Supplemental Material for "Information and Memory in Dynamic Resource Allocation"

Kuang Xu and Yuan Zhong

August, 2019

Appendix A: Proofs

A.1. Proof of Proposition 1

Proof. Consider the quadratic Lyapunov function $L : \mathbb{Z}_+^N \times \mathcal{G} \to \mathbb{R}_+$ defined by $L(\boldsymbol{Z}) = \|\boldsymbol{Q}\|^2$. We are interested in the conditional drift $\mathbb{E}[L(\boldsymbol{Z}(t+1)) - L(\boldsymbol{Z}(t)) \,|\, \boldsymbol{Z}(t)]$. We have

$$
\begin{aligned}
& L(\boldsymbol{Z}(t+1)) - L(\boldsymbol{Z}(t)) \\
&= \|\boldsymbol{Q}(t+1)\|^2 - \|\boldsymbol{Q}(t)\|^2 \\
&= \sum_{i=1}^N \left\{ \left[(Q_i(t) - D_i(t+1))^+ + A_i(t+1) - R_i(t+1) \right]^2 - Q_i^2(t) \right\} \\
&\leqslant \sum_{i=1}^N \left\{ \left[(Q_i(t) - D_i(t+1))^+ + A_i(t+1) \right]^2 - Q_i^2(t) \right\} \qquad (65) \\
&= \sum_{i=1}^N \left\{ \left[(Q_i(t) - D_i(t+1))^+ \right]^2 + A_i^2(t+1) + 2A_i(t+1)(Q_i(t) - D_i(t+1))^+ - Q_i^2(t) \right\} \\
&\leqslant \sum_{i=1}^N \left\{ (Q_i(t) - D_i(t+1))^2 + A_i^2(t+1) + 2A_i(t+1)Q_i(t) - Q_i^2(t) \right\} \qquad (66) \\
&= \sum_{i=1}^N \left\{ -2Q_i(t)D_i(t+1) + 2A_i(t+1)Q_i(t) + D_i^2(t+1) + A_i^2(t+1) \right\}. \qquad (67)
\end{aligned}
$$

Inequality (65) follows from properties (a) and (b) of the "residual service" $\boldsymbol{R}(\cdot)$, and Inequality (66) follows from the facts that $\left[(x)^+\right]^2 \leqslant x^2$ for any $x \in \mathbb{R}$, and that for any $x, y \in \mathbb{R}_+$, $0 \leqslant (x-y)^+ \leqslant x$.

Since both $\boldsymbol{A}(t)$ and $\boldsymbol{D}(t)$ have second moments that are uniformly upper bounded, let us suppose that for some constant $K > 0$, for all time t,

$$
\sum_{i=1}^N \mathbb{E}\left[A_i^2(t) + D_i^2(t) \right] \leqslant K.
$$

Conditioning on $\boldsymbol{Z}(t)$ and taking expectations on both sides of (67) gives

$$
\begin{aligned}
\mathbb{E}\left[L(\boldsymbol{Z}(t+1)) - L(\boldsymbol{Z}(t)) \,\middle|\, \boldsymbol{Z}(t) \right] &\leqslant -2 \left\langle \mathbb{E}\left[\boldsymbol{D}(t+1) \,\middle|\, \boldsymbol{Z}(t) \right], \boldsymbol{Q}(t) \right\rangle + 2 \left\langle \boldsymbol{\alpha}, \boldsymbol{Q}(t) \right\rangle + K \\
&\leqslant -2 \left\langle \boldsymbol{d}^*, \boldsymbol{Q}(t) \right\rangle + 2 \left\langle \boldsymbol{\alpha}, \boldsymbol{Q}(t) \right\rangle + K,
\end{aligned}
$$

