Proofs and additional results

In the following appendices we provide proofs of Theorem 2 (Appendix EC.2) and Theorem 3 (Appendix EC.3), and also obtain tight results using our methods in the setting of Akbarpour et al. (2014) (Appendix EC.4).

EC.1. Preliminary results

We first state a number of propositions and lemmas that will enable our proof of 2.

We begin by stating (without proof) the following version of the classical Chernoff bound *(see, e.g., Alon and Spencer 2008).*

Proposition EC.1 **(Chernoff bound)**. *Let* $X_i \in \{0,1\}$ *be independent with* $\mathbb{P}(X_i = 1) = p_i$ *for* $1 \le i \le n$. *Let* $\mu = \sum_{i=1}^{n} p_i$.

(i) For any $\delta \in [0,1]$ *we have*

$$\mathbb{P}(|X-\mu| \ge \mu\delta) \le 2\exp\{-\delta^2\mu/3\}. \tag{EC.1}$$

(ii) For any $R > 6\mu$ *we have*

$$\mathbb{P}(X \ge R) \le 2^{-R}. \tag{EC.2}$$

Next, we state a result that is based on the Lyapunov function technique. Given an irreducible aperiodic Markov chain $\{X_k\}$ on a countable statespace $\mathcal{X}$, suppose there exists a nonnegative function $V : \mathcal{X} \to \mathbb{R}_+$ that admits the following decomposition:

$$V(X_{k+1}) = V(X_k) + A_k - D_k, \tag{EC.3}$$

where $A_k \ge 0$ is an i.i.d. sequence such that A_k is independent of state X_k, while $D_k \ge 0$ is a random function of X_k and A_k, that does not depend on k directly. A_k and D_k are interpreted as the number of arrivals and departures in the time period of length k, respectively. Assume in addition that $\mathcal{B}(\alpha) \triangleq \{x \in \mathcal{X} \mid V(x) \le \alpha\} \subset \mathcal{X}$ is finite for every α. Note that as $V(x) \ge 0$, we have that $D_k \le V(x) + A_k$ a.s.

Proposition EC.2. *Suppose $\mathbb{E}[A_k^2]$ is finite and C_1 satisfies $\mathbb{E}[A_k^2] \le C_1 \mathbb{E}[A_k]^2 < \infty$. Suppose there exist $\alpha, \lambda, C_2 > 0$ such that for every $x \notin \mathcal{B}(\alpha)$,*

$$\mathbb{E}\left[A_k - \tilde{D}_k | X_k = x\right] \le -\lambda \mathbb{E}[A_k], \tag{EC.4}$$

where $\tilde{D}_k$ is defined to be $\min\{D_k, C_2 A_k\}$. Then X_k is positive recurrent with the unique stationary distribution X_∞ and

$$\mathbb{E}[V(X_\infty)] \le \max\left\{\alpha, \frac{\max\{1, C_2 - 1\}^2 C_1}{\lambda} \mathbb{E}[A_k]\right\}\left(2 + \frac{2}{\lambda}\right).$$

The reason for introducing a truncated downward jump process $\tilde{D}_k$ as opposed to using just D_k is that in general the statement of the proposition is not true. Namely, there exists a process such that the assumptions of the proposition above hold true when D_k replaces $\tilde{D}_k$ in (EC.4) and $\mathbb{E}[V(X_\infty)] = \infty$, as shown in Example EC.1 at the end of this section.

Proposition EC.2 is obtained as a corollary of Proposition EC.3 and Proposition EC.4 below.

Suppose that $\{X_k\}$ is a discrete-time irreducible Markov chain on a countable state space $\mathcal{X}$. First we give a condition for the positive recurrence of $\{X_k\}$ due to Foster (1953); see Asmussen (2003) for a contemporary reference. Recall that $\mathbb{E}_x$ denotes the expectation operator conditional on $X_0 = x$.

Proposition EC.3 **(Foster 1953)**. *If there exists a function $V: \mathcal{X} \to \mathbb{R}$, $\gamma > 0$, and a finite set $B \subset \mathcal{X}$ such that for all $x \in B$,*

$$\mathbb{E}_x[V(X_1) - V(X_0)] < \infty, \tag{EC.5}$$

and for all $x \in \mathcal{X} \setminus B$,

$$\mathbb{E}_x[V(X_1) - V(X_0)] \le -\gamma,$$

then $\{X_k\}$ is positive recurrent.

Now, suppose that $\{X_k\}$ is ergodic (irreducible, with all states being positive recurrent and aperiodic), and let X_∞ denote the unique steady-state distribution. We now give a bound on the first moment of $f(X_\infty)$ for any function f. The result below is from Anderson (2014) but is similar to Gamarnik and Zeevi (2006) and Glynn et al. (2008).

PROPOSITION EC.4 **(Anderson 2014).** *Suppose that X_t is ergodic, and that there exist $\alpha, \beta, \gamma > 0$, a set $B \subset \mathcal{X}$, and functions $U: \mathcal{X} \to \mathbb{R}_+$ and $f: \mathcal{X} \to \mathbb{R}_+$ such that for $x \in \mathcal{X} \setminus B$,*

$$\mathbb{E}_x[U(X_1) - U(X_0)] \leq -\gamma f(x), \tag{EC.6}$$

and for $x \in B$,

$$f(x) \leq \alpha, \tag{EC.7}$$

$$\mathbb{E}_x[U(X_1) - U(X_0)] \leq \beta. \tag{EC.8}$$

Then

$$\mathbb{E}[f(X_\infty)] \leq \alpha + \frac{\beta}{\gamma}.$$

Note that we need not assume that $\mathcal{B}$ is bounded. Finally, we can prove the specialization of the above results as used in the paper.

Proof of Proposition EC.2. First, we apply Proposition EC.3 to X_k using the same $V(x)$, $\mathcal{B}$, and $\gamma = \lambda \mathbb{E}[A_k]$ as in the statement of Proposition EC.2. For $x \notin \mathcal{B}$, we have

$$\mathbb{E}_x[V(X_1) - V(X_0)] = \mathbb{E}_x[A_0 - D_0] \leq \mathbb{E}_x\left[A_0 - \tilde{D}_0\right] \leq -\lambda \mathbb{E}_x[A_0] = -\gamma,$$

where in the inequalities we use $\tilde{D}_k \leq D_k$ and then (EC.4). For all x, we have

$$\mathbb{E}_x[V(X_1) - V(X_0)] = \mathbb{E}_x[A_0 - D_0] \leq \mathbb{E}_x[A_0] < \infty.$$

Thus as $\mathcal{B}$ is bounded, we can apply Proposition EC.3 to obtain positive recurrence of X_k. By assumption, X_k is irreducible and aperiodic. Hence, X_k is ergodic. Let X_∞ be the steady-state version of the Markov chain X_k.

Next, we apply Proposition EC.4 by taking $U(x) = V^2(x)$ and $f(x) = V(x)$. We let

$$\alpha' = \max\left\{\alpha, \frac{\max\{1, C_2 - 1\}^2 C_1}{\lambda} \mathbb{E}[A_k]\right\}, \tag{EC.9}$$

thus making our set of exceptions from Proposition EC.4 $\mathcal{B}' = \{x \in \mathcal{X} \mid V(x) \le \alpha'\}$. We have for $x \in \mathcal{B}'$ that

$$\begin{aligned}
\mathbb{E}_x[U(X_1) - U(X_0)] &= \mathbb{E}_x\left[(V(X_0) + A_0 - D_0)^2 - V(X_0)^2\right] && \text{(EC.10)}\\
&\le \mathbb{E}_x\left[(V(X_0) + A_0)^2 - V(X_0)^2\right] && \text{(EC.11)}\\
&= 2V(x)\mathbb{E}[A_0] + \mathbb{E}[A_0^2] \\
&\le 2\alpha'\mathbb{E}[A_0] + C_1\mathbb{E}[A_0]^2 && \text{(EC.12)}\\
&\le 2\alpha'\mathbb{E}[A_0] + \alpha'\lambda\mathbb{E}[A_0] && \text{(EC.13)}\\
&= \alpha'(2+\lambda)\mathbb{E}[A_0] \\
&\triangleq \beta' && \text{(EC.14)}
\end{aligned}$$

where (EC.10) follows from (EC.3), (EC.11) follows as $V(X_1) \ge 0$, (EC.12) follows from (EC.9) and the definition of C_1, and (EC.13) follows again from (EC.9) and as $\max\{1, C_2 - 1\}^2 \ge 1$ by definition. For $x \notin \mathcal{B}'$, we have that

$$\begin{aligned}
\mathbb{E}_x[U(X_1) - U(X_0)] &= \mathbb{E}_x\left[(V(X_0) + A_0 - D_0)^2 - V(X_0)^2\right] && \text{(EC.15)}\\
&\le \mathbb{E}_x\left[(V(X_0) + A_0 - \tilde{D}_0)^2 - V(X_0)^2\right] && \text{(EC.16)}\\
&= 2V(x)\mathbb{E}_x[A_0 - \tilde{D}_0] + \mathbb{E}_x\left[(A_0 - \tilde{D}_0)^2\right] \\
&\le -2V(x)\lambda\mathbb{E}[A_0] + \mathbb{E}\left[\max\{1, C_2 - 1\}^2 A_0^2\right] && \text{(EC.17)}\\
&\le -(V(x) + \alpha')\lambda\mathbb{E}_x[A_0] + \max\{1, C_2 - 1\}^2 C_1 \mathbb{E}\left[A_0\right]^2 && \text{(EC.18)}\\
&\le -(V(x) + \alpha')\lambda\mathbb{E}_x[A_0] + \alpha'\lambda\mathbb{E}\left[A_0\right] && \text{(EC.19)}\\
&= -V(x)\lambda\mathbb{E}_x[A_0],
\end{aligned}$$

where (EC.15) follows from (EC.3), (EC.16) follows as $V(X_1) \ge 0$ and $\tilde{D}_0 \le D_0$, (EC.17) follows from (EC.4) and as $\tilde{D}_k \le C_2 A_k$ a.s. implies that $|A_k - D_k| \le \max\{1, C_2 - 1\}A_k$ a.s., (EC.18) follows from (EC.9) and the definition of C_1, and finally (EC.19) follows again from (EC.9). Thus by taking $\gamma' = \lambda\mathbb{E}_x[A_0]$, we can now apply Proposition EC.4 with α', β', and γ' to obtain that

$$\mathbb{E}[V(\infty)] \le \alpha' + \frac{\beta'}{\gamma'} = \alpha' + \frac{\alpha'(2+\lambda)\mathbb{E}[A_0]}{\lambda\mathbb{E}[A_0]} = \alpha'\left(1 + \frac{2+\lambda}{\lambda}\right)$$

$$= \max\left\{\alpha, \frac{\max\{1, C_2-1\}^2 C_2}{\lambda}\mathbb{E}[A_k]\right\}\left(1+\frac{2+\lambda}{\lambda}\right)$$

showing the result. □

Last, we give a quick counterexample showing that without some assumptions beyond simply having a negative drift, we may not even have a finite first moment.

Example EC.1. Consider the following random walk X_t on the nonnegative integers parametrized by some $\gamma \in (0,1)$. From state 0, we always go up to state 1. For every other state $k = 1, 2, \ldots$, with probability $(1+\gamma)/(k+1)$, we go to state 0, and with the remaining probability, $(k-\gamma)/(k+1)$, we go up to state $k+1$. This walk has the property that for all $k \geq 1$,

$$\mathbb{E}_k[X_1 - X_0] = (k+1)\cdot\frac{k-\gamma}{k+1} + 0\cdot\frac{1+\gamma}{k+1} - k = -\gamma,$$

and thus is positive recurrent and has some stationary distribution $\pi_k = \mathbb{P}(X_\infty = k)$. However, we will show that $\mathbb{E}[X_\infty] = \infty$. A direct computation of the steady-state equations gives that $\pi_0 = \pi_1$, and for $n \geq 2$,

$$\pi_n = \frac{n-\gamma}{n+1}\pi_{n-1} = \pi_0 \prod_{k=2}^{n}\frac{k-\gamma}{k+1} = \frac{\pi_0}{\Gamma(2-\gamma)}\frac{\Gamma(n+1-\gamma)}{\Gamma(n+2)},$$

where Γ is the gamma function. Using the identity

$$\lim_{n\to\infty}\frac{\Gamma(n+\alpha)}{\Gamma(n)n^{\alpha}} = 1$$

for all $\alpha \in \mathbb{R}$, we have that

$$\pi_n = \Theta\left(\frac{1}{n^{1+\gamma}}\right).$$

Thus there exists $c > 0$ and $\ell \in \mathbb{Z}_+$ such that

$$\mathbb{E}[X_\infty] = \sum_{n=0}^{\infty} n\pi_n \geq \sum_{n=\ell}^{\infty}\frac{c}{n^{\gamma}} = \infty,$$

showing the claim. □

EC.2. Proof of Theorem 2

We prove the lower bound in Theorem 2 for the class of *monotone policies* that includes batching policies as a special case.

