta) - g(\beta - d\beta) \mid \Lambda_t^\beta = 1 \right) \left(|V_s| t (1 - d\beta)^{|V_s|t-1} \right)$$

where $V_s \subset V$ is the set of data sources. So, the total number of data packets generated in the two Markov chains upto time t differ by one and hence, the queues at nodes in the two chains differ by at most one data packet at any time step. Now, for the given coupled chains let t' be the time by which an extra packet is generated in chain $Q_t^{\boldsymbol{J},\beta}$. So, we have,

$$\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0 | Q_t^{\boldsymbol{J},\beta-d\beta}(u) = 0\right] = \frac{\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0 \cap Q_t^{\boldsymbol{J},\beta-d\beta}(u) = 0\right]}{\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta-d\beta}(u) = 0\right]} = \sum_{t'=1}^{t} d\beta P_{t'} \tag{9}$$

where $P_{t'}$ is the probability that the extra packet generated in chain $Q_t^{\boldsymbol{J},\beta}$ is present at node $u \in V \setminus \{u_s\}$. This means

$$\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right] - \mathbb{P}\left[Q_t^{\boldsymbol{J},\beta-d\beta}(u) > 0\right] \leq \sum_{t'=1}^{t} d\beta P_{t'}. \tag{10}$$

So, if $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0 | Q_t^{\boldsymbol{J},\beta-d\beta}(u) = 0\right]$ is defined, as, $d\beta \to 0$ from the above equation we have, $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right] - \mathbb{P}\left[Q_t^{\boldsymbol{J},\beta-d\beta}(u) > 0\right] \to 0$. Similarly, for the other side if $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta+d\beta}(v) > 0 | Q_t^{\boldsymbol{J},\beta}(u) = 0\right]$ is defined, so as $d\beta \to 0$, similar to Eq.(10) we have, $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta+d\beta}(u) > 0\right] - \mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right] \to 0$. Now, if both these conditions are true then the function is continuous as it has both left and right continuity respectively.

Now, consider all data rates $\beta < \beta^*$ where β^* is the critical rate below which all rates are stable and above which all are unstable. So, for such rates both the probabilities $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0 | Q_t^{\boldsymbol{J},\beta-d\beta}(u) = 0\right]$ and $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta+d\beta}(v) > 0 | Q_t^{\boldsymbol{J},\beta}(u) = 0\right]$ are defined, so as discussed above the function is continuous on both sides for all $\beta < \beta^*$. Now consider the case of data rates $\beta \geq \beta^*$. At β^*, we know $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta^*}(u) > 0\right] - \mathbb{P}\left[Q_t^{\boldsymbol{J},\beta^*-d\beta}(u) > 0\right]$ is defined (see Eq. (9)), as rate $\beta^* - d\beta$ is stable by definition, hence, the function is left continuous for this rate. However, for the other side since we know β^* is not stable i.e., $\lim_{t\to\infty} \mathbb{P}\left[Q_t^{\boldsymbol{J},\beta^*}(v) = 0\right] = 0$, hence, $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta^*+d\beta}(v) > 0 | Q_t^{\boldsymbol{J},\beta^*}(u) = 0\right]$ will not be defined and function is not right continuous. So, for $\beta \geq \beta^*$ function is left continuous but not right continuous. However, for all $u \in V \setminus \{u_s\}$, $\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right]$ is a continuous function (both limits exist) for all $\beta < \beta^*$. □

Proof of Theorem 1. We first proceed by proving the sufficient part i.e., existence of a non-trivial $\beta^* > 0$ such that the multi-dimensional Markov chain is ergodic for all $\beta < \beta^*$ and then the necessary part of the argument i.e., for all $\beta \geq \beta^*$ the chain is non-ergodic.

Sufficiency. Given a partition (P, U) of $V \setminus \{u_s\}$ queues we define a modification, $\bar{Q}_t^{\beta,U}$, of the $|V|-1$-dimensional chain $Q_t^{\boldsymbol{J},\beta}$. In this modification all nodes in U have the same behavior as in $Q_t^{\boldsymbol{J},\beta}$ but the nodes in $V \setminus \{u_s\} \setminus U$ are not allowed to have empty queues. Let us now first set $U = \emptyset$ (we will call these nodes "non-persistent") and $P = V \setminus \{u_s\}$ (we call these "persistent" nodes). For any $\beta \in (0,1)$, we know the one step basic queue evolution equation under the Data Collection Process for any u is as follows.

$$\mathbb{E}\left[Q_{t+1}^{\boldsymbol{J},\beta}(u) \mid Q_t^{\boldsymbol{J},\beta}(u)\right] = Q_t^{\boldsymbol{J},\beta}(u) - \mathbf{1}_{\{Q_t^{\boldsymbol{J},\beta}(u)>0\}} \sum_{v:v\sim u} P_w[u,v] + \sum_{v:v\sim u} P_w[v,u]\mathbf{1}_{\{Q_t^{\boldsymbol{J},\beta}(v)>0\}} + \beta \boldsymbol{J}(u).$$

So, at each node u we have an arrival from v with probability $P_w[v,u]$ in $\bar{Q}_t^{\beta,\emptyset}$ since the queue of v is always non-empty and the departure is the usual $\sum_{v:v\sim u} P_w[u,v]$.

Now, since we know $P_w[u_s, v] = 0$ for all $v \in V \setminus \{u_s\}$, so the sum of the outgoing probabilities from $V \setminus \{u_s\}$ is greater than the sum of the incoming probabilities, i.e., $\sum_{u\in V\setminus\{u_s\}} \sum_{v:v\sim u} P_w[u,v] > \sum_{u\in V\setminus\{u_s\}} \sum_{v:v\sim u, v\in V\setminus\{u_s\}} P_w[v,u]$. Therefore, there must be a vertex $u^* \in V\setminus\{u_s\}$ for which $\sum_{v:u^*\sim v} P_w[u^*,v] > \sum_{v:u^*\sim v, v\in V\setminus\{u_s\}} P_w[v,u^*]$. So, from Eq. (8) for this u^* we note that the expected drift is

$$-\sum_{v:u^*\sim v} P_w[u^*,v] + \sum_{v:u^*\sim v, v\in V\setminus\{u_s\}} P_w[v,u^*] + \beta \boldsymbol{J}(u^*)$$

which is negative for an appropriately small but non-zero value of β, let's call it β_{u^*}.

Now, to apply Loynes' scheme to vertex u^* we need to ensure that the sequence $(I_t(u^*), O_t(u^*))$ is strictly stationary where $I_t(u^*)$ is the number of incoming packets to u^* at time t and $O_t(u^*)$ is the number of outgoing packets from u^*. Since all nodes $u \in P$, so u^* as well as its neighbours always have a packet in the queue, so, both $O_t(u^*)$ and $I_t(u^*)$ are sequences of independent Bernoulli random variables and hence are stationary and ergodic. So, we can apply Loynes' scheme (Lemma 2) to claim that the one-dimensional process $\bar{Q}_t^{\beta_{u^*},\emptyset}(u^*)$ is stable, and, hence, $Q_t^{\beta_{u^*}}(u^*)$ is stable.