for some $\boldsymbol{d}^* \in \mathcal{D}^*(t+1)$. Since $\boldsymbol{\alpha} \in \mathrm{rel}\left(\mathrm{conv}^-(\mathcal{D})\right)$, there exist constants $\delta \in (0,1)$, and $p_{\boldsymbol{d}} \geqslant 0$ for each $\boldsymbol{d} \in \mathcal{D}$, such that $\sum_{\boldsymbol{d} \in \mathcal{D}} p_{\boldsymbol{d}} \leqslant 1 - \delta$ and $\boldsymbol{\alpha} \leqslant \sum_{\boldsymbol{d} \in \mathcal{D}} p_{\boldsymbol{d}} \boldsymbol{d}$. Therefore,

$$
\langle \boldsymbol{\alpha}, \boldsymbol{Q}(t) \rangle \leqslant \left\langle \sum_{\boldsymbol{d} \in \mathcal{D}} p_{\boldsymbol{d}} \boldsymbol{d}, \boldsymbol{Q}(t) \right\rangle = \sum_{\boldsymbol{d} \in \mathcal{D}} p_{\boldsymbol{d}} \langle \boldsymbol{d}, \boldsymbol{Q}(t) \rangle \leqslant (1 - \delta) \langle \boldsymbol{d}^*, \boldsymbol{Q}(t) \rangle.
$$

Thus,

$$\mathbb{E}\left[L(\boldsymbol{Z}(t+1))-L(\boldsymbol{Z}(t))\,\middle|\,\boldsymbol{Z}(t)\right]\leqslant -2\delta\left\langle \boldsymbol{d}^{*},\boldsymbol{Q}(t)\right\rangle+K\leqslant -2\delta c\|\boldsymbol{Q}(t)\|_{\infty}+K,$$

for some $c>0$, where the last inequality follows from Assumption 2.

Consider the finite set $\tilde{\mathcal{G}}\triangleq\left\{\boldsymbol{q}\in\mathbb{R}_{+}^{N}:\|\boldsymbol{q}\|_{\infty}\leqslant\frac{K}{2\delta c}+1\right\}\times\mathcal{G}$. Then we have

$$\mathbb{E}\left[L(\boldsymbol{Z}(t+1))-L(\boldsymbol{Z}(t))\,\middle|\,\boldsymbol{Z}(t)\right]\leqslant\begin{cases}K, & \text{if } \boldsymbol{Z}(t)\in\tilde{\mathcal{G}},\\ -2\delta c, & \text{if } \boldsymbol{Z}(t)\notin\tilde{\mathcal{G}}.\end{cases}$$

The positive recurrence of $\boldsymbol{Z}(\cdot)$ then follows from a standard application of the Foster-Lyapunov criteria (e.g., Tassiulas and Ephremides 1992, Hajek 2015). Q.E.D.

A.2. Proof of Theorem 2

We prove Theorem 2 in this subsection, and begin with the following technical result on the geometry of the set $\mathrm{conv}(\Pi)$.

Lemma 3. *Consider a schedule set* Π *that satisfies Assumption 1, and let* $\mathcal{E}$ *be the set of extreme points of* $\mathrm{conv}(\Pi)$*. Let* $\boldsymbol{d}^{(0)}\in\Pi$ *be an extreme point that is maximal in* $\mathcal{E}$*. Then,* $\boldsymbol{d}^{(0)}$ *is also maximal in* $\mathrm{conv}(\Pi)$*.*