Let $\mathcal{G}$ denote the *global compatibility graph* that includes all nodes that ever arrive to the system, and directed edges representing compatibilities between them.

Definition EC.1 (**Monotone policy**). A deterministic policy (under either chain removal or cycle removal) is said to be *monotone* if it satisfies the following property: consider any pair of nodes (i,j) and an arbitrary global compatibility graph $\mathcal{G}$ such that the edge (i,j) is present. Let $\bar{\mathcal{G}}$ be the graph obtained from $\mathcal{G}$ when edge (i,j) is removed. Let T_i and T_j be the times of the removal of nodes i and j respectively when the compatibility graph is $\mathcal{G}$ and let $T_{ij} = \min(T_i, T_j)$. Then the policy must act in an identical fashion on $\bar{\mathcal{G}}$ and $\mathcal{G}$ for all $t < T_{ij}$; i.e., the same cycles/chains are removed at the same times in each case, up to time T_{ij}. This property must hold for every pair of nodes (i,j) and every possible $\mathcal{G}$ containing the edge (i,j).

A randomized policy is said to be monotone if it randomizes between deterministic monotone policies.

Remark EC.1. Consider the greedy policy for cycle removal defined above. It is easy to see that we can suitably couple the execution of the greedy policy on different global compatibility graphs such that the resulting policy is monotone. The same applies to a batching policy that matches periodically (after arrival of x nodes), by finding a maximum packing of node disjoint cycles and removing them. (Note that such a policy is periodic Markov with a period equal to the batch size.)

Note that the class of monotone policies includes a variety of policies in addition to simple batching policies. For instance, a policy that assigns weights to nodes and finds an allocation with maximum weight (instead of simply maximizing the number of nodes matched) is also monotone.

Before proving Theorem 2, we need a couple of results. Next we state a straightforward combinatorial bound: in a directed graph, a set $\mathcal{M}$ of node disjoint three-way cycles is said to be *maximal* if no three-way cycle can be added to $\mathcal{M}$ such that the set remains node disjoint.

PROPOSITION EC.5. *Given an arbitrary directed graph G, let N be the number of three-way cycles in a maximum (i.e., largest in cardinality) set of node disjoint three-way cycles in G. Then, any maximal set of node disjoint three-way cycles consists of at least $N/3$ three-way cycles.*

Proof. We prove the result by contradiction. Assume that a maximal set of node disjoint three-way cycles $\mathcal{W}$ contains fewer than $N/3$ three-way cycles. Then there must be a three-way cycle X from a largest set of node disjoint three-way cycles such that for every three-way cycle $Y \in \mathcal{W}$, X and Y have no nodes in common. This yields a contradiction, as we could then add X to $\mathcal{W}$ to make a larger set of node disjoint three-way cycles, thus making $\mathcal{W}$ not maximal. □

Let $\mathcal{G}_t$ denote the *global compatibility graph* that includes all nodes that ever arrive to the system up to time t, and all directed edges representing compatibilities between them. Denote by $\mathcal{W}_t$ the set of nodes out of $0, 1, \dots, t$ still present in the system at time t. The following is a key property of monotone policies:

LEMMA EC.1. *Under any monotone policy, for every two nodes i, j arriving before time t (namely, $i, j \le t$) and every subset of nodes $\mathcal{W} \subset \{0, 1, \dots, t\}$ containing nodes i and j,*

$$\mathbb{P}((i,j) \in \mathcal{G}_t | \mathcal{W}_t = \mathcal{W}) \le p.$$

In words, pairs of nodes still present in the system at time t are no more likely to be connected at time t than at the time they arrived.

Proof. We assume that the removal policy is deterministic. The proof for the case of randomized policies follows immediately. Fix any two nodes i, j that arrive before time t (namely, $i, j \le t$). Given any directed graph $\mathcal{G}$ on nodes $0, 1, \dots, t$ (that is, nodes arriving up to time t) such that the edge (i, j) belongs to $\mathcal{G}$, denote by $\bar{\mathcal{G}}$ the same graph $\mathcal{G}$ with edge (i, j) deleted. Let $\mathcal{W}$ be any subset of nodes $0, 1, \dots, t$ containing i and j. Recall that we denote by $\mathcal{G}_t$ the directed graph generated by nodes $0, 1, \dots, t$ and by $\mathcal{W}_t$ the set of nodes observed at time t. Note that, since the policy is deterministic, graph $\mathcal{G}_t$ uniquely determines the set of nodes $\mathcal{W}_t$.

We have

$$\mathbb{P}(\mathcal{W}_t = \mathcal{W}) = \sum_{\mathcal{G}} \mathbb{P}(\mathcal{G}) + \sum_{\mathcal{G}} \mathbb{P}(\bar{\mathcal{G}}),$$

where the first sum is over graphs $\mathcal{G}$ containing edge (i,j) such that the set of nodes observed at time t is $\mathcal{W}$ when $\mathcal{G}_t = \mathcal{G}$, and the second sum is over graphs $\mathcal{G}$ containing edge (i,j), such that when $\mathcal{G}_t = \bar{\mathcal{G}}$, the set of nodes observed at time t is $\mathcal{W}$. Note, however, that by our monotonicity assumption, if $\mathcal{G}_t = \mathcal{G}$ implies $\mathcal{W}_t = \mathcal{W}$, then $\mathcal{G}_t = \bar{\mathcal{G}}$ also implies that $\mathcal{W}_t = \mathcal{W}$. Thus

$$\mathbb{P}(\mathcal{W}_t = \mathcal{W}) \geq \sum_{\mathcal{G}} (\mathbb{P}(\mathcal{G}) + \mathbb{P}(\bar{\mathcal{G}})),$$

where the sum is over graphs $\mathcal{G}$ containing edge (i,j) such that $\mathcal{G}_t = \mathcal{G}$ implies that $\mathcal{W}_t = \mathcal{W}$. At the same time note that $\mathbb{P}(\bar{\mathcal{G}}) = \mathbb{P}(\mathcal{G})(1-p)/p$ since it corresponds to the same graph except that edge (i,j) is deleted. We obtain

$$\mathbb{P}(\mathcal{W}_t = \mathcal{W}) \geq \sum_{\mathcal{G}} \mathbb{P}(\mathcal{G})(1 + (1-p)/p) = \sum_{\mathcal{G}} \mathbb{P}(\mathcal{G})/p.$$

We recognize the right-hand side as $\mathbb{P}(\mathcal{W}_t = \mathcal{W} | (i,j) \in \mathcal{G}_t)$. Now we obtain

$$\begin{aligned}\mathbb{P}((i,j) \in \mathcal{G}_t | \mathcal{W}_t = \mathcal{W}) &= \mathbb{P}(\mathcal{W}_t = \mathcal{W} | (i,j) \in \mathcal{G}_t)\mathbb{P}((i,j) \in \mathcal{G}_t)/\mathbb{P}(\mathcal{W}_t = \mathcal{W}) \\ &\leq \mathbb{P}((i,j) \in \mathcal{G}_t) \\ &\leq p,\end{aligned}$$

and the claim is established. □

The following corollary follows immediately by linearity of expectations.

Corollary EC.1. *Let $W_t = |\mathcal{W}_t|$ and let E_t be the number of edges between nodes in $\mathcal{W}_t$. Then, under a monotone policy, $\mathbb{E}[E_t | W_t] \leq W_t(W_t - 1)p$.*

Proof of Theorem 2: Upper bound for the greedy policy

Proof of Theorem 2: The performance of the greedy policy. Suppose that at time zero we observe $W \geq C^3/p^{3/2}$ nodes in the system with an arbitrary set of edges between them. Here C

is a sufficiently large constant to be fixed later. Call this set of nodes $\mathcal{W}$. Consider the next $T = 1/(Cp^{3/2})$ arrivals, and call this set of nodes $\mathcal{A}$. Without loss of generality label the times of these arrivals as $1, 2, \ldots, T$, and use the label t for the node that arrives at time t. Let $\mathcal{A}_t \subseteq \{1, 2, \ldots, t-1\}$ be the subset of nodes in $\mathcal{A}$ that have arrived but have not been removed before time t. Similarly, define $\mathcal{W}_t$ to be the set of nodes from $\mathcal{W}$ that are still in the system immediately before time t. Note that, in particular, $\mathcal{W}_1 = \mathcal{W}$.

Let N be the number of three-way cycles removed during the time period $[0, T]$, that include two nodes from $\mathcal{W}$. These are "good" three-way cycles (helping us obtain a negative drift). Let $\kappa = 1/C^2$ and consider the event

$$\mathcal{E}_1 \equiv \{|\mathcal{A}_{T+1}| - N \geq 2\kappa/p^{3/2}\}. \tag{EC.20}$$

Thus, $\mathcal{E}_1$ is a "bad" event corresponding to many of the new arrivals still being in the system at the end of T periods, leading to large $|\mathcal{A}_{T+1}|$. We discount $|\mathcal{A}_{T+1}|$ by the number of good three-way cycles N.

Introduce the event

$$\begin{aligned}\mathcal{E}_2 \equiv \{&\text{There exists a set of disjoint two- and three-way cycles in } \mathcal{A} \\ &\text{with cardinality at least } 3/(C^3p^{3/2})\ \}.\end{aligned} \tag{EC.21}$$

Event $\mathcal{E}_2$ is a "bad" event under which it is possible that many nodes in $\mathcal{A}$ depart due to cycles containing only other nodes from $\mathcal{A}$, and hence do not help with clearing any of the nodes in $\mathcal{W}$.

We now show that if neither of these bad events occurs, then the number of nodes in the system decreases by at least $3/(8Cp^{3/2})$ in expectation. The rest of the proof will then focus on showing that each of these events occurs with a small probability.

First suppose that the event $\mathcal{E}_1$ does not occur. Then

$$|\mathcal{A}_{T+1}| \leq \frac{2}{C^2p^{3/2}} + N \leq \frac{1}{16Cp^{3/2}} + N\,, \tag{EC.22}$$

for C sufficiently large. Also, event $\mathcal{E}_2^c$ implies that (again for C sufficiently large) at most $9/(C^3p^{3/2}) \le 1/(16Cp^{3/2})$ nodes in $\mathcal{A}$ leave due to internal three-way cycles or two-way cycles. Since $T = 1/(Cp^{\frac{3}{2}})$, it follows that by Eq. (EC.22), at least $7/(8Cp^{3/2}) - N$ other nodes in $\mathcal{A}$ also leave before $T+1$. These other nodes belong to cycles of one of the following types:

(i) A three-way cycle containing another node from $\mathcal{A}$ and a node from $\mathcal{W}$.

(ii) A two-way cycle with a node from $\mathcal{W}$.

(iii) A three-way cycle containing two nodes from $\mathcal{W}$. There are exactly N nodes of this type.

Exactly N nodes in $\mathcal{A}$ are removed due to cycles of type (iii) above, so from which we infer that at least $7/(8Cp^{3/2}) - 2N$ nodes in $\mathcal{A}$ are removed due to cycles of type (i) or (ii) above, meaning that at least $(1/2)(7/(8Cp^{3/2}) - 2N)$ nodes in $\mathcal{W}$ are removed as part of such cycles. Clearly, $2N$ nodes in $\mathcal{W}$ are removed as part of cycles of type (iii).