Now, we assume there is a non-empty set U of non-persistent users and a $\beta_U > 0$ such that $\bar{Q}_t^{\beta_U,U}(U)$ is stable and has a stationary distribution. To apply Loynes' scheme to a vertex, $u \in P = V \setminus \{u_\mathrm{s}\} \setminus U$ we need to ensure that the sequence $(I_t(u), O_t(u))$ is strictly stationary. Since $u \in P$ there is always a packet in the queue at u and so $O_t(u)$ is a sequence of independent Bernoulli random variables which takes value 1 with probability $\sum_{v:v\sim u} P_w[u,v]$ and 0 otherwise. We decompose $I_t(u)$ as the sum 0-1 random variables A_t^{uv}, where $A_t^{uv} = 1$ if u receives a packet from v at time t. Then

$$I_t(u) = \sum_{v\in U} A_t^{uv} + \sum_{v\in P} A_t^{uv}.$$

Since all $v \in P$ have a packet in their queue at all $t \geq 0$, each $\sum_{v\in P} A_t^{uv}$ is the sum of Bernoulli random variables and hence taken from a strongly stationary sequence. If we start the $\bar{Q}_t^{\beta_U,U}$ from an initial state picked according to this stationary distribution which ensures that the process stays in the stationary state for all $t \geq 0$. In particular, this implies that for any $v \in P$, number of incoming packets from v at time $t \geq 0$ is a sequence of random variables that is strongly stationary. Therefore $(I_t(u), O_t(u))$ is a strongly stationary sequence and we can apply Loynes' scheme. The expected drift at time $t \geq 0$ at any $u \in P$ for any $\beta \leq \beta_U$ is given by

$$-\sum_{v:u\sim v} P_w[u,v] + \sum_{u\sim v, v\in P} P_w[v,u] + \sum_{u\sim v, v\in U} P_w[v,u] P_w[\bar{Q}_t^{\beta,U}(u) > 0] + \beta \boldsymbol{J}(u). \quad (11)$$

Since the graph is connected and so there is at least one pair (w_1, w_2) such that $w_1 \in U, w_2 \in P$ and $P_w[w_1, w_2] > 0$, therefore we know that $\sum_{u\in P} \sum_{v\sim u} P_w[u,v] > \sum_{u\in P, v\in U} \sum_{v\sim u} P_w[v,u]$. This means that there is a $u^* \in P$ such that $\sum_{u^*\sim v} P_w[u^*, v] > \sum_{u^*\sim v, v\in P} P_w[v, u^*]$. For this u^* the first two terms in Eq. (11) add up to a value which is negative. Further from Lemma 3 we note that the third term is continuous and increasing in β and tends to 0 as $\beta \downarrow 0$. Hence, it is possible to find a value $\beta_{U\cup\{u^*\}}$ which lies in $(0, \beta_U)$ such that the expected drift is negative. So, from Loynes' scheme (Lemma 2) this implies that $\bar{Q}_t^{\beta,U}(U \cup \{u^*\})$ is stable for $\beta < \beta_{U\cup\{u^*\}}$. Moreover, from Lemma 1 since the stability of all the one-dimensional Markov Chains associated with the vertices in $U \cup \{u^*\}$ implies the stability of the overall multi-dimensional chain. Consequently, the same holds for $Q_t^{\boldsymbol{J},\beta}(U \cup \{u^*\})$. Therefore by induction there is a β^* such that for $\beta < \beta^*$, $Q_t^{\boldsymbol{J},\beta}$ is stable.

Necessity. Corresponding to the sequence by which the stability region is expanded to include all the vertices of $V \setminus \{u_\mathrm{s}\}$ there is a sequence $\beta_{u_1}, \beta_{u_2}, \ldots, \beta_{u_{|V\setminus\{u_\mathrm{s}\}|}}$ such that $\beta^* = \min\{\beta_{u_1}, \beta_{u_2}, \ldots, \beta_{u_{|V\setminus\{u_\mathrm{s}\}|}}\}$. Let w be the vertex for which $\beta_w = \beta^*$. Assume for the sake of simplicity of presentation that $\beta_w < \min\{\beta_u : u \in V \setminus \{u_\mathrm{s}\} \setminus \{w\}\}$. Hence we can choose any β such that $\beta_w < \beta < \min\{\beta_u : u \in V \setminus \{u_\mathrm{s}\} \setminus \{w\}\}$. For this β we know that $Q_t^{\beta, V\setminus\{u_\mathrm{s}\}\setminus\{w\}}(V \setminus \{u_\mathrm{s}\} \setminus \{w\})$ is stable. If we start this chain from its stationary distribution then the number of packets that are transmitted from $V \setminus \{u_\mathrm{s}\} \setminus \{w\}$ to w form a strongly stationary sequence. Since w is persistent in this setting the packets leaving it are also strongly stationary. Hence Loynes' scheme (Lemma 2) can be applied. By the choice of β we know that the expected drift at w is strictly positive and so $\bar{Q}_t^{\beta, V\setminus\{u_\mathrm{s}\}\setminus\{w\}}(w)$ is unstable and hence by Lemma 1, $\bar{Q}_t^{\beta, V\setminus\{u_\mathrm{s}\}\setminus\{w\}}$ is unstable.

In order to show that $Q_t^{\boldsymbol{J},\beta}$ is also unstable for this choice of β we will show there is a coupling of $Q_t^{\boldsymbol{J},\beta}$ and $\bar{Q}_t^{\beta, V\setminus\{u_\mathrm{s}\}\setminus\{w\}}$ with an appropriately chosen initial condition such that the two models behave *exactly* similarly. We know on the set of sample paths (of positive probability) on which the queue at w remains strictly positive the two coupled models behave exactly similarly because the difference only arises if the queue at w becomes 0 at time t, in which case $\bar{Q}_{t+1}^{\beta, V\setminus\{u_\mathrm{s}\}\setminus\{w\}}(w)$ is automatically set to 1 since w is persistent and $Q_{t+1}^{\boldsymbol{J},\beta}(w)$ remains 0. Now, we know that $\bar{Q}_t^{\beta, V\setminus\{u_\mathrm{s}\}\setminus\{w\}}(w)$ is unstable, so when we start $\bar{Q}_t^{\beta, V\setminus\{u_\mathrm{s}\}\setminus\{w\}}(V \setminus \{u_\mathrm{s}\} \setminus \{w\})$ according to its stationary distribution and we set the queue at w to 1, there is positive probability that this queue never reaches 0. So, for those cases $Q_t^{\boldsymbol{J},\beta}(w)$ behaves similarly as $\bar{Q}_t^{\beta, V\setminus\{u_\mathrm{s}\}\setminus\{w\}}(w)$ i.e., it is unstable. Therefore with these initial conditions $Q_t^{\boldsymbol{J},\beta}(w)$ is not substable since with positive probability $\lim_{t\to\infty} \mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(w) > m\right]$, for all finite m. Hence, $Q_t^{\boldsymbol{J},\beta}(w)$ is unstable and, by Lemma 1, $Q_t^{\boldsymbol{J},\beta}$ is unstable for our choice of β and, by the monotonicity of the process (see Lemma 3), it is unstable for all choices of $\beta \geq \beta^*$. $\square$

Discussion on proof technique. Our proof of Theorem 1 follows a general method for proving the existence of a "stability region" which is similar to the induction-based technique used by Georgiadis and Szpankowski to characterize the stability region of the multi-queue system described by token passing rings [7]. Szpankowski showed that this technique applies to a class of multi-queueing systems with certain properties [30], and the Data Collection Process falls in this class.

A more direct approach to proving Theorem 1 would involve taking the sum of all queues as a Lyapunov function and then proving that the 1-step drift of this Lyapunov function is negative when the sum of queues is greater than some $N > 0$. There are some challenges in this method. We want to show that

$$\mathbb{E}\left[\sum_{v \in V\setminus\{u_{\mathrm{S}}\}} Q_{t+1}^{\boldsymbol{J},\beta}(v) - \sum_{v \in V\setminus\{u_{\mathrm{S}}\}} Q_{t}^{\boldsymbol{J},\beta}(v) \mid \sum_{v \in V\setminus\{u_{\mathrm{S}}\}} Q_{t}^{\boldsymbol{J},\beta}(v) > N\right] < 0,$$

but in general this drift need not be negative. If none of the N packets currently in the system are in a queue adjacent to the sink, then this drift cannot be negative, and, in fact is strictly positive. Instead we may try to find an $\ell > 1$ such that

$$\mathbb{E}\left[\sum_{v \in V\setminus\{u_{\mathrm{S}}\}} Q_{t+\ell}^{\boldsymbol{J},\beta}(v) - \sum_{v \in V\setminus\{u_{\mathrm{S}}\}} Q_{t}^{\boldsymbol{J},\beta}(v) \mid \sum_{v \in V\setminus\{u_{\mathrm{S}}\}} Q_{t}^{\boldsymbol{J},\beta}(v) > N\right] < 0.$$

However, even determining this ℓ would involve making non-trivial arguments.