Proof. Suppose that $\boldsymbol{d}^{(0)}$ is not maximal in $\mathrm{conv}(\Pi)$. Then, let $\boldsymbol{\mu}\in\mathrm{conv}(\Pi)$ be such that $\boldsymbol{d}^{(0)}\leqslant\boldsymbol{\mu}$ and $\boldsymbol{\mu}\neq\boldsymbol{d}^{(0)}$. Then, $\boldsymbol{\mu}$ can be written as a convex combination of extreme points; i.e., $\boldsymbol{\mu}=p_1\boldsymbol{d}^1+\cdots+p_r\boldsymbol{d}^r$ with $p_s>0$ for all $s=1,2,\ldots,r$, $\sum_s p_s=1$, $\boldsymbol{d}^s\in\mathcal{E}$ for all s and the $\boldsymbol{d}^s$ are all distinct. WLOG we may assume that $\boldsymbol{d}^s\neq\boldsymbol{d}^{(0)}$ for all s, since, for example, if $\boldsymbol{d}^1=\boldsymbol{d}^{(0)}$, then $\frac{\boldsymbol{\mu}-p_1\boldsymbol{d}^1}{1-p_1}=\frac{p_2}{1-p_1}\boldsymbol{d}^2+\ldots\frac{p_r}{1-p_1}\boldsymbol{d}^r\in\mathrm{conv}(\Pi)$, $\frac{\boldsymbol{\mu}-p_1\boldsymbol{d}^1}{1-p_1}\geqslant\boldsymbol{d}^{(0)}$ and $\frac{\boldsymbol{\mu}-p_1\boldsymbol{d}^1}{1-p_1}\neq\boldsymbol{d}^{(0)}$, so we can re-choose $\boldsymbol{\mu}$ to be $\frac{\boldsymbol{\mu}-p_1\boldsymbol{d}^1}{1-p_1}$.

Consider the set $\Pi'=\{\boldsymbol{d}'\in\mathbb{Z}_{+}^{N}:\boldsymbol{d}'\leqslant\boldsymbol{d}^s$ for some $s=1,2,\ldots,r\}$. By Assumption 1, $\Pi'\subseteq\Pi$. Since $\boldsymbol{d}^{(0)}\leqslant\boldsymbol{\mu}=p_1\boldsymbol{d}^1+\cdots+p_r\boldsymbol{d}^r$, it is easy to see that we can find $p'_{\boldsymbol{d}'}\geqslant 0$ for $\boldsymbol{d}'\in\Pi'$ such that $\sum_{\boldsymbol{d}'\in\Pi'}p'_{\boldsymbol{d}'}=1$ and $\sum_{\boldsymbol{d}'\in\Pi'}p'_{\boldsymbol{d}'}\boldsymbol{d}'=\boldsymbol{d}^{(0)}$. Now $\boldsymbol{d}^{(0)}$ is an extreme point, so we must have $p'_{\boldsymbol{d}'}>0\Rightarrow\boldsymbol{d}'=\boldsymbol{d}^{(0)}$. Consider any $\boldsymbol{d}'\in\Pi'$ such that $p'_{\boldsymbol{d}'}>0$. By definition, $\boldsymbol{d}'\leqslant\boldsymbol{d}^s$ for some s. Thus, $\boldsymbol{d}^{(0)}=\boldsymbol{d}'\leqslant\boldsymbol{d}^s$. However, $\boldsymbol{d}^s\neq\boldsymbol{d}^{(0)}$ by assumption, which contradicts the maximality of $\boldsymbol{d}^{(0)}$ in $\mathcal{E}$. This establishes the lemma. Q.E.D.

Lemma 4. *Consider a schedule set* Π *that satisfies Assumption 1, and let* $\mathcal{E}$ *be the set of extreme points of* $\mathrm{conv}(\Pi)$*. Then, Assumption 3 holds if and only if* $\mathrm{conv}(\Pi)$ *has two distinct extreme points that are maximal.*

Proof. Suppose that Assumption 3 does not hold. Then, the schedule set Π has a unique maximal point, which dominates all other elements of Π. Thus, this unique maximal point dominates all points in $\mathrm{conv}(\Pi)$ as well, making it the unique maximal point in $\mathrm{conv}(\Pi)$. Now suppose that

Assumption 3 holds. We will show that there are at least two distinct maximal extreme points in $\mathcal{E}$. Suppose that $\mathcal{E}$ has a unique maximal extreme point $\boldsymbol{d}^{(0)}$. Then $\boldsymbol{d}^{(0)}$ dominates all other extreme points in $\mathcal{E}$, hence all points in $\text{conv}(\mathcal{E})$ as well. But $\text{conv}(\Pi) = \text{conv}(\mathcal{E})$, so $\boldsymbol{d}^{(0)}$ also dominates all points in $\text{conv}(\Pi)$. This implies that Π has a unique maximal element as well, contradicting Assumption 3. Thus, there are at least two distinct maximal extreme points in $\mathcal{E}$, which, by Lemma 3, are also maximal in $\text{conv}(\Pi)$. Q.E.D.