It follows that

$$|\mathcal{W}_{T+1}| \le |\mathcal{W}| - 2N - \frac{1}{2}\left(\frac{7}{8Cp^{3/2}} - 2N\right) \le |\mathcal{W}| - \frac{7}{16Cp^{3/2}} - N\,. \tag{EC.23}$$

Combining Eqs. (EC.22) and (EC.23), we deduce that

$$|\mathcal{W}_{T+1}| + |\mathcal{A}_{T+1}| \le |\mathcal{W}| - \frac{3}{8Cp^{3/2}}\,. \tag{EC.24}$$

We also have that the number of nodes in the system increases by at most $T = 1/(Cp^{3/2})$. We will show that $\mathbb{P}(\mathcal{E}_1 \cup \mathcal{E}_2) \le \varepsilon = 1/9$. Before establishing this claim we show how this claim implies the result. We have

$$\mathbb{E}[|\mathcal{W}_{T+1}| + |\mathcal{A}_{T+1}| - |\mathcal{W}|] \le \varepsilon T - (1-\varepsilon)\frac{3}{8Cp^{3/2}} = -\frac{2}{9Cp^{3/2}}; \tag{EC.25}$$

i.e., the number of nodes decreases by at least $2/(9Cp^{3/2})$ in expectation. We now apply Proposition EC.2 to the embedded Markov chain observed at times that are multiples of T. Namely, let $T_i = i \cdot T$, and take $X_i = \mathcal{G}(T_i) = (\mathcal{V}(T_i), \mathcal{E}(T_i))$ and define $V(X_i) = |\mathcal{V}(T_i)|$. If we let $\mathcal{D}_i$ be the set of nodes that are deleted in some cycle during the time interval $[T_i, T_{i+1})$, we obtain a decomposition

$$V(X_{i+1}) = |\mathcal{V}(T_{i+1})| = |\mathcal{V}(T_i)| + T - |\mathcal{D}_i| = V(X_i) + T - |\mathcal{D}_i|. \tag{EC.26}$$

Since $T > 0$ is deterministic it is trivially independent of $\mathcal{G}(T_i)$. Thus the assumptions on decomposing V from (EC.3) are satisfied. The assumption that $\{\mathcal{G} \mid V(\mathcal{G}) < n\}$ is finite for every n is satisfied as there are only finitely many graphs with n nodes. We take $\alpha = C^3/p^{3/2}$, leading to $\mathcal{B} = \{\mathcal{G} \mid |\mathcal{V}(\mathcal{G})| \leq C^3/p^{3/2}\}$. We can take $C_1 = 1$ as T is deterministic. We can take $C_2 = 3$, as trivially $|\mathcal{D}_i| \leq 3T$ since each newly arriving node can be in at most one three-way cycle (and hence $\tilde{D}_k = D_k$ in Proposition EC.2). Finally, we can take $\lambda = 2/9$, as by (EC.25),

$$\mathbb{E}[T - |\mathcal{D}_i|] \leq -\frac{2}{9Cp^{3/2}} = -\frac{2}{9}\mathbb{E}[T].$$

Thus, by applying Proposition EC.2, we obtain that

$$\begin{aligned}
\mathbb{E}[|\mathcal{V}(T_\infty)|] &\leq \max\left\{\alpha, \frac{\max\{1, C_2 - 1\}^2 C_1}{\lambda}\mathbb{E}[A_k]\right\}\left(2 + \frac{2}{\lambda}\right) \\
&= \max\left\{\frac{C^3}{p^{3/2}}, \frac{4}{2/9}\frac{1}{Cp^{3/2}}\right\}\left(2 + \frac{2}{2/9}\right) \\
&= \frac{11C^3}{p^{3/2}},
\end{aligned}$$

for C sufficiently large. Finally, since the embedded chain is observed over deterministic time intervals, the bound above applies to the steady-state bound. We conclude that

$$\mathbb{E}[|\mathcal{V}(\infty)|] \leq \frac{11C^3}{p^{3/2}}.$$

It remains to bound $\mathbb{P}(\mathcal{E}_1)$ and $\mathbb{P}(\mathcal{E}_2)$ to complete the proof. We do this below. We claim that $\mathbb{P}(\mathcal{E}_2) \leq \varepsilon/4$. We first show that there is likely to be a maximal set of node disjoint three-way cycles in $\mathcal{A}$ of size less than $2/(3C^3p^{3/2})$. This will imply, using Proposition EC.5, that the maximum number of node disjoint three-way cycles in $\mathcal{A}$ is at most $2/(C^3p^{3/2})$. Reveal the graph on $\mathcal{A}$ and simultaneously construct a maximal set of node disjoint three-way cycles as follows. Reveal node 1. Then reveal node 2. Then reveal node 3 and whether it forms a three-way cycle with the existing nodes. If it does remove this three-way cycle. Continuing this way, at any stage t if a three-way cycle is formed, choose uniformly at random such a three-way cycle and remove it.

Since this process corresponds to a monotone policy (see Definition EC.1), it follows from Corollary EC.1 that the residual graph immediately before step t contains no more than $2\binom{t-1}{2}p$ edges in expectation, as the number of nodes is no more than $t-1$. It follows that the conditional probability of three-way cycle formation at step t is no more than $\mathbb{E}[\text{Number of three-way cycles formed}] = 2\binom{t-1}{2}p^3$. It follows that we can set up a coupling such that the total number of three-way cycles removed (this is a maximal set of edge disjoint three-way cycles resulting from our particular greedy policy) is no more than $Z = \sum_{t=1}^{T} X_t$, where $X_t \sim \text{Bernoulli}(2\binom{t-1}{2}p^3)$ are independent. Now $\mathbb{E}[Z] = 2\binom{T}{3}p^3 \leq 1/(3C^3p^{3/2})$. Using Proposition EC.1 (i), we obtain that $\mathbb{P}(Z \geq 2/(3C^3p^{3/2})) < \varepsilon/8$, for large enough p, establishing the desired bound on the number of node disjoint three-way cycles. We have shown that the probability of having more than $2/(C^3p^{3/2})$ node disjoint three-way cycles in $\mathcal{A}$ is less than $\varepsilon/8$.

Let Z' be the number of two-way cycles internal to $\mathcal{A}$. Then $Z' \sim \text{Bin}(\binom{T}{2}, p^2)$. Hence, $\mathbb{E}[Z'] \leq 1/(C^2p)$ and $\mathbb{P}(Z' \geq 1/(C^3p^{3/2})) \leq \varepsilon/8$ for sufficiently small p using Proposition EC.1 (ii). It follows that the probability of having more than $1/(C^3p^{3/2})$ node disjoint two-way cycles in $\mathcal{A}$ is less than $\varepsilon/8$. Now $\mathbb{P}(\mathcal{E}_2) \leq \varepsilon/4$ follows by union bound.

We now show $\mathbb{P}(\mathcal{E}_1) \leq 3\varepsilon/4$. To prove this, we find it convenient to define two additional events. Denote by $\mathcal{N}(\mathcal{S}_1, \mathcal{S}_2)$ the (directed) neighborhood of the nodes in $\mathcal{S}_1$ in the set of nodes $\mathcal{S}_2$, i.e., $\mathcal{N}(\mathcal{S}_1, \mathcal{S}_2) = \{j \in \mathcal{S}_2 : \exists i \in \mathcal{S}_1 \text{ s.t. } (i,j) \in \mathcal{E}\}$. Abusing notation, we use $\mathcal{N}(i, \mathcal{S})$ to denote the neighborhood of node i in $\mathcal{S}$. Further, we find it convenient to define $\mathcal{B}_t = \mathcal{N}(t, \mathcal{A}_t)$. Define

$$\mathcal{E}_{3,t} \equiv \{|\mathcal{A}_t| \geq \kappa/p^{3/2}, \text{ and } |\mathcal{B}_t| < \kappa/(2p^{1/2})\}, \tag{EC.27}$$

and $\mathcal{E}_3 = \cup_{0 \leq t \leq T} \mathcal{E}_{3,t}$. Thus, $\mathcal{E}_3$ is the "bad" event that at some time t, the set $\mathcal{A}_t$ of arrivals that has not yet departed is "large" and yet the arrival at time t forms much fewer edges with $\mathcal{A}_t$ than expected. We will show that $\mathcal{E}_3$ is unlikely. Define

$$\mathcal{E}_{4,t} \equiv \{|\mathcal{A}_t| \geq \kappa/p^{3/2}, \text{ and } |\mathcal{N}(\mathcal{B}_t, \mathcal{W}_t)| < C^3\kappa/(8p)\}, \tag{EC.28}$$

and let $\mathcal{E}_4 = \cup_{0 \leq t \leq T} \mathcal{E}_{4,t}$. Thus, $\mathcal{E}_4$ is the "bad" event that at some time t, the set $\mathcal{A}_t$ of arrivals that has not yet departed is "large" and yet the arrival at time t reaches fewer than the expected number of nodes in $\mathcal{W}_t$ in two-hops via a node in $\mathcal{A}_t$. We will show that $\mathcal{E}_4$ is unlikely if $\mathcal{E}_3$ does not occur. We will also show that if $\mathcal{E}_4$ does not occur, this ensures that the event $\mathcal{E}_1$ (which we are trying to control) is unlikely to occur, since if $|\mathcal{A}_t| \geq \kappa/p^{3/2}$, then the next new arrival is likely to be removed in a three-way cycle (We will show that a three-way cycle of type (i) is likely to be formed. If such a three-way cycle is executed then $|\mathcal{A}_{t+1}| = |\mathcal{A}_t| - 1$; otherwise if a three-way cycle of type (iii) is executed then that leads to $N_{t+1} = N_t + 1$, where N_t is the number of three-way cycles of type (iii) executed before time t. In either case, $|\mathcal{A}_{t+1}| - N_{t+1} = (|\mathcal{A}_t| - N_t - 1)$. In words, this implies that $(|\mathcal{A}_t| - N_t$ is pushed downwards in expectation whenever it exceeds $\kappa/p^{3/2}$.) We make use of

$$\begin{aligned} \mathcal{E}_1 \; &\subseteq \; (\mathcal{E}_4^c \cap \mathcal{E}_1) \cup \mathcal{E}_4 \; \subseteq \; (\mathcal{E}_4^c \cap \mathcal{E}_1) \cup \mathcal{E}_3 \cup (\mathcal{E}_4 \cap \mathcal{E}_3^c) \\ \Rightarrow \; \mathbb{P}(\mathcal{E}_1) \; &\leq \mathbb{P}(\mathcal{E}_4^c \cap \mathcal{E}_1) + \mathbb{P}(\mathcal{E}_3) + \mathbb{P}(\mathcal{E}_4 \cap \mathcal{E}_3^c). \end{aligned}$$

Reveal the edges between t and $\mathcal{A}_t$ when node t arrives. The existence of each edge is independent of the other edges and the current revealed graph. Thus we can bound the probability of the event $\mathcal{E}_{3,t}$ using Proposition EC.1 (i) by $2\exp(-1/(12C^2p^{1/2}))$ for large enough C. It follows that for sufficiently small p, we have

$$\mathbb{P}(\mathcal{E}_3) \leq 2T\exp(-1/(12C^2p^{1/2})) \leq \varepsilon/4\,. \tag{EC.29}$$

We now bound $\mathbb{P}(\mathcal{E}_4^c \cap \mathcal{E}_1)$. Let N_t be the number of three-way cycles removed before time t of type (iii) (recall that three-way cycles of type (iii) include two nodes from $\mathcal{W}$). Define $Z_t \equiv |\mathcal{A}_t| - N_t$. Define

$$\mathcal{E}_{5,t} \equiv \{\text{Node } t \text{ is immediately part of a three-way cycle of type (i)}\}\,.$$

Assume that three-way cycles are given priority over two-way cycles. Note that:

- If $\mathcal{E}_{5,t}$ then $|\mathcal{A}_{t+1}| = |\mathcal{A}_t| - 1, N_{t+1} = N_t$ if such a three-way cycle is removed and $|\mathcal{A}_{t+1}| = |\mathcal{A}_t|, N_{t+1} = N_t + 1$ if a three-way cycle of type (iii) is removed instead. In either case, we have $Z_{t+1} = Z_t - 1$. If a three-way cycle consisting of nodes from $\mathcal{A}$ is removed then we have $|\mathcal{A}_{t+1}| = |\mathcal{A}_t| - 2, N_{t+1} = N_t + 1$, i.e., $Z_{t+1} = Z_t - 2$. Overall, $Z_{t+1} \leq Z_t - 1$ in any of these cases.

- With probability one we have $|\mathcal{A}_{t+1}| \le |\mathcal{A}_t| + 1$ and $N_{t+1} \ge N_t$. It follows that $Z_{t+1} \le Z_t + 1$.

Now suppose that $Z_t \ge \kappa/p^{3/2}$ and $\mathcal{E}_{4,t}^c$. Clearly $Z_t \ge \kappa/p^{3/2} \Rightarrow |\mathcal{A}_t| \ge \kappa/p^{3/2}$ and hence $\mathcal{E}_{4,t}^c \Rightarrow |\mathcal{N}(\mathcal{B}_t, \mathcal{W}_t)| \ge C^3\kappa/(8p) = C/(8p)$. Revealing the edges between from $\mathcal{N}(\mathcal{B}_t, \mathcal{W}_t)$ to t, we see that

$$\mathbb{P}(\mathcal{E}_{5,t} | Z_t \ge \kappa/p^{3/2}, \mathcal{E}_{4,t}^c) \ge 1 - (1-p)^{C/(4p)} \ge 3/4\,, \tag{EC.30}$$

for large enough C and small enough p, independent of everything so far. Hence, informally, if $\mathcal{E}_{4,t}^c$ then Z_t is bounded above by a random walk with a downward drift whenever $Z_t \ge \kappa/p^{3/2}$. We now formalize this.

Define the random walk $(\widetilde{Z}_t)_{t\ge 1}$ as follows. Let $\widetilde{Z}_1 = 0$. Whenever $\widetilde{Z}_t = 0$, we have $\widetilde{Z}_{t+1} = 1$; otherwise

$$\widetilde{Z}_{t+1} = \begin{cases} \widetilde{Z}_t + 1 \text{ with probability } 1/4 \\ \widetilde{Z}_t - 1 \text{ with probability } 3/4 \end{cases} \tag{EC.31}$$

Hence $(\widetilde{Z}_t)_{t=1}^{T+1}$ is a downward biased random walk reflected upwards at 0.