This was one reason why we chose to use Szpankowski's technique to prove Theorem 1. The other reason was that the induction on the graph is a pleasing technique which extends stability one vertex at a time to the entire graph.

4. Characterizing the critical rate

In section 3, we proved the ergodicity of the Markov chain associated with the Data Collection Process and showed that its stationary distribution exists. Now, in this section we will show that at steady-state the Data Collection Process satisfies a special class of linear equations the we call the "one-sink" Laplacian system. Using this equivalence we will derive a lower bound on the critical rate. Lastly, we will also present an upper bound on the critical rate.

4.1. Equivalence to one-sink Laplacian systems

The basic one step queue evolution equation under the Data Collection Process for any node $u \in V$ is as follows.

$$\mathbb{E}\left[Q_{t+1}^{\boldsymbol{J},\beta}(u) \mid Q_t^{\boldsymbol{J},\beta}(u)\right] = Q_t^{\boldsymbol{J},\beta}(u) - \mathbf{1}_{\{Q_t^{\boldsymbol{J},\beta}(u)>0\}} \sum_{v:v\sim u} P_w[u,v] + \sum_{v:v\sim u} P_w[v,u]\mathbf{1}_{\{Q_t^{\boldsymbol{J},\beta}(v)>0\}} + A_t(u), \quad (12)$$

where the second and third term on the right-hand side of the above equation represents the transmissions sent to and received from the neighbours respectively and $A_t(u)$ is the number of packets generated at u, which is 1 with probability $\beta\boldsymbol{J}(u)$ if $u \in V_s$, for the sink $\beta\boldsymbol{J}(u_{\mathrm{s}}) = -\beta\sum_{v\in V_s}\mathbf{J}(v)$, and for all other nodes $\boldsymbol{J}(u) = 0$, where $u \notin \{V_s \cup \{u_{\mathrm{s}}\}\}$. Now, taking expectations on both sides of Eq. (12) and let $\boldsymbol{\eta}_t^\beta(u) = \mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right]$ be the queue occupancy probability of node u and observing that $\mathbb{E}\left[A_t(u)\right] = \beta\mathbf{J}(u)$, where $\boldsymbol{J}$ is the relative rate vector, we have

$$\mathbb{E}\left[Q_{t+1}^{\boldsymbol{J},\beta}(u)\right] = \mathbb{E}\left[Q_t^{\boldsymbol{J},\beta}(u)\right] - \boldsymbol{\eta}_t^\beta(u)\sum_{v:v\sim u} P_w[u,v] + \sum_{v:v\sim u} P_w[v,u]\boldsymbol{\eta}_t^\beta(v) + \beta\mathbf{J}(u). \quad (13)$$

From Theorem 1, we know that for an appropriately chosen value of β the Data Collection Process has a steady state. Moreover, at steady state $\mathbb{E}\left[Q_t^{\boldsymbol{J},\beta}(u)\right]$ is a constant, so if we let $\boldsymbol{\eta}^\beta(u) = \lim_{t\to\infty}\mathbb{P}\left[Q_t^{\boldsymbol{J},\beta}(u) > 0\right]$ be the queue occupancy probability of node u at the stationarity, then we have the steady-state equation for the given node as

$$-\boldsymbol{\eta}^\beta(u)\sum_{v:v\sim u} P_w[u,v] + \sum_{v:v\sim u} P_w[v,u]\boldsymbol{\eta}^\beta(v) + \beta\mathbf{J}(u) = 0. \quad (14)$$

We can also represent the steady-state equations of all $|V| = n$ nodes in matrix form as follows. For this, let us first order the nodes such that the nth node represents the sink. Let $\boldsymbol{\eta}$ be an n element column vector representing the steady-state queue occupancy probability $\boldsymbol{\eta}^\beta(u)$ of nodes $u \in V$. We drop the superscript β assuming a stable rate. So, we have $\boldsymbol{\eta} = [\boldsymbol{\eta}(1)\ \boldsymbol{\eta}(2)\ \cdots\ \boldsymbol{\eta}(n-1)\ 0]$. This is defined assuming that sink collects all data it receives and has no notion of maintaining queue. Let $\boldsymbol{J}$ be another n element column vector such that $\boldsymbol{J}(i) > 0$ if $i \in V_s$, $\boldsymbol{J}(u_{\mathrm{s}}) = -\sum_{i\in V_s}\boldsymbol{J}(i)$ and 0 elsewhere, and I be the usual $n \times n$ identity matrix. So, given the transition matrix P_w for the random walk defined by w on graph G, the steady-state queue equations at the nodes can be written in matrix form as

$$\boldsymbol{\eta}^T(I - P_w) = \beta \boldsymbol{J}^T. \tag{15}$$

As we know transition matrix $P_w = D^{-1}A$ where D is the diagonal matrix of generalized degrees and A is the adjacency matrix, so matrix $(I - P_w)$ is also a Laplacian as we can rewrite it as $(I - P_w) = D^{-1}(D - A) = D^{-1}L$. So, the above equation (Eq. (15)) can be rewritten as

$$\boldsymbol{x}^T L = \beta \boldsymbol{J}^T \tag{16}$$

where $\boldsymbol{x}^T = \boldsymbol{\eta}^T D^{-1}$ is a row vector such that $\boldsymbol{x}(u) = \boldsymbol{\eta}(u)/\deg(u)$ for all u where $\boldsymbol{\eta}(u)$ is the steady-state queue occupancy probability and $\deg(u) = \sum_{v:(u,v)\in E} w_{uv}$ is the generalized degree of node u. Eq. (16) is similar to Laplacian systems of the form $L\boldsymbol{x} = \boldsymbol{b}$ with a constraint that only one element in $\boldsymbol{b}$ is negative. We call such systems "one-sink" Laplacian systems. In our subsequent work [8][9] we discuss this connection in detail.

4.2. A lower bound

Now having established the steady-state equation for the Data Collection Process, we will use it for characterizing the critical data rate. In particular, we will prove a lower bound on such rate.

Proof of Theorem 2. For a given graph $G = (V, E, w)$, with source set $V_s \subseteq V \setminus \{u_{\rm s}\}$ and transition matrix P_w for random walk defined by w on graph G, recall that the steady-state queue equations at nodes can be written in matrix form as

$$\boldsymbol{\eta}^T(I - P_w) = \beta \boldsymbol{J}^T. \tag{17}$$

Now, in order to bound the maximum stable data rate β at which the source nodes generate data in terms of the underlying graph parameters, we will consider eigendecomposition of the left hand side of Eq. (17). For this, we will deviate from the usual inner product on the vector space $\mathbb{R}^V$ i.e., $\langle f, g\rangle = \sum_{x\in V} f(x)g(x)$ and define another inner product on $\mathbb{R}^V$ which is given by $\langle f, g\rangle_\mu := \sum_{x\in V} f(x)g(x)\mu(x)$ where μ is the stationary distribution of random walk defined by w on graph satisfying $\mu = \mu P_w$. From Lemma 12.2 [19], it is known that the inner product space $(\mathbb{R}^V, \langle\cdot,\cdot\rangle)_\mu$ has an orthonormal basis of real-valued eigenfunctions $\{f_j\}_{j=1}^{|V|}$ corresponding to real eigenvalues $\{\lambda_j\}$. Using this lemma and writing the vector $\boldsymbol{\eta}^T$ in terms of the eigenvectors, we have $\boldsymbol{\eta}^T = \sum_{i=1}^{|V|} \langle \boldsymbol{\eta}^T, f_i \rangle_\mu f_i$. This gives us that $\boldsymbol{\eta}^T(I - P_w) = \sum_{i=1}^{|V|}(1-\lambda_i^w)\langle \boldsymbol{\eta}^T, f_i\rangle_\mu f_i$, where λ_i^w is the i^{th} eigenvalue of transition matrix P_w. Moreover, from Lemma 12.1 of [19], we also know that the absolute value of any eignevalue of a transition matrix can be at most 1, so, $\lambda_1^w = 1 > \lambda_2^w \geq \cdots \geq \lambda_n^w$. So, we have

$$\boldsymbol{\eta}^T(I - P_w) = \sum_{i=2}^{|V|}(1 - \lambda_i^w)\langle \boldsymbol{\eta}^T, f_i\rangle_\mu f_i \tag{18}$$

$$\geq (1 - \lambda_2^w)\left(\sum_{i=2}^{|V|}\langle \boldsymbol{\eta}^T, f_i\rangle_\mu f_i\right). \tag{19}$$