LEMMA 5. *Fix any finite $k, v \in \mathbb{Z}_+$, and $\varepsilon \in (0,1)$. Consider some $\phi \in \Phi_k$ and $\psi \in \Psi_v$, and an ε-majorizing channel $\mathcal{C} = (\mathcal{X}, \mathcal{Y}, C)$ that satisfies Assumption 4. Recall $\tilde{\Lambda}(\phi,\psi)$, the capacity region under the policy pair (ϕ,ψ), defined in (18). For any $\boldsymbol{\lambda} \in \tilde{\Lambda}(\phi,\psi)$, let $\mathbb{P}_{\boldsymbol{\lambda}}(\cdot)$ denote the stationary probability of the chain $\boldsymbol{W}(\cdot)$. Then, there exists some $\delta > 0$ such that the following is true. For any $y \in \mathcal{Y}$ and $m \in \mathcal{M}_r$, and for any $\boldsymbol{\lambda} \in \tilde{\Lambda}(\phi,\psi)$,*

$$\mathbb{P}_{\boldsymbol{\lambda}}(Y(t) = y, M_r(t) = m) \geqslant \delta. \tag{68}$$

Proof. Write $C = \varepsilon C^0 + (1-\varepsilon)C^1$ as in Eq. (24), with C^0 having identical rows. Let $\delta' = \min\{C^0_{x,y} : x \in \mathcal{X}, y \in \mathcal{Y}\}$. By Assumption 4, $\delta' > 0$.

Let $\boldsymbol{\lambda} \in \tilde{\Lambda}(\phi,\psi)$, and consider any $y', y'' \in \mathcal{Y}$, and $m', m'' \in \mathcal{M}_r$ with

$$\mathbb{P}_{\boldsymbol{\lambda}}\left(Y(t+1) = y', M_r(t+1) = m' \,\middle|\, Y(t) = y'', M_r(t) = m''\right) > 0. \tag{69}$$

We claim that there exists $\tilde{\delta} > 0$ such that whenever (69) holds, then

$$\mathbb{P}_{\boldsymbol{\lambda}}\left(Y(t+1) = y', M_r(t+1) = m' \,\middle|\, Y(t) = y'', M_r(t) = m''\right) \geqslant \tilde{\delta}. \tag{70}$$

To prove the claim, note that $M_r(t+1)$ is generated from $Y(t)$ and $M_r(t)$ by ψ alone (see Eq. (14)). Thus, $M_r(t+1)$ is conditionally independent from $Y(t+1)$, given $Y(t)$ and $M_r(t)$, and we can write

$$\begin{aligned}
&\mathbb{P}_{\boldsymbol{\lambda}}\left(Y(t+1) = y', M_r(t+1) = m' \,\middle|\, Y(t) = y'', M_r(t) = m''\right) \\
=\ &\mathbb{P}_{\boldsymbol{\lambda}}\left(Y(t+1) = y' \,\middle|\, Y(t) = y'', M_r(t) = m''\right) \times \\
&\times \mathbb{P}_{\psi}\left(M_r(t+1) = m' \,\middle|\, Y(t) = y'', M_r(t) = m''\right),
\end{aligned} \tag{71}$$

where the notation $\mathbb{P}_\psi$ is used to emphasize the fact the conditional probability

$$\mathbb{P}_{\psi}\left(M_r(t+1) = m' \,\middle|\, Y(t) = y'', M_r(t) = m''\right)$$

depends only on the policy ψ, but not on $\boldsymbol{\lambda}$. Since $C = \varepsilon C^0 + (1-\varepsilon)C^1$ and $\delta' = \min_{x,y} C^0_{x,y} > 0$, $\mathbb{P}_{\boldsymbol{\lambda}}\left(Y(t+1) = y' \,\middle|\, Y(t) = y'', M_r(t) = m''\right) \geqslant \varepsilon\delta'$. Thus,

$$\begin{aligned}
&\mathbb{P}_{\boldsymbol{\lambda}}\left(Y(t+1) = y', M_r(t+1) = m' \,\middle|\, Y(t) = y'', M_r(t) = m''\right) \\
\geqslant\ &\varepsilon\delta' \mathbb{P}_{\psi}\left(M_r(t+1) = m' \,\middle|\, Y(t) = y'', M_r(t) = m''\right).
\end{aligned} \tag{72}$$