Proposition EC.6. *There exists $C < \infty$ such that for any $T \in \mathbb{N}$ and $\nu > 0$, we have $\mathbb{P}(\widetilde{Z}_{T+1} \ge \nu) \le CT\exp(-\nu/C)$.*

The proof is omitted, as this is a standard result for random walks with a negative drift. Using Proposition EC.6, we have that for sufficiently small p,

$$\mathbb{P}(\widetilde{Z}_{T+1} \ge \kappa/(2p^{3/2})) \le \varepsilon/4\,.$$

Let τ be the first time at which event $\mathcal{E}_{4,t}$ occurs for $t \le T$, and let $\tau = T+1$ if $\mathcal{E}_4$ does not occur. We now show that the following claim holds:

Claim EC.1. *We can couple Z_t and $\widetilde{Z}_t$ such that for all $t < \tau$, whenever $Z_t \ge \kappa/p^{3/2}$ we have $\widetilde{Z}_{t+1} - \widetilde{Z}_t \ge Z_{t+1} - Z_t$.*

Proof of Claim. If $\mathcal{E}_{5,t}$ occurs, then (see above) we know that $Z_{t+1} = Z_t - 1$ and $\widetilde{Z}_{t+1} - \widetilde{Z}_t \ge -1$ holds by definition of $\widetilde{Z}$. Hence, it is sufficient to ensure that $\widetilde{Z}_{t+1} = \widetilde{Z}_t + 1$ whenever $\mathcal{E}_{5,t}^c$ occurs. But this is easy to satisfy since Eq. (EC.30) implies that

$$\mathbb{P}(\mathcal{E}_{5,t}^c | Z_t \ge \kappa/p^{3/2}, \mathcal{E}_{4,t}^c) \le 1/4\,,$$

whereas $\mathbb{P}(\widetilde{Z}_{t+1} = \widetilde{Z}_t + 1) = 1/4$. This completes the proof of the claim. □

The following claim is an immediate consequence.

Claim EC.2. *We have* $Z_t \le \widetilde{Z}_t + \lceil \kappa/p^{3/2} \rceil$ *for all* $t \le \tau$.

Proof of Claim. The claim follows from Claim EC.1 and a simple induction argument. □

It follows that

$$\begin{aligned}\mathbb{P}(Z_{T+1} \ge 2\kappa/p^{3/2}, \tau = T+1) &\le \mathbb{P}(\widetilde{Z}_{T+1} \ge \kappa/p^{3/2}, \tau = T+1)\\ &\le \mathbb{P}(\widetilde{Z}_{T+1} \ge \kappa/p^{3/2})\\ &\le \varepsilon/4\,.\end{aligned}$$

Thus we obtain

$$\mathbb{P}(\mathcal{E}_4^c \cap \mathcal{E}_1) \le \varepsilon/4\,. \tag{EC.32}$$

Finally, we bound $\mathbb{P}(\mathcal{E}_4 \cap \mathcal{E}_3^c)$. For any $\mathcal{S} \subseteq \mathcal{A}$, let $\mathcal{W}_{\sim\mathcal{S}} \subseteq \mathcal{W}$ be the set of waiting nodes that would have been removed before $T+1$ if (hypothetically) the nodes in $\mathcal{S}$ had had no incident edges in either direction, but we had left all other compatibilities unchanged. Define the event $\mathcal{E}_6$ as follows: for all $\mathcal{S} \subseteq \mathcal{A}$ such that $|\mathcal{S}| = \kappa/(2p^{1/2})$, the bound

$$|\mathcal{N}(\mathcal{S}, \mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}})| \ge C/(8p) \tag{EC.33}$$

holds.

Claim EC.3. *The event* $\mathcal{E}_6$ *occurs with high probability.*

Before proving the claim, we show that it implies that $\mathbb{P}(\mathcal{E}_4 \cap \mathcal{E}_3^c) \le \varepsilon/4$. Suppose that $\mathcal{E}_6$ and $\mathcal{E}_3^c$ occur. Consider any t such that $|\mathcal{A}_t| \ge \kappa/p^{3/2}$. Since $\mathcal{E}_3^c$, we have that $|\mathcal{B}_t| \ge \kappa/(2p^{1/2})$. Take any $\mathcal{S} \subseteq \mathcal{B}_t$ such that $|\mathcal{S}| = \kappa/(2p^{1/2})$. Notice that for our monotone greedy policy (see Remark EC.1) the set of waiting nodes that are removed before time t must be a subset of $\mathcal{W}_{\sim\mathcal{S}}$; i.e., we have we have that $\mathcal{W}_t \supseteq \mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}}$. Since $\mathcal{E}_6$ occurs, it follows that $|\mathcal{N}(\mathcal{S}, \mathcal{W}_t)| \ge C/(8p) \Rightarrow |\mathcal{N}(\mathcal{B}_t, \mathcal{W}_t)| \ge C/(8p)$. Thus we have $\mathcal{E}_4^c$. This argument establishes that

$$\begin{aligned}&\mathcal{E}_6 \cap \mathcal{E}_3^c \subseteq \mathcal{E}_4^c \cap \mathcal{E}_3^c\\ \Rightarrow\ \ &\mathcal{E}_6^c \cap \mathcal{E}_3^c \supseteq \mathcal{E}_4 \cap \mathcal{E}_3^c\,.\end{aligned}$$

It follows that $\mathbb{P}(\mathcal{E}_4 \cap \mathcal{E}_3^c) \le \mathbb{P}(\mathcal{E}_6^c \cap \mathcal{E}_3^c) \le \mathbb{P}(\mathcal{E}_6^c) \le \varepsilon/4$ using Claim EC.3, as required.

Proof of Claim EC.3. Consider any $\mathcal{S} \subseteq \mathcal{A}$ such that $|\mathcal{S}| = \kappa/(2p^{1/2})$. Clearly, since each node in $\mathcal{A}$ can eliminate at most 2 nodes in $\mathcal{W}$, we have that $|\mathcal{W}_{\sim\mathcal{S}}| \leq 2|\mathcal{A}\backslash\mathcal{S}| \leq 2|\mathcal{A}| = 2/(Cp^{3/2})$. It follows that $|\mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}}| \geq C^3/p^{3/2} - 2/(Cp^{3/2}) \geq C^3/(2p^{3/2})$ for large enough C. Now notice that by definition $\mathcal{W}_{\sim\mathcal{S}}$ is a function of only the edges between nodes in $\mathcal{W} \cup (\mathcal{A}\backslash\mathcal{S})$, and is independent of the edges coming out of $\mathcal{S}$. Thus, for each node $i \in \mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}}$ independently, we have that each node in $\mathcal{S}$ has an edge to i independently with probability p. We deduce that $i \in \mathcal{N}(\mathcal{S}, \mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}})$ with probability $1-(1-p)^{\kappa/(2p^{1/2})} \geq \kappa p^{1/2}/3$ for small enough p, i.i.d. for each $i \in \mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}}$. It follows from Proposition EC.1 (i) that

$$|\mathcal{N}(\mathcal{S}, \mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}})| < \frac{C^3}{2p^{3/2}} \cdot \frac{\kappa p^{1/2}}{3} \cdot \frac{3}{4} = \frac{\kappa C^3}{8p} = \frac{C}{8p}$$

occurs with probability at most $2\exp\{-(1/4)^2 \cdot C/(6p) \cdot (1/3)\} \leq \exp(-C/(300p))$ for small enough p. Now, the number of candidate subsets $\mathcal{S}$ is $\binom{1/(Cp^{3/2})}{\kappa/p^{1/2}} \leq (1/(Cp^{3/2}))^{\kappa/p^{1/2}} \leq \exp(1/p^{\varepsilon+1/2})$ for small enough p. It follows from union bound that $|\mathcal{N}(\mathcal{S}, \mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}})| < \frac{C}{8p}$ for one (or more) of these subsets $\mathcal{S}$ with probability at most $\exp(-C/(300p)) \cdot \exp(1/p^{\varepsilon+1/2}) \leq \exp(-C/(400p)) \xrightarrow{p\to 0} 0$. Thus, with high probability, $|\mathcal{N}(\mathcal{S}, \mathcal{W}\backslash\mathcal{W}_{\sim\mathcal{S}})| < \frac{C}{8p}$ occurs for no candidate subset $\mathcal{S}$; i.e., event $\mathcal{E}_6$ occurs with high probability. $\square$

Proof of Theorem 2: Lower bound

Proof of Theorem 2: Lower bound for monotone policies. Denote by m the expected steady-state number of nodes in the system, which by Little's law equals the expected steady-state waiting time. Suppose that $m \leq 1/(Cp^{3/2})$, where C is any constant larger than 36. Fix a node i, and reveal the number of nodes N in the system when i arrives. Notice that $N \leq 3m$ occurs with probability at least $1 - 1/3 = 2/3$ in steady-state by Markov's inequality. Assume that $N \leq 3m$ holds. Let $\mathcal{W}$ denote the nodes waiting in the system when i arrives (note that $|\mathcal{W}| = N$), and let $\mathcal{A}$ be the nodes that arrive in the next $3m$ time slots after node i arrives. Now, if node i leaves the system within $3m$ time slots of arriving, then i must form a two- or three-way cycle with nodes in $\mathcal{A} \cup \mathcal{W}$. The probability of forming such a cycle is bounded above by

$$\mathbb{E}[\text{Number of two-way cycles between } i \text{ and } \mathcal{A}\cup\mathcal{W}|N] \tag{EC.34}$$

$$+\,\mathbb{E}[\text{Number of three-way cycles containing } i \text{ and two nodes from } \mathcal{A}\cup\mathcal{W}|N]\,. \tag{EC.35}$$

Clearly,

$$\mathbb{E}[\text{Number of two-way cycles between } i \text{ and } \mathcal{A}\cup\mathcal{W}|N] \le 6mp^2 \le 1/C \tag{EC.36}$$

for p sufficiently small. To bound the other term we notice that

$$\begin{aligned}&\mathbb{E}[\text{Number of three-way cycles containing } i \text{ and two nodes from } \mathcal{A}\cup\mathcal{W}|N]\\ &= p^2\cdot\mathbb{E}[\text{Number of edges between nodes in } \mathcal{A}\cup\mathcal{W}|N]\,.\end{aligned} \tag{EC.37}$$

We use Corollary EC.1 to bound the expected number of edges between nodes in $\mathcal{W}$ at the time when i arrives by $N(N-1)p$ and notice that other compatibilities (j_1, j_2) for $\{j_1, j_2\} \not\subseteq \mathcal{W}$ are present independently with probability p. Hence, we have

$$\mathbb{E}[\text{Number of edges between nodes in } \mathcal{A}\cup\mathcal{W}|N] \le |\mathcal{W}\cup\mathcal{A}|(|\mathcal{W}\cup\mathcal{A}|-1)p \le 6m(6m-1)p\,.$$

Using Eq. (EC.37) we infer that

$$\begin{aligned}&\mathbb{E}[\text{Number of three-way cycles containing } i \text{ and two nodes from } \mathcal{A}\cup\mathcal{W}|N]\\ &\le 6m(6m-1)p^3 \le 36/C^2 \le 1/C\end{aligned} \tag{EC.38}$$

for $C > 36$. Using Eqs. (EC.36) and (EC.38) in (EC.35), we deduce that the probability of node i being removed within $3m$ slots is no more than $2/C$.

Combining, we get that the unconditional probability that node i stays in the system for more than $3m$ slots is at least $(2/3)(1-2/C) > 1/3$ for large enough C. This violates Markov's inequality, implying that our assumption, $m \le 1/(Cp^{3/2})$, is false. This establishes the stated lower bound. □

EC.3. Proof of Theorem 3

Preliminaries. Before proving the main theorem we need some preparation. We begin by stating a result on long chains in a static Erdős–Rényi random graph. The following result was first shown by Ajtai et al. (1981) and refined in a series of papers; see Krivelevich et al. (2013) for a historical account and the tightest result.

Proposition EC.7 **(Krivelevich et al. 2013)**. *Fix any $\varepsilon > 0$ and any $\delta > 0$. There exist C and n_0 such that for all $c > C$ and all $n > n_0$ the following occurs: consider an* $\mathrm{ER}(n, c/n)$ *directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, and let D be the length of the longest directed cycle. We have,*

$$\mathbb{P}(D > (1 - (2+\delta)ce^{-c})n) > 1 - \varepsilon\,.$$

In words, (for large c) we have a cycle containing a large fraction of the nodes with high probability. From this, we can easily obtain a similar result about the longest path starting from a specific node.