Note, that $f_1, \ldots, f_{|V|}$ form an orthonormal basis so, $\sum_{i=1}^{|V|}\langle \boldsymbol{\eta}^T, f_i\rangle_\mu^2 = \|\boldsymbol{\eta}^T\|_\mu^2$. Hence we have

$$\sum_{i=2}^{n}\langle \boldsymbol{\eta}^T, f_i\rangle_\mu^2 = \|\boldsymbol{\eta}^T\|_\mu^2 - \langle \boldsymbol{\eta}^T, f_1\rangle_\mu^2. \tag{20}$$

The eigenfunction f_1 corresponding to the eigenvalue 1 can be taken to be a constant vector $\mathbf{1}$, so $\langle \boldsymbol{\eta}^T, f_1\rangle_\mu = \sum_{i=1}^{n} \boldsymbol{\eta}(i)\mu(i)$, where $\mu(i) = \sum_{v\in V}\mu(v)P_w[v,i]$. Also, $\|\boldsymbol{\eta}^T\|_\mu^2 = \sum_{i=1}^{n}\boldsymbol{\eta}^2(i)\mu(i)$. So, using these results in Eq. (20) we have

$$\sum_{i=2}^{n}\langle \boldsymbol{\eta}^T, f_i\rangle_\mu^2 = \sum_{i=1}^{n}\boldsymbol{\eta}^2(i)\mu(i) - \left(\sum_{i=1}^{n}\boldsymbol{\eta}(i)\mu(i)\right)^2 = \mathrm{Var}_\mu(\boldsymbol{\eta}(i)) = \sum_{i=1}^{n}(\boldsymbol{\eta}(i) - \bar{\boldsymbol{\eta}}_\mu)^2\mu(i) \tag{21}$$

where, $\bar{\boldsymbol{\eta}}_\mu = \sum_{i=1}^{n}\boldsymbol{\eta}(i)\mu(i)$ is the expected queue occupancy probability of nodes under stationary distribution μ. Now, taking square of norm of Eq. (19) and using Eq. (21), we have

$$\|\boldsymbol{\eta}^T(I - P_w)\|_\mu^2 \geq (1 - \lambda_2^w)^2\mathrm{Var}_\mu(\boldsymbol{\eta}(i)). \tag{22}$$

Using Eq. (22) in the square of norm of Eq. (17), we have

$$\beta \geq \frac{(1 - \lambda_2^w)}{\|\boldsymbol{J}^T\|_\mu}\sqrt{\mathrm{Var}_\mu(\boldsymbol{\eta}(i))}. \tag{23}$$

Moreover, as $\sum_{i\in V_s}\boldsymbol{J}^2(i) \leq (\sum_{i\in V_s}\boldsymbol{J}(i))^2$, so we have

$$\|\boldsymbol{J}^T\|_\mu = \sqrt{\sum_{i\in V_s}\boldsymbol{J}^2(i)\mu(i) + \left(\sum_{i\in V_s}\boldsymbol{J}(i)\right)^2\mu(u_\mathrm{s})} \leq \sum_{i\in V_s}\boldsymbol{J}(i)\sqrt{\mu_m + \mu(u_\mathrm{s})} \tag{24}$$

where $\mu_m = \max_{i \in V_s} \mu(i)$

Now to get a bound on $\mathrm{Var}_\mu(\boldsymbol{\eta}(i)) = \sum_{i=1}^{n} (\boldsymbol{\eta}(i) - \bar{\boldsymbol{\eta}}_\mu)^2 \mu(i)$, we consider two nodes whose queue occupancy probability we know precisely (1) the sink, u_s, which has $\boldsymbol{\eta}(u_\mathrm{s}) = 0$ (as it has no notion of maintaining queue and it sinks data packets as soon as it receives them), and (2) a node $u_{\max}$ with maximum queue occupancy probability for a given β, let it be $\boldsymbol{\eta}^\beta_{\max} = \max_{u \in V \setminus \{u_\mathrm{s}\}} \boldsymbol{\eta}^\beta(u)$. Now, let $\beta = (1-\delta)\beta^*$ where β^* is the critical data rate and $\delta \in (0,1)$. From Eq. (15) we know $\boldsymbol{\eta}$ is linear in β and $\boldsymbol{\eta}^{\beta *}_{\max} = 1$, so $\boldsymbol{\eta}^\beta_{\max} = \frac{\beta}{\beta^*}$ and hence, we have $\boldsymbol{\eta}^\beta_{\max} = 1 - \delta$.

Let $\bar{\boldsymbol{\eta}}_\mu = \sum_{i=1}^n \boldsymbol{\eta}(i)\mu(i)$ be the expected queue occupancy probability of nodes under the stationary distribution μ. We can bound the variance $\mathrm{Var}_\mu(\boldsymbol{\eta}(i))$ using the contributions of the nodes u_s and $u_{\max}$ as follows.

$$\mathrm{Var}_\mu(\boldsymbol{\eta}(i)) \geq \left(1 - \delta - \bar{\boldsymbol{\eta}}_\mu\right)^2 \mu(u_{\max}) + \left(\bar{\boldsymbol{\eta}}_\mu - 0)^2\right) \mu(u_\mathrm{s}) \geq \frac{(1-\delta)^2 \mu(u_{\max})\mu(u_\mathrm{s})}{\mu(u_{\max}) + \mu(u_\mathrm{s})}. \quad (25)$$

where the last inequality holds as $(1 - \delta - \bar{\boldsymbol{\eta}}_\mu)^2 \mu_{u_{\max}} + (\bar{\boldsymbol{\eta}}_\mu - 0)^2 \mu_{u_\mathrm{s}}$ achieves optimum at $\bar{\boldsymbol{\eta}}_\mu = \frac{(1-\delta)\mu_{u_{\max}}}{\mu_{u_{\max}} + \mu_{u_\mathrm{s}}}$. Now, using the fact that when $\beta \to \beta^*$, $\delta \to 0$ in Eq. (25) and then using the resultant bound on $\mathrm{Var}_\mu(\boldsymbol{\eta}_i)$ and the value of $\|\boldsymbol{J}^T\|_\mu$ (Eq. (24)) in Eq. (23) we have

$$\beta^* \geq \frac{(1 - \lambda_2^w)}{\sum_{i \in V_s} \boldsymbol{J}(i)} \sqrt{\frac{\mu(u_{\max})\mu(u_\mathrm{s})}{(\mu(u_{\max}) + \mu(u_\mathrm{s}))(\mu_m + \mu(u_\mathrm{s}))}}. \quad (26)$$

Now, we know $\mu(i) = \frac{\deg(i)}{\sum_{u \in V} \deg(u)}$, and $\frac{d_{\min}}{\sum_{u \in V} \deg(u)} \leq \mu(i) \leq \frac{d_{\max}}{\sum_{u \in V} \deg(u)}$ where, $d_{\min}$ and $d_{\max}$ are the generalized minimum and maximum degrees of graph respectively. So using the appropriate bounds on $\mu(i)$ in Eq. (26) we have

$$\beta^* \geq \frac{(1 - \lambda_2^w)}{\sum_{i \in V_s} \boldsymbol{J}(i)} \frac{\sqrt{d_{\min} \deg(u_\mathrm{s})}}{(d_{\max} + \deg(u_\mathrm{s}))} \quad (27)$$

where λ_2^w is the second smallest eigenvalue of the transition matrix of random walk defined by the weight function w and $\deg(u_\mathrm{s})$ is the generalized degree of the sink node. □

4.3. An upper bound

We also prove an upper bound on β^* for the case where $V_s = V \setminus \{u_\mathrm{s}\}$. Here is the upper bound result.