Corollary EC.2. *Fix any $\varepsilon > 0$. There exists $C < \infty$ and n_0 such that for all $c > C$ and all $n > n_0$ the following occurs: consider a set $\mathcal{V}$ of n vertices including a fixed node $v \in \mathcal{V}$, and draw an* $\mathrm{ER}(n, c/n)$ *directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Let P_v denote the length of a longest path starting at v. Then*

$$\mathbb{P}(P_v < n(1-\varepsilon)) \leq \varepsilon.$$

The proof follows relatively easily from Proposition EC.7. The idea is as follows. A sufficient condition to form a long chain from a node v is for v to be a member of the long cycle that will occur with high probability according to the proposition. Note that with constant probability e^{-c}, v will be isolated and thus not be part of the cycle, but we can make this probability small by taking C large.

Proof of Corollary EC.2. Given ε from the statement of Corollary EC.2, let $\bar{C}$ and $\bar{n}_0$ be values guaranteed to exist by applying Proposition EC.7 with $\delta = 1$ and the probability of a long chain existing is at least $1 - \varepsilon/2$.

There exists C^* such that for all $c > C^*$, $3ce^{-c} < \varepsilon/2$ as the function $f(x) = xe^{-x}$ is strictly decreasing for $x > 1$. We claim that given our ε, Corollary EC.2 holds by taking $C = \max\{\bar{C}, C^*\}$ and $n_0 = \bar{n}_0$.

Given our $\mathrm{ER}(n, c/n)$ graph where $n > n_0$ and $c > C$ and a fixed node v, let A be the event that the graph contains a cycle of length at least $(1 - 3ce^{-c})n$, and let $B \subset A$ be the event that v is in the cycle. Observe that it suffices to prove that $P(B) > 1-\varepsilon$ to establish the result, as $3ce^{-c} < \varepsilon$ by our assumption that $c > C \geq C^*$ and the definition of C^*. Thus we compute that

$$\mathbb{P}(B) = P(B|A)\mathbb{P}(A) \geq (1-\varepsilon/2)(1-\varepsilon/2) \geq 1-\varepsilon,$$

showing the result, where $\mathbb{P}(A) \geq 1-\varepsilon/2$ follows from Proposition EC.7 and $P(B|A) \geq 1-\varepsilon/2$ follows as the cycle is equally likely to pass through every node, and so, when the cycle hits $1-\varepsilon/2$ fraction of the nodes, it has probability $1-\varepsilon/2$ of hitting v. □

We extend the result above to the case of bipartite random graphs. Note that $1/p$ in the statement below roughly corresponds to n in the statements above.

COROLLARY EC.3. *Fix any $\kappa > 1$ and $\varepsilon > 0$. Then there exists $p_0 > 0$ and $C > 0$ such that the following holds: consider any $c_L \in [1/\sqrt{\kappa}, \kappa]$, any $c_R > C$, and any $p < p_0$. Let $\mathcal{L}$ be a set of c_L/p vertices and let $\mathcal{R}$ be a set of c_R/p vertices. Fix a node $v \in \mathcal{L}$. Draw $\mathcal{G} = (\mathcal{L}, \mathcal{R}, \mathcal{E})$ as an* $\mathrm{ER}(c_L/p, c_R/p, p)$ *bipartite random graph. We have,*

$$\mathbb{P}\left(P_v < 2\frac{c_L}{p}(1-\varepsilon)\right) \leq \varepsilon,$$

where, again, P_v is the length of a longest path starting at v.

The idea of the proof is as follows. First, we show with a simple calculation that a constant fraction of the nodes in $\mathcal{R}$ will have both an indegree and an outdegree of one, as $p \to 0$. We consider paths that use only this subset of nodes from $\mathcal{R}$. Such a path is equivalent to a path in a modified graph on the set of nodes $\mathcal{L}$ where there is an edge between two nodes u and v if and only if there is a path of length two between them via an intermediate node in $\mathcal{R}$ that has an indegree and an outdegree of one. Such a graph behaves (approximately) as an Erdős–Rényi graph on the nodes of $\mathcal{L}$, with the number of edges proportionate to $|\mathcal{R}|$. Thus, by ensuring that $|\mathcal{R}|$ is sufficiently large, we can apply Corollary EC.2 to obtain the result.

Proof of Corollary EC.3. Fix $\kappa > 1$ and $\varepsilon > 0$ from the statement of the corollary. For C and p_0 to be chosen later, let $c_L \in [1/\sqrt{\kappa}, \kappa]$, $c_R > C$, and let $p < p_0$ be arbitrary. Given our graph

$\mathcal{G} = (\mathcal{L}, \mathcal{R}, \mathcal{E})$, that is, $\text{ER}(c_L/p, c_R/p, p)$, consider the subgraph $\mathcal{G}' = (\mathcal{L}, \mathcal{R}', \mathcal{E}')$ of $\mathcal{G}$, where $\mathcal{R}'$ is the set of vertices in $\mathcal{R}$ with indegree of one and outdegree of one in $\mathcal{G}$, and $\mathcal{E}'$ are the edges in $\mathcal{E}$ such that both endpoints are in $\mathcal{G}'$. From this graph, we create a new directed non-bipartite digraph $\mathcal{G}'' = (\mathcal{L}, \mathcal{E}'')$, where there is an edge from $u \in \mathcal{L}$ to $v \in \mathcal{L}$ iff there is at least one node $r \in \mathcal{R}'$ such that $(u, r) \in \mathcal{E}'$ and $(r, v) \in \mathcal{E}'$. Observe that a path of length k in $\mathcal{G}''$ gives a path of length $2k$ in $\mathcal{G}$ by following the two edges in $\mathcal{G}'$ for each edge in the path on $\mathcal{G}''$. Hence, it suffices to find a path of length $(1-\varepsilon)c_L/p$ in $\mathcal{G}''$.

For any node $r \in \mathcal{R}$, let I_r be the indicator variable that r has an indegree of one and an outdegree of one. Note that these variables are independent. Further, we have,

$$\mu(p) \triangleq \mathbb{P}(I_r = 1) = \mathbb{P}(\text{Bin}(|\mathcal{L}|, p) = 1)^2 = \left(\frac{c_L}{p} p (1-p)^{\frac{c_L}{p}-1}\right)^2 \to c_L^2 \exp(-2c_L),$$

as $p \to 0$. As each of the I_r are independent, we have that

$$|\mathcal{R}'| \stackrel{d}{=} \text{Bin}\left(\frac{c_R}{p}, \mu(p)\right).$$

Letting

$$A_1(\delta_1) = \left\{(1-\delta_1)\frac{c_R}{p}\mu(p) < |\mathcal{R}'| < (1+\delta_1)\frac{c_R}{p}\mu(p)\right\}$$

we have by Proposition EC.1 that for all p,

$$\mathbb{P}(A_1(\delta_1)) \geq 1 - 2\exp\left(-\delta_1^2 \frac{c_R}{p}\frac{\mu(p)}{3}\right).$$

We can view the edges of $\mathcal{G}''$ as being generated by the following process: for each $r \in \mathcal{R}'$, pick a source and then a destination uniformly at random from $\mathcal{L}$ and add an edge from the source to the destination unless either:

- the source and destination are the same node,
- an edge between the source and destination already exists in the graph.

Thus $|\mathcal{E}''|$ is the number of nonempty bins if we throw $|\mathcal{R}'|$ balls into $(c_L/p)^2$ bins and then throw out the c_L/p bins that correspond to self-edges. (Alternatively, we can think of this process as throwing c_R/p balls, but each ball "falls through" only with probability $1-\mu(p)$. This problem was studied extensively in Samuel-Cahn 1974, but here we need only a coarse analysis.) Trivially, $|\mathcal{E}''| \le |\mathcal{R}'|$. We now show that, typically, the number of nonempty bins is almost equal to the number of balls thrown. For each $r \in \mathcal{R}'$, let X_r be the indicator that there is $\ell \in \mathcal{L}'$ such that $(\ell, r) \in \mathcal{E}'$ and $(r, \ell) \in \mathcal{E}'$. It is easy to see that the X_r are i.i.d. Bernoulli(p/c_L). For each $\{r,s\} \subset \mathcal{R}'$, let $Y_{\{rs\}}$ be the indicator that the nodes r and s are "colliding" on both their source and destination choices in $\mathcal{L}'$; i.e. there is $\ell, m \in \mathcal{L}'$, $\ell \ne m$, such that $(\ell,r),(\ell,s),(r,m),(s,m) \in \mathcal{E}'$. It is easy to see that $\mathbb{P}(Y_{\{rs\}}=1) \le p^2/c_L^2$ for each $\ell, m \in \mathcal{L}'$. We have,

$$|\mathcal{E}''| \ge |\mathcal{R}'| - \sum_{r\in\mathcal{R}'} X_r - \sum_{\{r,s\}\subset\mathcal{R}'} Y_{\{rs\}}.$$

We compute that for any fixed $\mathcal{R}'$,

$$\mathbb{E}\left[\sum_{r\in\mathcal{R}'} X_r + \sum_{\{r,s\}\subset\mathcal{R}'} Y_{\{rs\}}\right] \le |\mathcal{R}'|\frac{p}{c_L} + \binom{|\mathcal{R}'|}{2}\frac{p^2}{c_L^2} \le |\mathcal{R}'|\frac{p}{c_L} + \left(|\mathcal{R}'|\frac{p}{c_L}\right)^2$$

Letting

$$A_2(\delta_2) = \left\{\sum_{r\in\mathcal{R}'} X_r + \sum_{\{r,s\}\subset\mathcal{R}'} Y_{rs} \le \delta_2|\mathcal{R}'|\right\},$$

we have that

$$\mathbb{P}(A_2(\delta_2)) \ge 1 - \frac{p}{\delta_2 c_L} - |\mathcal{R}'|\delta_2^{-1}\left(\frac{p}{c_L}\right)^2.$$

Letting

$$B(\delta_1,\delta_2) = \left\{(1-\delta_1)(1-\delta_2)\frac{c_R}{p}\mu(p) < |\mathcal{E}''| < (1+\delta_1)\frac{c_R}{p}\mu(p)\right\},$$

we have that $B(\delta_1,\delta_2) \supset A_1(\delta_2)\cap A_2(\delta_2)$, and thus by taking complements and then applying the union bound,

$$\begin{aligned}\mathbb{P}(B(\delta_1,\delta_2)) &\ge 1 - \mathbb{P}(A_1(\delta_1)^c) - \mathbb{P}(A_2(\delta_2)^c)\\ &\ge 1 - 2\exp\left(-\delta_1^2\frac{c_R}{p}\mu(p)/3\right) - \frac{p}{\delta_2 c_L} - (1+\delta_1)c_R\mu(p)\frac{p}{\delta_2^2 c_L^2},\end{aligned}$$

thus giving us a high probability bound on the size of $|\mathcal{E}''|$ as $p \to 0$.

For our fixed ε, let $\tilde{C}$ and $\tilde{n}_0$ be C and n_0 from Corollary EC.2 such that for any $c > \tilde{C}$ and $n > \tilde{n}_0$, given a node in graph $\mathrm{ER}(n, c/n)$, there exists a path of length at least $n(1-\varepsilon/2)$ with probability at least $1-\varepsilon/2$. We now specify p_0 from the corollary to be such that for all $c_L \in [1/\sqrt{\kappa}, \kappa]$, we have $c_L/p_0 > \tilde{n}_0$, i.e., $p_0 < 1/(n_0\sqrt{\kappa})$.

Let $\tilde{\mathcal{G}} = (\mathcal{L}, \tilde{\mathcal{E}})$ be an $\mathrm{ER}(c_L/p, \tilde{C}p/c_L)$ directed random graph. We now couple $\mathcal{G}''$ (a directed $\mathrm{ER}(n, M)$ graph, where M is random but independent of the edges selected) and $\tilde{\mathcal{G}}$ (a directed $\mathrm{ER}(n,p)$ graph) in the standard way so that when $|\tilde{\mathcal{E}}| \le |\mathcal{E}''|$ then $\tilde{\mathcal{E}} \subset \mathcal{E}''$, and when $|\mathcal{E}''| \le |\tilde{\mathcal{E}}|$ then $\mathcal{E}'' \subset \tilde{\mathcal{E}}$. Thus if $\tilde{\mathcal{G}}$ has a long path and $|\tilde{\mathcal{E}}| < |\mathcal{E}''|$, then $\mathcal{G}''$ will have at least as long a path as well, as it will contain more edges on the same nodes. Let $\tilde{P}$ be the length of a longest path starting at v in $\tilde{\mathcal{G}}$. Letting

$$A_3 = \left\{ \tilde{P} > \left(1 - \frac{\varepsilon}{2}\right) \frac{c_L}{p} \right\}$$

and recalling that $p_0 < 1/(n_0\sqrt{\kappa})$ implies that $c_L/p > \tilde{n}_0$. We have by Proposition EC.7 that

$$\mathbb{P}(A_3) \ge 1 - \frac{\varepsilon}{2}.$$

We now need to show that $\mathcal{G}''$ will have more edges than $\tilde{\mathcal{G}}$ with high probability for all c_R sufficiently large (which we can control by choice of C from the statement of the corollary). We have that $|\tilde{\mathcal{E}}| \sim \mathrm{Bin}((c_L/p-1)c_L/p, \tilde{C}p/c_L)$, and thus, by Proposition EC.1, if

$$A_4(\delta_4) = \left\{ \tilde{C}(c_L/p-1)(1-\delta_4) < |\tilde{\mathcal{E}}| < \tilde{C}(c_L/p-1)(1+\delta_4) \right\}$$

then

$$\mathbb{P}(A_4(\delta_4)) \ge 1 - 2\exp\left(-\delta_4^2 \tilde{C}\left(\frac{c_L}{p} - 1\right)/3\right).$$

Now, for any fixed choice of $\delta_1, \delta_2, \delta_4$, there exists C sufficiently large such that if $c_R > C$ then for all $p < p_0$ and all $c_L \in [1/\sqrt{\kappa}, \kappa]$,

$$\tilde{C}\left(\frac{c_L}{p} - 1\right)(1+\delta_4) < (1-\delta_1)(1-\delta_2)\frac{c_R}{p}\mu(p)$$

(recall that $\mu(p)$ converges to a constant depending on only c_L as $p \to 0$). For large enough c_R, we have that

$$\{|\mathcal{E}''| > |\tilde{\mathcal{E}}|\} \subset B(\delta_1, \delta_2) \cap A_4(\delta_4),$$

as B makes $|\mathcal{E}''|$ big and A_4 ensures that $|\tilde{\mathcal{E}}|$ is small. Putting everything together, we have that

$$\left\{P > 2\frac{c_L}{p}(1-\varepsilon)\right\} \supset B(\delta_1, \delta_2) \cap A_3 \cap A_4(\delta_4),$$

and by taking complements and then applying the union bound, we obtain

$$\mathbb{P}\left(P > 2\frac{c_L}{p}(1-\varepsilon)\right) \geq 1 - \mathbb{P}(B(\delta_1,\delta_2)^c) - \mathbb{P}(A_3^c) - \mathbb{P}(A_4(\delta_4)^c) = 1 - \frac{\varepsilon}{2} - O(p),$$

showing the result. □

The required bounds on c_L, namely $c_L \in [1/\sqrt{\kappa}, \kappa]$, will correspond to bounds on p times the "typical" interval between successive times when the chain advances under a greedy policy. These intervals are distributed i.i.d. Geometric(p), and hence typically lie in the range $[1/(p\sqrt{\kappa}), \kappa/p]$ for large κ, as stated in Lemma EC.2 below. The $1/\sqrt{\kappa}$ term in the lower bound of this "typical" range is a somewhat arbitrary choice we make that facilitates a proof of Theorem 3 (a variety of other decreasing functions of κ would work as well).

LEMMA EC.2. *There exist p_0 and κ_0 such that for all $p < p_0$ and all $\kappa > \kappa_0$, if $X \sim$ Geometric(p), then*

$$\mathbb{E}\left[X \mathbb{I}_{\{X < \frac{1}{p\sqrt{\kappa}} \text{ or } X > \frac{\kappa}{p}\}}\right] \leq \frac{2}{\kappa p}.$$

Proof. By the memoryless property of the geometric distribution, for all $t > 0$,

$$\mathbb{E}[X \mid X > t] = t + \mathbb{E}[X] = t + \frac{1}{p}. \tag{EC.39}$$

Thus for all sufficiently large κ we have

$$\mathbb{E}\left[X \mathbb{I}_{X > \frac{\kappa}{p}}\right] = (1-p)^{\frac{\kappa}{p}} \left(\frac{\kappa}{p} + \frac{1}{p}\right) \tag{EC.40}$$

$$\leq e^{-\kappa} \frac{1+\kappa}{p} \tag{EC.41}$$

$$\leq \frac{1}{2\kappa p}, \tag{EC.42}$$

where (EC.40) follows from (EC.39), (EC.41) follows as $(1-p)^{1/p} \le e^{-1}$ for all p (take logarithms), and finally (EC.42) holds provided that $\kappa \ge \kappa_0$ for appropriately large κ_0.

For the remaining term, we have that for all sufficiently large κ and sufficiently small p,

$$\begin{aligned}
\mathbb{E}\left[X\mathbb{I}_{X\le\frac{1}{\sqrt{\kappa p}}}\right] &= \mathbb{E}[X] - \mathbb{E}\left[X\mathbb{I}_{X>\frac{1}{\sqrt{\kappa p}}}\right] \\
&= \frac{1}{p} - (1-p)^{\frac{1}{\sqrt{\kappa p}}}\left(\frac{1}{\sqrt{\kappa p}} + \frac{1}{p}\right) && \text{(EC.43)} \\
&\le \frac{1}{p} - \left(\frac{1-p}{e}\right)^{\frac{1}{\sqrt{\kappa}}}\left(\frac{1}{\sqrt{\kappa p}} + \frac{1}{p}\right) && \text{(EC.44)} \\
&\le \frac{1}{p} - (1-p)^{\frac{1}{\sqrt{\kappa}}}\left(1 - \frac{1}{\sqrt{\kappa}}\right)\left(\frac{1}{\sqrt{\kappa p}} + \frac{1}{p}\right) && \text{(EC.45)} \\
&= \frac{1}{p} - (1-p)^{\frac{1}{\sqrt{\kappa}}}\frac{1}{p} + (1-p)^{\frac{1}{\sqrt{\kappa}}}\frac{1}{\kappa p} \\
&\le \frac{2}{\sqrt{\kappa}} + \frac{1}{\kappa p} && \text{(EC.46)} \\
&\le \frac{3}{2\kappa p}. && \text{(EC.47)}
\end{aligned}$$

where (EC.43) follows from (EC.39). To obtain (EC.44), by Taylor's theorem, $(1-p)^{1/p} = e^{-1}(1-p/2) + o(p)$ as $p \to 0$; thus, for sufficiently small p, we have that $(1-p)^{1/p} \ge e^{-1}(1-p)$. In (EC.45) we use that $e^{-x} \ge 1-x$, in (EC.46), we use that for all x sufficiently small, $1 \ge (1-x)^n \ge 1-2xn$, and (EC.47) follows by taking κ sufficiently large. Thus the result is shown. □

EC.3.1. Proof of Theorem 3

We introduce the following notation. Let $\mathcal{G}(t) = (\mathcal{V}(t), \mathcal{E}(t), h(t))$ be the directed graph at time t describing the compatibility graph at time t. Here $h(t)$ is a special node not included in $\mathcal{V}(t)$ that is the head of the chain, which can only have outgoing edges. We denote by $\mathcal{G}(\infty) = (\mathcal{V}(\infty), \mathcal{E}(\infty), h(\infty))$ the steady-state version of this graph (which exists as we show below).

According to the greedy policy, whenever $h(t)$ forms a directed edge to a newly arriving node, a largest possible chain starting from $h(t)$ is made. Thus before the new node arrives, $h(t+1)$ will always have an indegree and outdegree of zero (as explained in Section 2), and we can only advance the chain when a newly arriving node has an incoming edge from $h(t)$. We refer to these periods between chain advancements as *intervals*. Let τ_i for $i = 1, 2, \ldots$, denote the length of the ith interval. Note that $\tau_i \sim \text{Geometric}(p)$. Let $T_0 = 0$ and $T_i = \sum_{j=1}^{i} \tau_j$ for $i = 1, 2, \ldots$, be the time at the end of the ith interval. Additionally, let $\mathcal{A}_i$ be the set of nodes that arrived during the ith interval $[T_{i-1}, T_i]$, in particular $|\mathcal{A}_i| = \tau_i$, and let $\mathcal{W}_i$ be the set of nodes that were "waiting" at the start of the i-th interval, namely, at time T_{i-1}. Thus, right before the chain is advanced, every node in the graph is either in $\mathcal{W}_i$, $\mathcal{A}_i$ or it is $h(t)$ itself.

Proof of Theorem 3: Performance of the greedy policy. We apply Proposition EC.2, taking as our Markov chain $X_i = \mathcal{G}(T_i)$, and our Lyapunov function $V(\cdot)$ to be $V(\mathcal{G}(T_i)) \triangleq |\mathcal{V}(T_i)|$. For a constant $C > 0$ to be specified later, we let α from Proposition EC.2 be $\alpha = C/p$. Thus our finite set of exceptions is $\mathcal{B} = \{\mathcal{G} = (\mathcal{V}, \mathcal{E}, h) \colon |\mathcal{V}| \le C/p\}$, the directed graphs with at most C/p nodes. Obviously our state space is countable and $\mathcal{B}$ is finite. Let $\mathcal{P}_i$ be the path of nodes that are removed from the graph in the ith interval. Thus

$$|\mathcal{V}(T_i)| = |\mathcal{V}(T_{i-1})| + |\mathcal{A}_i| - |\mathcal{P}_i|.$$

By taking $A_i = |\mathcal{A}_i| = \tau_i$ and $D_i = |\mathcal{P}_i|$, we have that $V(\cdot)$ satisfies the form of (EC.3) and the independence assumptions on A_i and D_i. As $\tau_i \sim \text{Geometric}(p)$, we have $\mathbb{E}[|\mathcal{A}_i|^2] \le 2/p^2 = 2\mathbb{E}[|\mathcal{A}_i|]^2$, and so we can take $C_1 = 2$. We set $C_2 = 2$, and so to apply Proposition EC.2, we must find $\lambda > 0$ such that for every graph $\mathcal{G} \notin \mathcal{B}$,

$$\mathbb{E}_{\mathcal{G}}\left[|\mathcal{A}_i| - \min\{|\mathcal{P}_i|, 2|\mathcal{A}_i|\}\right] \le -\lambda \mathbb{E}[|\mathcal{A}_i|], \tag{EC.48}$$

where $\mathbb{E}_{\mathcal{G}}[\cdot]$ denotes the expectation conditioned on the event $\mathcal{G}_{i-1} = \mathcal{G}$. We create an auxiliary bipartite graph $\mathcal{G}_i' = (\mathcal{A}_i, \mathcal{W}_i, \mathcal{E}_i')$, where $\mathcal{A}_i$ are the nodes on the left, $\mathcal{W}_i$ are the nodes on the right, and $\mathcal{E}_i'$ is the subset of $\mathcal{E}(T_i)$ consisting of edges (u, v) such that either $u \in \mathcal{A}_i, v \in \mathcal{W}_i$ or vice versa (thus ensuring that $\mathcal{G}_i'$ is bipartite). We let $v_i' \in \mathcal{A}_i$ be the node newly arrived at T_i that $h(T_i - 1)$ connected to. Finally, we let $\mathcal{P}_i'$ be the longest path in $\mathcal{G}_i'$ starting at v_i'. Trivially, $|\mathcal{P}_i'| \le |\mathcal{P}_i|$. Observe that $\mathcal{G}_i'$ is a $\text{ER}(|\mathcal{A}_i|, |\mathcal{W}_i|, p)$ bipartite random graph. Thus we apply Corollary EC.3 to show that $|\mathcal{P}_i'|$ is appropriately large with high probability. In particular, given arbitrary

$\varepsilon > 0$ and $\kappa > \kappa_0 > 1$, where κ_0 is to be specified later, find C and p_0 according to Corollary EC.3. Then $\mathcal{G}(T_{i-1}) \notin B$ implies that $|\mathcal{W}_i| = |\mathcal{V}(T_{i-1})| \geq C/p$. Then, if $p < p_0$ and $a \in \left[\frac{1}{p\sqrt{\kappa}}, \frac{\kappa}{p}\right]$, then by Corollary EC.3,

$$\mathbb{P}\left(|\mathcal{P}'_i| < 2|\mathcal{A}_i|(1-\varepsilon) \,\middle|\, \mathcal{A}_i| = a\right) \leq \varepsilon. \tag{EC.49}$$

We define the events E_i and F_i by

$$E_i = \left\{\mathcal{A}_i \notin \left[\frac{1}{p\sqrt{\kappa}}, \frac{\kappa}{p}\right]\right\} \qquad\qquad F_i = \{|\mathcal{P}'_i| < 2|\mathcal{A}_i|(1-\varepsilon)\}.$$