Proof of Proposition 1. Given any vertex $u \in V$, recall its measure is defined as, $\rho(u) := \sum_{v \in V} P_w[u,v]$, and for any $U \subset V$ we have $\rho(U) = \sum_{u \in U} \rho(u)$. Similarly, for edge boundary as $\partial U := \{(u,v) : u \in U, v \notin U\}$, we have $\rho(\partial U) = \sum_{u \in U, v \notin U} P_w[u,v]$. Now, let us define constants $h(U) := \frac{\rho(\partial U)}{\rho(U)}$ and $\hat{h}(G) := \min_{U \subset V, u_{\mathrm{s}} \notin U} h(U) \leq h(G)$ where $h(G)$ is the edge expansion of graph G.

We know, for any given set $U \subset V$, where $u_{\mathrm{s}} \notin U$ the maximum data flow that can move out of this set is the flow across the boundary ∂U, so

$$\beta \rho(U) \leq \rho(\partial U) \tag{28}$$

$$\beta \leq \min_{U} h(U) = \hat{h}(G) \leq h(G). \tag{29}$$

Now, for set $U = V_s = V \setminus \{u_{\mathrm{s}}\}$, we have $\hat{h}(G)) \leq \sum_{u:u \sim u_{\mathrm{s}}} \frac{P_w[u,u_{\mathrm{s}}]}{n-1}$. So, from eq. (28) $\beta \leq \sum_{u:u \sim u_{\mathrm{s}}} \frac{P_w[u,u_{\mathrm{s}}]}{n-1}$. Hence, the upper bound on the critical data rate is given by,

$$\beta \leq \min \left\{ \hat{h}(G), \sum_{u:u \sim u_{\mathrm{s}}} \frac{P_w[u,u_{\mathrm{s}}]}{n-1} \right\}. \tag{30}$$

□

Note that our derived upper and lower bound on the critical data rate relates directly to the two sides of Cheeger's inequality [4].

5. Geometric rate of convergence

Next, we characterize the rate of convergence of Markov chain $\left\{ Q_t^{\boldsymbol{J},\beta} \right\}_{t \geq 0}$ for the stable regime i.e., $\beta < \beta^*$. In particular, we first prove a general result about the total variation distance between the probability distributions of two Markov chains and their rate of convergence. Then, as a special case of this result we show that the convergence of Markov chain $\left\{ Q_t^{\boldsymbol{J},\beta} \right\}_{t \geq 0}$ is geometric, i.e., starting from any initial state, the distance from the stationarity reduces exponentially. Note that we drop the superscript $\boldsymbol{J}, \beta$ from the Markov chain representation as a stable data rate value for proving the convergence rate is assumed.

Proof of Theorem 3. We first note that our Markov chain Q_t is *stochastically ordered* (c.f. [22]). To understand what this means we define a natural partial order on $(\mathbb{N} \cup$

$\{0\})^{|V\setminus\{u_s\}|}$ as follows: $\boldsymbol{x} \preceq \boldsymbol{y}$ if $\boldsymbol{x}_v \leq \boldsymbol{y}_v$ for all $v \in V \setminus \{u_s\}$. A function f : $(\mathbb{N} \cup \{0\})^{|V\setminus\{u_s\}|} \to \mathbb{R}$ is said to be increasing if $\boldsymbol{x} \preceq \boldsymbol{y}$ implies that $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$. Given two random processes X and Y supported on $(\mathbb{N}\cup\{0\})^{|V\setminus\{u_s\}|}$ or $(\mathbb{N}\cup\{0\})^{|V|-1}$, we say X is stochastically dominated by Y if $\mathbb{E}\left[f(X)\right] \leq \mathbb{E}\left[f(Y)\right]$ for every increasing function f. For our Data Collection chain we state the stochastic orderedness property as follows.

Claim 1. *Given two instances of the Data Collection Process Q_t and Q'_t such that $Q_0 \preceq Q'_0$, Q_t is stochastically dominated by $Q'_t, t \geq 0$. In particular this means that $\mathbb{P}\left[Q_t(v) > 0\right] \leq \mathbb{P}\left[Q'_t(v) > 0\right]$ for all $v \in V \setminus \{u_s\}$.*

The proof of this claim follows by constructing a coupling between the two chains such that each of them performs exactly the same transmission actions. In case one of the chains is empty then the transmission action is a dummy action. It is easy to see that stochastic ordering follows naturally for the Data Collection chain.

To use this claim, for our irreducible and aperiodic Markov chain Q_t described by the Data Collection Process defined on $(\mathbb{N}\cup\{0\})^{|V|-1}$ having transition matrix $\mathcal{P}$ and a stationary distribution π, let us define two other irreducible and aperiodic Markov chains Q^1_t and Q^2_t, each with state space $(\mathbb{N}\cup\{0\})^{|V|-1}$. Initially, suppose the data is generated in the two chains in a coupled way such that one of them dominates the other, i.e., either $Q^1_0(v) \leq Q^2_0(v)$ for all $v \in V \setminus \{u_s\}$ or vice-versa.

Now, consider the coupling (Q^1_t, Q^2_t) on $(\mathbb{N}\cup\{0\})^{|V|-1} \times (\mathbb{N}\cup\{0\})^{|V|-1}$ defined over random sequences $\{0,1\} \times \{\prod_{v\in V\setminus\{u_s\}} \Gamma(v)\}$ where $\Gamma(v)$ is the set of one-step destinations from node v, such that both the chains Q^1_t and Q^2_t are populated in a coupled way. Such Markov chains are said to be stochastically ordered chains in the queueing theory and have a property that the Markov chain which dominates the other chain will always maintain dominance over it.

Under this coupling we allow the two chains to run in a way that any data generation or data transmission decision made by any queue in one chain is followed by the corresponding queue in the other chain as well. However, to distinguish the newly generated packets in two chains from the existing ones, we assign colors to the data packets: the existing packets in Q^1_t chain are colored red and in Q^2_t chain are colored blue, and the newly generated packets in both the chains are colored green. Moreover, in both the chains green (newly generated) packets get a preference in the transmission. Now, let $Q_t^{1,green}(u)$ represent the number of green packets in the queue of a given node u in Q_t^1 and $\boldsymbol{\eta}_u$ be the steady-state queue occupancy probability of Markov chain Q_t. Since, the number of green packets in both the chains starts from zero and the chains are stochastically ordered, green packet queue occupancy is always bounded by that of the chain with stationary distribution i.e., $\mathbb{P}\left[Q_t^{1,green}(u) \geq 1\right] \leq \boldsymbol{\eta}_u$. Same holds true for the other chain Q_t^2 as well.

To ensure both chains get coupled all the red and blue (old) packets in Q_t^1 and Q_t^2 respectively need to be sunk. We consider Q_t^1 chain and the same will hold for Q_t^2 as well. We know by our preference in transmission, the probability that red packets move out of queue in one time step in Q_t^1 is equal to the probability that there are no green packets in the given queue, i.e., $1 - \mathbb{P}\left[Q_t^{1,green}(u) \geq 1\right]$. Also, we have $1 - \mathbb{P}\left[Q_t^{1,green}(u) \geq 1\right] \geq 1 - \boldsymbol{\eta}_u \geq \min_u 1 - \boldsymbol{\eta}_u \geq 1 - \boldsymbol{\eta}_{max}$, where $\boldsymbol{\eta}_{max} = \max_{u \in V \setminus \{u_s\}} \boldsymbol{\eta}_u$. Now, let $N^{(red)}$ and $N^{(blue)}$ be the total number of red and blue data packets in chains Q_t^1 and Q_t^2 respectively at the beginning which are assumed to be finite. Also, let $T_{N^{(red)}}$ and $T_{N^{(blue)}}$ be the time taken by the the respective number of packets to get sunk. We have the following lemma that bounds this time.

Lemma 4. *Given a Data Collection Process on graph G with $N^{(*)} < \infty$ as the total number of data packets present in the queues of all nodes initially, then the time taken by all such packets to reach the sink, let it be $T_{N^{(*)}}$ is bounded as*

$$\mathbb{P}\left[T_{N^{(*)}} \geq \frac{2t_{hit}}{1 - \boldsymbol{\eta}_{\max}}(\log 1/\epsilon + 1) + N^{(*)} - 1\right] \leq \frac{\epsilon}{2} \tag{31}$$

where t_{hit} is the worst-case hitting time of random walk on G and $\boldsymbol{\eta}_{\max}$ is the maximum queue occupancy probability at stationarity.

The proof of the lemma proceeds by coupling the Data Collection Process to a random walk with the property that the time taken by this random walk to hit u_s is at most $N^{(*)}$ steps less than the time for the Data Collection Process to sink all $N^{(*)}$ packets. The result then follows by bounding the hitting time of this random walk.

Proof of Lemma 4. Let us say that the set of $N^{(*)}$ initial packets has color red and the new packets generated thereafter are green in color. And let us say that the

Markov chain Q_t^* describes the Data Collection Process that begun with the set of initial packets at some initial locations.

We now define three $N^{(*)}$-dimensional processes based on Q_t^*.

- V_t^1: $V_t^1(i) \in V$ is the position of red packet p_i at step t in Q_t^*.
- V_t^2: Here there are no green packets delaying the red packets, but at each time step each vertex transmits no packet with probability $\boldsymbol{\eta}_{\max}$, the maximum queue occupancy probability of Q_t^* at stationarity. $V_t^2(i)$ again gives the position of the i-th red packet at time t.
- V_t^3: In this process, we have only red packets but all of them have been assigned distinct deterministically chosen ranks, let's say $1, \ldots, N^{(*)}$. Under this process, at any queue the choice of packet to send is deterministic: the packet with the lowest rank is chosen for transmission. Other than the choice of packet, all other transmission decisions are similar to V_t^2. $V_t^3(i)$ again gives the position of the i-th red packet at time t.

Note that all three processes are absorbed by the state $(u_{\mathrm{s}}, \ldots, u_{\mathrm{s}})$, i.e., when all red data packets get sunk. Let τ_i be the time to absorption of the process V_t^i, $1 \le i \le 3$. Then for every $k \ge 0$, using a simple coupling argument we can prove

$$\mathbb{P}\left[\tau_1 \ge k\right] \le \mathbb{P}\left[\tau_2 \ge k\right] = \mathbb{P}\left[\tau_3 \ge k\right]. \tag{32}$$

Now for V_t^3, let us define a collection of random variables $\{Y_t\}_{t\ge 0}$ by constructing what is refered to as a *delay sequence* in [17]. We set $Y_0 = u_{\mathrm{s}}$. Let packet p_1 be the last packet to be absorbed in the sink in V_t^3. If p_1 came to the sink from node u then $Y_1 = u$. This way, we move back $\ell_1 \ge 1$ steps till we reach the first queue where p_1 was delayed because a lower ranked packet, say p_2 was preferred for transmission. Let's say this node is v_1, i.e., $Y_{\ell_1} = v_1$. After this, we set Y_{ℓ_1+1} to the node from which the preferred packet p_2 came to v_1. Now we trace p_2 back till we reach the node v_2 where it was delayed by another packet p_3. Let ℓ_2 represent the number of hops encountered during this trace. Similarly, we continue this process till we reach a packet p_k which was never delayed and can be traced back to its initial position. At this point, we have defined $Y_0, \ldots, Y_D$ where $D = \sum_{i=1}^{k-1} \ell_i + \ell_k$ where ℓ_i is the number of hops travelled by packet p_i till it reached a node where it got delayed by packet p_{i+1}, and ℓ_k is the number of hops travelled by packet p_k till it reaches its initial position v_k.

Now consider the collection of random variables $\{X_t\}_{t=0}^{D}$ with $X_t = Y_{D-t}$ for $0 \leq t \leq D$. From the way the Y_ts are constructed, it is clear that $\{X_t\}_{t=0}^{D}$ is a random walk beginning at $X_0 = Y_D$ followed till it hits u_{s}. This is a lazy walk that stays at the same vertex with transition matrix $\boldsymbol{\eta}_{\max} I + (1-\boldsymbol{\eta}_{\max})\mathcal{P}$. Let τ be the time taken for a random walk starting at any node to hit u_{s} under the Data Collection Process. We have the following claim relating this hitting time to the absorption time of process V_t^3.

Claim 2. *For all $L \geq 1$, $\mathbb{P}\left[\tau_3 \geq L + N^{(*)} - 1\right] \leq \mathbb{P}\left[\tau \geq L\right]$.*

Proof of Claim 2. Consider a trace of V_t^3 and the corresponding random walk $\{X_t\}$. We will synchronize the progress of a single packet in V_t^3 to the progress of $\{X_t\}$, i.e., at every time t we will identify a packet p_i and an $0 \leq s \leq k$ such that $X_t = V_{t+s}^3(i)$. The particular packet chosen for synchronization will change as we move along. The first candidate will be p_k, the last packet in the delay sequence. Note that the last packet p_k in the delay sequence makes its first move at same time for both V_t^3 and $\{X_t\}$, i.e., $V_0^3(k) = X_0 = v_k$, where v_k is the initial position of packet p_k. So, both the process and the random walk are synchronized with $s = 0$ in the beginning and will remain synchronized till ℓ_k hops, i.e., $V_{\ell_k}^3(k) = X_{\ell_k} = v_{k-1}$, where packet p_k delays packet p_{k-1} by one unit. After this, our choice of packet to follow will switch to p_{k-1}. Since the random walk follows packet p_{k-1}'s progress so it is delayed by one unit while the V_t^3 process keeps on moving, i.e., the two are synchronized with $s = 1$. Continuing in this way we eventually synchronize $\{X_t\}$ with p_1 with $s = k-1$. So, we have, by the coupling of V_t^3 and X_t that

$$\mathbb{P}\left[\tau_3 \geq L + k - 1\right] \leq \mathbb{P}\left[\tau \geq L\right]. \tag{33}$$

Since k, the number of packets encountered in the delay sequence is a random variable upper bounded by $N^{(*)}$ the claim follows. □

Since $\{X_t\}$ is a lazy walk, the time taken to hit u_{s} is at most $\frac{t_{\mathrm{hit}}}{1-\boldsymbol{\eta}_{\max}}$ where t_{hit} is the hitting time for the random walk associated with $\mathcal{P}$. So, by Markov's inequality $\mathbb{P}\left[\tau \geq \frac{2t_{\mathrm{hit}}}{1-\boldsymbol{\eta}_{\max}}\right] \leq \frac{1}{2}$. Now, consider the probability of a random walk not hitting the sink u_{s} in $2(\log 1/\epsilon + 1)$ times $\frac{t_{\mathrm{hit}}}{1-\boldsymbol{\eta}_{\max}}$, i.e., we consider $\frac{2t_{\mathrm{hit}}}{1-\boldsymbol{\eta}_{\max}}(\log 1/\epsilon + 1)$ time and divide it into $(\log 1/\epsilon + 1)$ slots of $\frac{2t_{\text{hit}}}{1-\boldsymbol{\eta}_{\max}}$ each. By the Markov property of random walks, we know that the random walks in each of these slots are independent. So, we have the following result.

$$\mathbb{P}\left[\tau \geq \frac{2t_{\text{hit}}}{1-\boldsymbol{\eta}_{\max}}(\log 1/\epsilon + 1)\right] \leq \frac{\epsilon}{2}. \tag{34}$$

Finally, using Eq. (32) in Claim 2 we get the result. □

Now, since both the chains Q_t^1 and Q_t^2 operate in parallel, the expected time for the two chains to couple i.e., all red and blue packets get sunk is the maximum of the time taken by each to get their respective packets sunk. So, using Lemma 4 for both the chains we have the expected time for Q_t^1, Q_t^2 to couple, let it be $\tau_{couple}^{1,2} = \max\{T_{N^{(red)}}, T_{N^{(blue)}}\}$ as

$$\mathbb{P}\left[\tau_{couple}^{1,2} \geq \frac{2t_{\text{hit}}}{1-\boldsymbol{\eta}_{\max}}\left(\log\frac{1}{\epsilon}+1\right) + \max\{N^{(red)}, N^{(blue)}\} - 1\right] \leq \epsilon. \tag{35}$$

Note that this expected coupling time result is similar to the delay result of Leighton et al. [18, 17] depicting the pipelining behaviour of Data Collection Process.