We define $Z_i \triangleq 2|\mathcal{A}_i|(1-\varepsilon)\mathbb{I}_{E_i^c \cap F_i^c}$. Thus $Z_i \leq |\mathcal{P}'_i| \leq |\mathcal{P}_i|$ by the definition of the event F_i, and $Z_i \leq 2|\mathcal{A}_i|$ by construction. We now use this to get an upper bound (EC.48) as follows. First, we have that

$$\begin{aligned}\mathbb{E}_{\mathcal{G}}\left[|\mathcal{A}_i| - \min\{|\mathcal{P}_i|, 2|\mathcal{A}_i|\}\right] &\leq \mathbb{E}_{\mathcal{G}}\left[|\mathcal{A}_i| - Z_i\right] \\ &= \mathbb{E}\left[|\mathcal{A}_i|\mathbb{I}_{E_i}\right] + \mathbb{E}\left[|\mathcal{A}_i| - Z_i \,\middle|\, E_i^c\right]\mathbb{P}(E_i^c),\end{aligned} \tag{EC.50}$$

where in (EC.50) we used that Z_i is zero on E_i. Now noting that for all $a \in \left[\frac{1}{p\sqrt{\kappa}}, \frac{\kappa}{p}\right]$, i.e., in the event E_i^c, we have that

$$\mathbb{P}\left(Z_i = 0 \,\middle|\, |\mathcal{A}_i| = a\right) = \mathbb{P}\left(F_i \,\middle|\, |\mathcal{A}_i| = a\right) \leq \varepsilon,$$

by (EC.49), and therefore

$$\mathbb{P}\left(Z_i = 2(1-\varepsilon)|\mathcal{A}_i| \,\middle|\, |\mathcal{A}_i| = a\right) = \mathbb{P}\left(F_i^c \,\middle|\, |\mathcal{A}_i| = a\right) \geq 1 - \varepsilon.$$

as well. We now compute that

$$\begin{aligned}&\mathbb{E}\left[|\mathcal{A}_i| - Z_i \,\middle|\, E_i^c\right] \\ &= \sum_{a \in \left[\frac{1}{p\sqrt{\kappa}}, \frac{\kappa}{p}\right]} \mathbb{E}\left[|\mathcal{A}_i| - Z_i \,\middle|\, |\mathcal{A}_i| = a\right]\mathbb{P}\left(|\mathcal{A}_i| = a \,\middle|\, E_i^c\right) \\ &= \sum_{a \in \left[\frac{1}{p\sqrt{\kappa}}, \frac{\kappa}{p}\right]} \mathbb{E}\left[|\mathcal{A}_i| - Z_i \,\middle|\, F_i \cap |\mathcal{A}_i| = a\right]\mathbb{P}\left(F_i \,\middle|\, |\mathcal{A}_i| = a\right)\mathbb{P}\left(|\mathcal{A}_i| = a \,\middle|\, E_i^c\right)\end{aligned}$$

$$
\begin{aligned}
&\quad + \sum_{a\in\left[\frac{1}{p\sqrt{\kappa}},\frac{\kappa}{p}\right]} \mathbb{E}\left[|\mathcal{A}_i| - Z_i \,\middle|\, F_i^c \cap |\mathcal{A}_i| = a\right] \mathbb{P}\left(F_i^c \,\middle|\, |\mathcal{A}_i| = a\right) \mathbb{P}\left(|\mathcal{A}_i| = a \,\middle|\, E_i^c\right) \\
&\le \sum_{a\in\left[\frac{1}{p\sqrt{\kappa}},\frac{\kappa}{p}\right]} \mathbb{E}\left[|\mathcal{A}_i| \,\middle|\, F_i \cap |\mathcal{A}_i| = a\right] \cdot \varepsilon \cdot \mathbb{P}\left(|\mathcal{A}_i| = a \,\middle|\, E_i^c\right) \\
&\quad + \sum_{a\in\left[\frac{1}{p\sqrt{\kappa}},\frac{\kappa}{p}\right]} \mathbb{E}\left[|\mathcal{A}_i| - 2(1-\varepsilon)|\mathcal{A}_i| \,\middle|\, F_i^c \cap |\mathcal{A}_i| = a\right] \cdot (1-\varepsilon) \cdot \mathbb{P}\left(|\mathcal{A}_i| = a \,\middle|\, E_i^c\right) \\
&\le \varepsilon \mathbb{E}\left[|\mathcal{A}_i| \,\middle|\, E_i^c\right] + (1-\varepsilon)\mathbb{E}\left[(-1+2\varepsilon)|\mathcal{A}_i| \,\middle|\, E_i^c\right] \\
&= (-1+4\varepsilon-2\varepsilon^2)\mathbb{E}\left[|\mathcal{A}_i| \,\middle|\, E_i^c\right] \\
&\le (-1+4\varepsilon)\mathbb{E}\left[|\mathcal{A}_i| \,\middle|\, E_i^c\right].
\end{aligned}
$$

Now combining this with (EC.50), we have that

$$
\begin{aligned}
\mathbb{E}\left[|\mathcal{A}_i|\mathbb{I}_{E_i}\right] + \mathbb{E}\left[|\mathcal{A}_i| - Z_i \,\middle|\, E_i^c\right]\mathbb{P}(E_i^c) &\le \mathbb{E}\left[|\mathcal{A}_i|\mathbb{I}_{E_i}\right] + (-1+4\varepsilon)\mathbb{E}\left[|\mathcal{A}_i| \,\middle|\, E_i^c\right]\mathbb{P}(E_i^c) \\
&= \mathbb{E}\left[|\mathcal{A}_i|\mathbb{I}_{E_i}\right] + (-1+4\varepsilon)\mathbb{E}\left[|\mathcal{A}_i|\mathbb{I}_{E_i^c}\right]\mathbb{P}(E_i^c) \\
&\le \frac{2}{\kappa p} + (-1+4\varepsilon)\left(\frac{1}{p} - \frac{2}{\kappa p}\right) \qquad \text{(EC.51)} \\
&\le -\frac{1}{p} + \frac{4\varepsilon}{p} + \frac{4}{\kappa p}, \\
&= -\frac{1}{p}\left(1 - 4\varepsilon - \frac{4}{\kappa}\right),
\end{aligned}
$$

where in (EC.51) we used Lemma EC.2 twice. Now, we let $\delta \triangleq 4\varepsilon + 4/\kappa$, and observe that we can make δ arbitrarily and the inequality will still hold for sufficiently small p by our choice of ε and κ_0. As we have that

$$
\mathbb{E}_{\mathcal{G}}\left[|\mathcal{A}_i| - \min\{|\mathcal{P}_i|, 2|\mathcal{A}_i|\}\right] \le -\frac{1}{p}(1-\delta) = -\mathbb{E}[|\mathcal{A}_i|](1-\delta),
$$

we can apply Proposition EC.2 with $\lambda = (1-\delta)$ to obtain that

$$
\begin{aligned}
\mathbb{E}[|\mathcal{V}(T_\infty)|] &\le \max\left\{\alpha, \frac{\max\{1, C_2-1\}^2 C_1}{\lambda}\mathbb{E}[A_k]\right\}\left(2 + \frac{2}{\lambda}\right) \\
&= \max\left\{\frac{C}{p}, \frac{2}{1-\delta}\frac{1}{p}\right\}\left(2 + \frac{2}{1-\delta}\right).
\end{aligned}
$$

Finally, recall that we are working with the "embedded Markov chain" as we are observing the process at times T_i only. We can relate the actual Markov chain to the embedded Markov chain as follows:

$$
\begin{aligned}
\mathbb{E}[|\mathcal{V}(\infty)|] &= \lim_{t\to\infty} \frac{1}{t}\sum_{s=0}^{t} |\mathcal{V}(s)| && \text{(EC.52)}\\
&= \lim_{n\to\infty} \frac{1}{T_n}\sum_{s=0}^{T_n} |\mathcal{V}(s)| && \text{(EC.53)}\\
&= \lim_{n\to\infty} \frac{n}{T_n} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}\sum_{s=T_{i-1}}^{T_i} |\mathcal{V}(s)| \\
&= p\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}\left(|\mathcal{V}(T_{i-1})|(T_i-T_{i-1}) + \frac{(T_i-T_{i-1})(T_i-T_{i-1}+1)}{2}\right) && \text{(EC.54)}\\
&\le p\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}|\mathcal{V}(T_{i-1})|(T_i-T_{i-1}) + \frac{1}{n}\sum_{i=1}^{n}(T_i-T_{i-1})^2\right). && \text{(EC.55)}
\end{aligned}
$$

Here (EC.52) follows from the positive recurrence of $\mathcal{G}(t)$. We have (EC.53) because $a_n \to a$ implies that for every subsequence a_{n_i}, we have that $a_{n_i} \to a$ as well, and using that, almost surely $T_n \to \infty$. We obtain the left term in (EC.54) by observing that T_n is the sum of n independent Geometric(p) random variables and then applying the SLLN. For the right term of (EC.54), we simply use that $|\mathcal{V}(s+1)| = |\mathcal{V}(s)|+1$ for $s \in [T_{i-1}, T_i - 1]$, and then the identity $\sum_{i=1}^{n} i = n(n+1)/2$.

We now consider each sum from (EC.55) independently. For the first sum, observing that $|\mathcal{V}(T_{i-1})|(T_i - T_{i-1})$ is a function of our positive recurrent Markov chain $\mathcal{G}(T_i)$, we have that there exists a random variable X^* such that $|\mathcal{V}(T_{i-1})|(T_i - T_{i-1}) \Rightarrow X^*$ and the average value of $|\mathcal{V}(T_{i-1})|(T_i - T_{i-1})$ converges to $\mathbb{E}[X^*]$ a.s. The convergence in distribution $|\mathcal{V}(T_{i-1})|(T_i - T_{i-1}) \Rightarrow X^*$ implies the existence of $|\tilde{\mathcal{V}}(\tilde{T}_{i-1})|(\tilde{T}_i - \tilde{T}_{i-1})$ that converges to X^* a.s. Putting these together, we have that

$$
\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} |\mathcal{V}(T_{i-1})|(T_i-T_{i-1}) &= \mathbb{E}[X^*] && \text{(EC.56)}\\
&= \mathbb{E}\left[\lim_{i\to\infty} |\tilde{\mathcal{V}}(\tilde{T}_{i-1})|(\tilde{T}_i-\tilde{T}_{i-1})\right] \\
&\le \liminf_{i\to\infty} \mathbb{E}\left[|\mathcal{V}(T_{i-1})|(T_i-T_{i-1})\right] && \text{(EC.57)}\\
&= \liminf_{i\to\infty} \mathbb{E}\left[|\mathcal{V}(T_{i-1})|\right]\mathbb{E}[T_i-T_{i-1}] && \text{(EC.58)}
\end{aligned}
$$

$$= \frac{1}{p}\mathbb{E}[|\mathcal{V}(T_\infty)|]. \tag{EC.59}$$

Here we have (EC.56) by the ergodic theorem for Markov chains, (EC.57) by Fatou's lemma, (EC.58) by the independence of $T_i - T_{i-1}$ from $\mathcal{V}(T_{i-1})$, and (EC.59) by Theorem 2 from Tweedie (1983) (alternatively, (EC.59) can be shown with a little extra work by using a simpler result from Holewijn and Hordijk 1975).

For the second sum, as $T_i - T_{i-1} = \tau_i$ are i.i.d. Geometric(p), by the SLLN,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(T_i - T_{i-1})^2 = \mathbb{E}[\tau_1^2] = \frac{2-p}{p^2} \leq \frac{2}{p^2}$$

Thus

$$\mathbb{E}[|\mathcal{V}(\infty)|] \leq p\left(\frac{1}{p}\mathbb{E}[|\mathcal{V}(T_\infty)|] + \frac{2}{p^2}\right) = \mathbb{E}[|\mathcal{V}(T_\infty)|] + \frac{2}{p},$$

showing the result, as we have for the embedded process that $\mathbb{E}[|\mathcal{V}(T_\infty)|] = \Omega(1/p)$. □

Finally, we mention that in moving from the "embedded Markov chain" back to the original Markov chain we make use of the fact that τ_i is light tailed in the sense that $\mathbb{E}[\tau_i^2] = O((\mathbb{E}[\tau_i])^2)$, to obtain a bound of $O(1/p)$ of the steady-state expected number of nodes in the system.

Proof of Theorem 3: Lower bound. Let $C = 24$. We will show that the expected steady-state waiting time W is at least $1/(Cp)$ for all p, giving the result. Assume to the contrary that there exists p such that $W \leq 1/(Cp)$. By Little's law we have that $W = \mathbb{E}[|\mathcal{V}(\infty)|] \leq 1/(Cp)$ as well. Let i be a node entering at steady-state. Let $\mathcal{W}$ be the set of nodes in the system when i arrives, leading to $\mathcal{W} \overset{d}{=} |\mathcal{V}(\infty)|$, and define the event $E_1 = \{|\mathcal{W}| \leq 3W\}$. By Markov's inequality, $\mathbb{P}(E_1) \geq 2/3$. Note that i cannot leave the system until it has an indegree of at least one. Let $\mathcal{A}$ be the first $3W$ arrivals after i, and let the event E_2 be the event that either a node from $\mathcal{W}$ or a node from $\mathcal{A}$ has an edge pointing to i. We have that

$$\mathbb{P}(E_2) = \mathbb{P}(\text{Bin}(|\mathcal{W}| + 3W, p) \geq 1),$$

making

$$\mathbb{P}(E_2 \mid E_1) \leq \mathbb{P}(\text{Bin}(6W, p) \geq 1) \leq \mathbb{P}(\text{Bin}(6/(Cp), p) \geq 1) \leq \frac{6}{C} = \frac{1}{4},$$

using the definition of E_1, and then that $W \leq 1/(Cp)$, then Markov's inequality, and finally that $C = 24$. Thus

$$\begin{aligned} W &= \mathbb{E}[\text{Waiting time of node } i] \\ &\geq 3W\mathbb{P}(E_2^c) \geq 3W\mathbb{P}(E_2^c|E_1)\mathbb{P}(E_1) \geq 3W(1-1/4)(2/3) = 3W/2 > W, \end{aligned}$$

providing the contradiction. □

EC.4. Two-way cycles with departures

We argued heuristically in Section 1 that our problem formulation, with no departures and looking for a policy that minimizes the expected waiting time, is closely related to an alternate possible formulation where agents die at a certain rate and the goal is to minimize the likelihood that an agent will die before being served. The argument is based on the fact that minimizing the expected waiting time in our formulation is the same as minimizing the expected number of nodes in the system (by Little's law), and, in the alternate formulation, the rate of agents perishing is also minimized by minimizing the expected number of nodes in the system.