Now, to bound the distance between the two chains Q_t^1 and Q_t^2 we use the following result from Levin et al. [19].

Lemma 5. (Theorem 5.2, Levin et al. [19].) *Let $\{(X_t, Y_t)\}$ be a coupling with initial states $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ such that $X_0 = \boldsymbol{x}$ and $Y_0 = \boldsymbol{y}$ and coupling time defined as $\tau_{couple} := \min\{t : X_s = Y_s \text{ for all } s \geq t\}$, then,*

$$||\mathcal{P}^t[\boldsymbol{x}, \cdot] - \mathcal{P}^t[\boldsymbol{y}, \cdot]||_{TV} \leq P_{\boldsymbol{x},\boldsymbol{y}}\{\tau_{couple} > t\}.$$

Let $\boldsymbol{x}, \boldsymbol{y} \in (\mathbb{N} \cup \{0\})^{|V|-1}$ be the initial states of Q_t^1 and Q_t^2 chain then for $||\mathcal{P}^t[\boldsymbol{x}, \cdot] - \mathcal{P}^t[\boldsymbol{y}, \cdot]||_{TV} \leq \epsilon$, using Lemma 5 and the expected coupling time from Eq. (35) we have

$$||\mathcal{P}^t[\boldsymbol{x}, \cdot] - \mathcal{P}^t[\boldsymbol{y}, \cdot]||_{TV} \leq 2^{\left(1+\frac{\max\{N^{(red)}, N^{(blue)}\}-1}{\frac{2t_{\text{hit}}}{1-\boldsymbol{\eta}_{\max}}}\right)} \cdot \left(\frac{1}{2}\right)^{\frac{t}{\frac{2t_{\text{hit}}}{1-\boldsymbol{\eta}_{\max}}}}. \tag{36}$$

Now, assume the stable data rate at which we are running these stochastic processes is $\beta = (1-\delta)\beta^*$ where β^* is the critical data rate and $\delta \in (0,1)$. From Eq. (15) we know $\boldsymbol{\eta}$ is linear in β. Also, for unstable data rates there exists a node whose steady-state queue occupancy is 1 (see [23] and Lemma 5 of [7]), i.e., $\boldsymbol{\eta}_{\max}^{\beta^*} = 1$, so we have $\boldsymbol{\eta}_{\max}^{\beta} = \frac{\beta}{\beta^*}$, hence, $1 - \boldsymbol{\eta}_{\max}^{\beta} = \delta$. Using this in Eq. (36) we prove the desired result. □

To use Theorem 3 to prove the geometric ergodicity result (Corollary 1) we pick $\boldsymbol{y}$ according to the stationary distribution π of the Data Collection Process Markov chain.

Proof of Corollary 1. Let us consider two instances of Data Collection Process Q_t^1 and Q_t^2 such that the former starts from some finite state $\boldsymbol{x} \in (\mathbb{N} \cup \{0\})^{|V|-1}$ and the latter starts from stationarity, i.e., initially all queues in Q_t^1 are occupied by some finite number of packets and that of Q_t^2 are filled according to the stationary distribution π. Then, from Theorem 3 we have

$$||\mathcal{P}^t[\boldsymbol{x}, \cdot] - \pi||_{TV} \leq 2^{\left(1+\frac{\left(\max\{N^{(\boldsymbol{x})}, N^{(\pi)}\}-1\right)\delta}{2t_{\text{hit}}}\right)} \cdot \left(\frac{1}{2}\right)^{\frac{\delta}{2t_{\text{hit}}} \cdot t} \tag{37}$$

where t_{hit} is the worst-case hitting time of random walk on graph, $N^{(\boldsymbol{x})}$ and $N^{(\pi)}$ are the total number of data packets in state $\boldsymbol{x}$ and at stationarity respectively and δ is the relative distance from the critical data rate. Now, if we compare Eq. (37) with the Definition 2 (Eq. (5)) we prove geometric ergodicity property for the Markov chain $Q_t^{\boldsymbol{J},\beta}$.

Now for random variable $N^{(\pi)}$, let $\mathbb{E}\left[N^{(\pi)}\right]$ be its expectation, i.e., the expected number of data packets in $Q_t^{\boldsymbol{J},\beta}$ at stationarity which by Little's law [20] is equal to the product of the data generation rate and the expected latency of a data packet to reach the sink at the stationarity, i.e., $\mathbb{E}\left[N^{(\pi)}\right] = \frac{\beta t_{\text{hit}}}{1-\boldsymbol{\eta}^{\beta}_{\max}} = \frac{(1-\delta)\beta^* t_{\text{hit}}}{\delta}$ (from linearity of $\boldsymbol{\eta}$ and $\beta = (1-\delta)\beta^*$) where β^* is the critical data rate and $\delta \in (0, 1)$. Now, let $\epsilon_{\boldsymbol{x}} = \max\left\{\alpha \in [0,1] : N^{(\boldsymbol{x})} \leq \frac{(1-\delta)\beta^* t_{\text{hit}}}{\delta} \cdot \left(\log \frac{1}{\alpha} + 1\right)\right\}$. So by the definition of $\epsilon_{\boldsymbol{x}}$ we have two regimes: $\epsilon \leq \epsilon_{\boldsymbol{x}}$ where the $\mathbb{E}\left[N^{(\pi)}\right]$ term is dominant and $\epsilon > \epsilon_{\boldsymbol{x}}$ where the $N^{(\boldsymbol{x})}$ is dominant.

For the simple case of $\epsilon > \epsilon_{\boldsymbol{x}}$, using Eq. (37) we have

$$||\mathcal{P}^t[\boldsymbol{x}, \cdot] - \pi||_{TV} \leq 2^{\left(1+\frac{N^{(\boldsymbol{x})}\delta}{2t_{\text{hit}}}\right)} \cdot \left(\frac{1}{2}\right)^{\frac{\delta}{2t_{\text{hit}}} \cdot t}. \tag{38}$$

Similarly for $\epsilon \leq \epsilon_{\boldsymbol{x}}$ we have

$$||\mathcal{P}^t[\boldsymbol{x}, \cdot] - \pi||_{TV} \leq 2^{\left(1+\frac{(1-\delta)\beta^*}{2}(\log 1/\epsilon+1)\right)} \cdot \left(\frac{1}{2}\right)^{\frac{\delta}{2t_{\text{hit}}} \cdot t}. \tag{39}$$

Setting the RHS of Eq. (39) to ϵ and solving for t we get that

$$||\mathcal{P}^t[x, \cdot] - \pi||_{TV} \leq 2 \cdot \left(\frac{1}{2}\right)^{\frac{\delta}{t_{\text{hit}}((1-\delta)\beta^*+2)} \cdot t}. \tag{40}$$

Combining (38) and (40) gives us the result.