In this section, we formalize this connection in the case of two-cycle removal. Consider a modified model, with agents departing/dying at rate of λp^2, and ask what policy minimizes the chance that an agent will depart unsuccessfully. We follow our proof of Theorem 1 very closely to obtain a very similar result, i.e., that the greedy policy is asymptotically optimal.

Note that a death rate of order p^2 is the scaling regime in which the likelihood an agent will die before being served remains bounded away from 0 and 1 for small p. This is exactly the scaling regime considered in Akbarpour et al. (2014); in fact the model considered in this section is identical to the model considered in that paper (with no criticality information) up to a redefinition of parameters. (Our λ corresponds to $1/d$ there, our p^2 corresponds to d/m there, and time moves m times as fast in that paper.) Following our approach, we obtain a tight lower bound, not obtained in the concurrent paper Akbarpour et al. (2014).

Define the *loss* of a policy to be the expected fraction of agents who die without being served in steady-state.

Theorem EC.1. *Fix $\lambda \in (0,\infty)$. Let x_* be the unique solution of $f(x)=0$, where $f(x)=-\lambda x+2\exp(-x)-1$. Consider the two-cycle removal setting and each agent dying with probability λp^2 in each period, independently across agents and periods. Then the loss of the greedy policy (see Definition 1) is $\lambda x_*(1+o(1))$. This is optimal, in the sense that for every periodic Markov policy (see Definition 2), the loss is at least $\lambda x_*(1+o(1))$.*

Proof of Theorem EC.1. We first compute the expected steady-state waiting time under the greedy policy. Let $\mathcal{V}(t)$ be the number of nodes in the system just before a new node arrives at time t. If an incoming node at time t forms a two-way cycle, it is executed and the remaining $|\mathcal{V}(t)|-1$ nodes each die with probability λp^2 before the next period. If no two-way cycle is formed, each of the $|\mathcal{V}(t)|+1$ nodes dies with probability λp^2 before the next period. We deduce that for all $t \geq 0$, each of the $|\mathcal{V}(t)|-1$ nodes

$$\begin{aligned}\mathbb{E}[|\mathcal{V}(t+1)|-|\mathcal{V}(t)|] &= [1-(1-p^2)^{|\mathcal{V}(t)|}](1-\lambda p^2)(|\mathcal{V}(t)|-1)\\ &\quad +(1-p^2)^{|\mathcal{V}(t)|}(1-\lambda p^2)(|\mathcal{V}(t)|+1)-|\mathcal{V}(t)|\\ &= -\lambda|\mathcal{V}(t)|p^2+(2(1-p^2)^{|\mathcal{V}(t)|}-1)(1-\lambda p^2)\\ &= -\lambda|\mathcal{V}(t)|p^2+2\exp(-p^2|\mathcal{V}(t)|)-1+O(p^2)\end{aligned}$$

Let x_* be the unique solution of $f(x)=0$, where $f(x)=-\lambda x+2\exp(-x)-1$. Note that $f'(x)=-\lambda-2\exp(-x)<-\lambda$ for all x. The idea is that for $|\mathcal{V}(t)|>x_*/p^2(1+O(p^2))$, the random walk has a negative drift, and hence we should be able to establish that $\mathbb{E}[|\mathcal{V}(t)|]$ does not exceed x_*/p^2 by much. On the other hand, for $|\mathcal{V}(t)|<x_*/p^2(1+O(p^2))$, the random walk has a positive drift, and hence we will obtain a tight characterization of the performance of the greedy policy as $\mathbb{E}_{\mathrm{gr}}[|\mathcal{V}(t)|]=(x_*/p^2)(1+o(1))$. As a side remark, we point out that $x_* \in (1/(2+\lambda),\ln 2)$ can be immediately deduced. (The extreme values of this interval correspond to the lower and upper bounds in Akbarpour et al. (2014).)

Let $\varepsilon>0$ be arbitrary. Suppose that $|\mathcal{V}(t)|>(1+\varepsilon)x_*/p^2$. Then $|\mathcal{V}(t+1)|-|\mathcal{V}(t)|$ is stochastically dominated by a random variable defined as

$$-1+2\mathrm{Bernoulli}((1-p^2)^{|\mathcal{V}(t)|})-\mathrm{Binomial}(|\mathcal{V}(t)|-1,\lambda p^2)$$

with the Bernoulli and Binomial r.v.'s being independent of each other. For small enough p this is further dominated by

$$Z_t = -1 + 2\text{Bernoulli}(\exp\{-(1+\varepsilon/2)x_*\}) - \text{Binomial}((1+\varepsilon/2)x_*/p^2, \lambda p^2).$$

Note that $\mathbb{E}[Z_t] = f(x_*(1+\varepsilon/2)) \leq -\lambda\varepsilon/2$ and $\mathbb{P}(|Z_t| > n) \leq C\exp(-n)$ for some $C < \infty$ that does not depend on p, using that $Z_t \leq 1$ and $\mathbb{E}[\exp(-Z_t)]$ is bounded above independently of p, and then using Markov's inequality on $\exp(-Z_t)$. In other words, Z_t has negative expectation and has subexponential tails where our control over the distribution of Z_t is independent of p.

Let $S_0 = 0$ and for $t \geq 1$, $S_{t+1} = (S_t + Z_t)^+$, and so S_t is a random walk with negative drift, and well-behaved independent steps, reflected at zero. This is a positive recurrent renewal process, that renews each time it hits 0. We deduce (e.g., from Proposition EC.2) that

$$\mathbb{E}[S_\infty] = C' = C'(\varepsilon, \lambda) < \infty,$$

where S_∞ is distributed as per the stationary distribution of S. We can couple the random walk $|\mathcal{V}(t)|$ with S_t such that $|\mathcal{V}(t)| \leq (1+\varepsilon)x_*/p^2 + S_t$ for all t. (When $|\mathcal{V}(t)| < (1+\varepsilon)x_*/p^2$, then we simply use the fact that the number of nodes can increase by at most 1 per time step. On the other hand, if $|\mathcal{V}(t)| \geq (1+\varepsilon)x_*/p^2$, then Z_t stochastically dominates the step size $|\mathcal{V}(t+1)| - |\mathcal{V}(t)|$ by construction.) This yields

$$\mathbb{E}[|\mathcal{V}(\infty)|] \leq (1+\varepsilon)\frac{x_*}{p^2} + \mathbb{E}[S_\infty] \leq (1+\varepsilon)\frac{x_*}{p^2} + C' \leq (1+2\varepsilon)\frac{x_*}{p^2}.$$

for small enough p, since C' does not depend on p. Thus, for every $\varepsilon > 0$, we have that

$$\lim_{p\to 0} \frac{\mathbb{E}[|\mathcal{V}(\infty)|] - x_*/p^2}{1/p^2} \leq 2x_*\varepsilon.$$

As ε is arbitrary, the result follows.

Now we establish the lower bound on $|\mathcal{V}(\infty)|$. Let v be a newly arriving node at time t, and let $\mathcal{W}$ be the nodes currently in the system that are waiting to be matched. Let I be the indicator that at the arrival time of v (just before cycles are potentially deleted), no two-way cycles between

v and any node in $\mathcal{W}$ exist. Let $\tilde{I}$ be the indicator that at the arrival time of v, no two-way cycles *that will eventually be removed* from between v and any node in $\mathcal{W}$ exist (in particular, $\tilde{I}$ depends on the future). Thus $\tilde{I} \geq I$ a.s. Let $\tilde{V}(t)$ be the number of vertices in the system before time t such that a cycle that eventually removes them has not yet arrived. We let $\tilde{V}(\infty)$ be distributed as $\tilde{V}(t)$ when the system begins in steady-state. By stationarity,

$$0 = \mathbb{E}[\tilde{V}(t+1) - \tilde{V}(t)] = \mathbb{E}_\infty[2\tilde{I} - 1 - \text{Number of deaths}] \geq \mathbb{E}_\infty[2\tilde{I} - 1] - \lambda p^2 \mathbb{E}[\tilde{V}(\infty) + 1],$$

since just after time t at most $\tilde{V}(t) + 1$ nodes are still waiting for the cycle that eventually removes them, and each of them dies with probability λp^2. Define $\tilde{x} = p^2 \mathbb{E}[\tilde{V}(\infty)]$. Thus, we obtain

$$\mathbb{E}_\infty[\tilde{I}] \leq 1/2 + (p^2 + \tilde{x})\lambda/2 \leq 1/2 + (p^2 + x)\lambda/2\,,$$

where $x = p^2 \mathbb{E}[|\mathcal{V}(\infty)|]$, since $\tilde{V} \leq |\mathcal{V}|$. Intuitively, in steady-state, the expected change in the number of vertices not yet "matched" must be zero. Thus we obtain

$$\begin{aligned} \frac{1}{2} + (p^2 + x)\lambda/2 = \mathbb{E}[\tilde{I}] &\geq \mathbb{E}[I] = \mathbb{E}[\mathbb{E}[I \mid |\mathcal{V}(\infty)|]] = \mathbb{E}\left[(1-p^2)^{|\mathcal{V}(\infty)|}\right] \\ &\geq (1-p^2)^{\mathbb{E}[|\mathcal{V}(\infty)|]} \geq \exp(-x) + O(p^2), \end{aligned}$$

by Jensen's inequality. Thus,

$$f(x) = -\lambda x + 2\exp(-x) - 1 \leq O(p^2),$$

leading to $x \geq x_* + O(p^2)$ using $f(x_*) = 0$ and $f'(x) < -\lambda$ for all x. It follows that the loss rate (for any periodic Markov policy in steady-state), which is the same as the expected number of nodes that die in each period, is bounded below by

$$\begin{aligned} &\lambda p^2 \mathbb{E}[\#\text{nodes just after new node arrives and cycles are executed, in steady-state}] \\ &= \lambda p^2 \mathbb{E}[\,|\mathcal{V}(\infty)| + 1 - \mathbb{E}[\#\text{nodes removed via cycles in one time unit, in steady-state}]\,] \\ &\geq \lambda p^2 \mathbb{E}[|\mathcal{V}(\infty)|] \\ &= \lambda x \\ &\geq \lambda x_* + O(p^2)\,, \end{aligned}$$

which implies the stated lower bound. Here we used that in expectation at most 1 node can be served via a cycle per time unit in steady-state, since the arrival rate is 1. □

References

Ajtai M, Komlós J, Szemerédi E (1981) The longest path in a random graph. *Combinatorica* 1(1):1–12.

Akbarpour M, Li S, Gharan SO (2014) Dynamic matching market design. *arXiv preprint arXiv:1402.3643* .

Alon N, Spencer J (2008) *The Probabilistic Method* (A John Wiley & Sons).

Anderson R (2014) *Stochastic models and data driven simulations for healthcare operations*. Ph.D. thesis, Massachusetts Institute of Technology.

Asmussen S (2003) *Applied probability and queues* (Springer Verlag), ISBN 0387002111.

Foster F (1953) On the stochastic matrices associated with certain queuing processes. *The Annals of Mathematical Statistics* 24(3):355–360.

Gamarnik D, Zeevi A (2006) Validity of heavy traffic steady-state approximations in generalized jackson networks. *The Annals of Applied Probability* 16(1):56–90, ISSN 1050-5164.

Glynn PW, Zeevi A, et al. (2008) Bounding stationary expectations of markov processes. *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*, 195–214 (Institute of Mathematical Statistics).

Holewijn P, Hordijk A (1975) On the convergence of moments in stationary markov chains. *Stochastic Processes and Their Applications* 3(1):55–64.

Krivelevich M, Lubetzky E, Sudakov B (2013) Longest cycles in sparse random digraphs. *Random Structures & Algorithms* 43(1):1–15.

Samuel-Cahn E (1974) Asymptotic distributions for occupancy and waiting time problems with positive probability of falling through the cells. *The Annals of Probability* 2(3):515–521.

Tweedie R (1983) The existence of moments for stationary markov chains. *Journal of Applied Probability* 191–196.