We observe that if we set $\boldsymbol{x}$ to $\mathbf{0}$ (all zeros), i.e., all queues are initially empty, then $\epsilon_{\mathbf{0}}$ is 1 so only Eq. (40) applies and we determine the mixing time by setting the RHS to $1/M$ for a given value of $M > 0$. □

6. Some future directions

The fact that the Data Collection Process mixes fast to its stationary distribution when started from the all-empty setting can be exploited to solve systems of equations such as Eq. (16) simply by allowing the process to get close enough to stationarity and then estimate the $\boldsymbol{\eta}$ by keeping track of the number of time slots for which each queue is occupied. This opens up the possibilities of distributed algorithms for effective resistance and other problems, some of which we have explored in [8][9][10]. Even if we consider graph problems on very large graphs, Laplacian systems of equations become tractable via this method since random walks can be simulated very fast in modern computing systems for graphs with nodes in the millions (see, e.g., [27]).

The key shortcoming of our work is that the Data Collection Process in the subcritical region models only one-sink Laplacian systems of equations. A model that captures the full generality of Laplacian systems of equations will open a more general class of problems that can be attacked algorithmically using this method.

Acknowledgements

The authors would like to thank Tanmay Gupta and Abhishek Patel for their contributions to this work. We would also like to thank Prof. Mark Yarrow and Prof. Peter Taylor for their support.

References

[1] Becchetti, L., Bonifaci, V. and Natale, E. (2018). Pooling or sampling: Collective dynamics for electrical flow estimation. In *Proc. of the 17th Intl. Conf. on Autonomous Agents and MultiAgent Systems.* AAMAS '18. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC.

pp. 1576–1584.

[2] Boyd, S., Ghosh, A., Prabhakar, B. and Shah, D. (2005). Gossip algorithms: Design, analysis and applications. In *Proc. of the 24th Annual Joint Conf. of the IEEE Computer and Comm. Societies.* INFOCOM '05. IEEE. pp. 1653–1664 vol. 3.

[3] Brightwell, G. and Winkler, P. (1990). Maximum hitting time for random walks on graphs. *Random Structures & Algorithms* **1,** 263–276.

[4] Cheeger, J. (1969). A lower bound for the smallest eigenvalue of the Laplacian. In *Proc. of the Princeton conference in honor of Professor S. Bochner*. pp. 195–199.

[5] Cohen, M. B., Kyng, R., Miller, G. L., Pachocki, J. W., Peng, R., Rao, A. B. and Xu, S. C. (2014). Solving SDD linear systems in nearly $m\ log^{1/2}n$ time. In *Proc. of the 46th annual ACM Symp. on Theory of Computing.* STOC '14. ACM. pp. 343–352.

[6] Doyle, P. and Snell, J. Random walks and electric networks. mathematical association of america, washington, dc 1984.

[7] Georgiadis, L. and Szpankowski, W. (1992). Stability of token passing rings. *Queueing systems* **11,** 7–33.

[8] Gillani, I. A. and Bagchi, A. (2020). A queueing network-based distributed Laplacian solver. In *Proc. of the 32nd ACM Symp. on Parallelism in Algorithms and Architectures.* SPAA '20. ACM, New York, NY, USA. p. 535–537.

[9] Gillani, I. A. and Bagchi, A. (2021). A queueing network-based distributed Laplacian solver for directed graphs. *Inf. Proc. Lett.* **166,** 106040.

[10] Gillani, I. A., Bagchi, A. and Ranu, S. (2021). A group-to-group version of random walk betweenness centrality. In *Proc. of 8th ACM IKDD CODS and 26th COMAD.* CODS-COMAD '21. pp. 127–135.

[11] Gillani, I. A., Vyavahare, P. and Bagchi, A. (2021). Lower bounds for in-network computation of arbitrary functions. *Distrib. Comput.* https://doi.org/10.1007/s00446-021-00394-7.

[12] Kamra, A., Misra, V., Feldman, J. and Rubenstein, D. (2006). Growth codes: Maximizing sensor network data persistence. In *Proc. of the 2006 Conf. on Applications, Technologies, Architectures, and Protocols for Computer Comm.*. SIGCOMM '06. ACM. pp. 255–266.

[13] Kempe, D., Dobra, A. and Gehrke, J. (2003). Gossip-based computation of aggregate information. In *Proc. of the 44th Annual IEEE Symp. on Foundations of Computer Science*. FOCS '03. IEEE. pp. 482–491.

[14] Kleinrock, L. (1975). Queueing systems. *John Wiley & Sons, Inc.* **6,**.

[15] Koutis, I. and Xu, S. C. (2016). Simple parallel and distributed algorithms for spectral graph sparsification. *ACM Trans. Parallel Comput.* **3,**.

[16] Kyng, R. and Sachdeva, S. (2016). Approximate gaussian elimination for laplacians-fast, sparse, and simple. In *Proc. of the 57th Annual IEEE Symp. on Foundations of Computer Science*. FOCS '16 IEEE. pp. 573–582.

[17] Leighton, F. T., Maggs, B. M., Ranade, A. G. and Rao, S. B. (1994). Randomized routing and sorting on fixed-connection networks. *Journal of Algorithms* **17,** 157–205.

[18] Leighton, F. T., Maggs, B. M. and Rao, S. B. (1994). Packet routing and job-shop scheduling in $O(congestion + dilation)$ steps. *Combinatorica* **14,** 167–186.

[19] Levin, D. A., Peres, Y. and Wilmer, E. L. (2009). *Markov chains and mixing times.* American Mathematical Soc.

[20] Little, J. D. C. (1961). Proof for the queuing formula: $L = \lambda w$. *Operations research* **9,** 383–387.

[21] Loynes, R. M. (1962). The stability of a queue with non-independent inter-arrival and service times. In *Mathematical Proc. of the Cambridge Philosophical Society.* vol. 58 Cambridge University Press. pp. 497–520.

[22] Lund, R. B. and Tweedie, R. L. (1996). Geometric convergence rates for stochastically ordered markov chains. *Mathematics of operations research* **21,** 182–194.

[23] Luo, W. and Ephremides, A. (1999). Stability of N interacting queues in random-access systems. *IEEE Trans. on Info. Theory* **45,** 1579–1587.

[24] Meyn, S. P. and Tweedie, R. L. (1993). *Markov chains and stochastic stability.* Springer Science & Business Media.

[25] Mosk-Aoyama, D. and Shah, D. (2006). Computing separable functions via gossip. In *Proc. of the 25th Annual ACM Symposium on Principles of Distributed Computing.* PODC '06. ACM. pp. 113–122.

[26] Palacios, J. L. (1994). Expected hitting and cover times of random walks on some special graphs. *Random Structures & Algorithms* **5,** 173–182.

[27] Sengupta, N., Bagchi, A., Ramanath, M. and Bedathur, S. (2019). ARROW: Approximating reachability using random-walks over web-scale graphs. In *Proc. of the 35th IEEE International Conference on Data Engineering.* ICDE '19. IEEE. pp. 470–481.

[28] Spielman, D. A. and Teng, S.-H. (2004). Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proc. of the 36th Annual ACM Symp. on Theory of Computing.* STOC '04. ACM, New York, NY, USA.

[29] Szpankowski, W. (1990). Towards computable stability criteria for some multidimensional stochastic processes. *Stochastic Analysis of Computer and Communication Systems* 131–172.

[30] Szpankowski, W. (1994). Stability conditions for some distributed systems: Buffered random access systems. *Advances in Applied Probability* **26,** 498–515.

[31] Tripathi, V., Talak, R. and Modiano, E. (2019). Age optimal information gathering and dissemination on graphs. In *Proc. of the 38th Annual Joint Conf. of the IEEE Computer and Comm. Societies.* INFOCOM '19 IEEE. pp. 2422–2430.

[32] Tsybakov, B. and Mikhailov, W. (1979). Ergodicity of slotted aloha system. *Probl. Pered. Inf.* **15,** 73–87.

[33] V. Kanade, F. M.-T. and Sauerwald, T. (2019). On coalescence time in graphs: When is coalescing as fast as meeting? In *Proc. of the 30th Annual ACM-SIAM Symp. on Discrete Algorithms.* SODA '19 SIAM. pp. 956–965.