Also, by (69) and (71), we have

$$\mathbb{P}_{\psi}\left(M_r(t+1)=m' \,\middle|\, Y(t)=y'', M_r(t)=m''\right)>0. \tag{73}$$

The preceding conditional probability only depends on ψ, so we can find a uniform lower bound $\delta''>0$ with $\mathbb{P}_{\psi}\left(M_r(t+1)=m' \,\middle|\, Y(t)=y'', M_r(t)=m''\right)\geqslant \delta''$. Therefore, we have

$$\mathbb{P}_{\boldsymbol{\lambda}}\left(Y(t+1)=y', M_r(t+1)=m' \,\middle|\, Y(t)=y'', M_r(t)=m''\right)\geqslant \varepsilon\delta'\delta''. \tag{74}$$

By choosing $\tilde{\delta}=\varepsilon\delta'\delta''$, we have established the claim.

To prove the lemma, let $y_0, y\in\mathcal{Y}$ and $m_0, m\in\mathcal{M}_r$. By the irreducibility of $\boldsymbol{W}(\cdot)$, there exists $T>1$ such that

$$\mathbb{P}_{\boldsymbol{\lambda}}(Y(T)=y, M_r(T)=m \,|\, Y(1)=y_0, M_r(1)=m_0)>0.$$

Using Ineq. (70) in the claim above, it is easy to show that

$$\mathbb{P}_{\boldsymbol{\lambda}}(Y(T)=y, M_r(T)=m \,|\, Y(1)=y_0, M_r(1)=m_0)\geqslant \tilde{\delta}^T. \tag{75}$$

Since the set $\mathcal{Y}\times\mathcal{M}_r$ is finite, we can choose $y_0\in\mathcal{Y}$ and $m_0\in\mathcal{M}_r$ such that $\mathbb{P}_{\boldsymbol{\lambda}}(Y(1)=y_0, M_r(1)=m_0)>|\mathcal{Y}\times\mathcal{M}_r|^{-1}$, and T accordingly, so that

$$\mathbb{P}_{\boldsymbol{\lambda}}(Y(1)=y, M_r(1)=m)=\mathbb{P}_{\boldsymbol{\lambda}}(Y(T)=y, M_r(T)=m)\geqslant \tilde{\delta}^T/|\mathcal{Y}\times\mathcal{M}_r|. \tag{76}$$

Letting $\delta=\min_{y\in\mathcal{Y}, m\in\mathcal{M}}\tilde{\delta}^T/|\mathcal{Y}\times\mathcal{M}_r|$, we have

$$\mathbb{P}_{\boldsymbol{\lambda}}(Y(1)=y, M_r(1)=m)\geqslant\delta, \tag{77}$$

which establishes the lemma. Q.E.D.

Proof of Theorem 2. We prove Theorem 2 by contradiction. Towards this end, suppose that there exists $\varepsilon>0$ and an ε-majorizing channel $\mathcal{C}=(\mathcal{X},\mathcal{Y},C)$, such that $\rho^*_{k,v}(\mathcal{C})=1$ for some $k\geqslant 0$ and $v\in\mathbb{Z}_+$.

Let $\varepsilon'>0$. Then there exist some $k'\in\mathbb{Z}_+$, and $(\phi,\psi)\in\Phi_{k'}\times\Psi_v$ such that

$$\left(1-\frac{\varepsilon'}{2}\right)\Lambda\subseteq\tilde{\Lambda}(\phi,\psi). \tag{78}$$

By Lemma 4, let $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ be two distinct extreme points of $\operatorname{conv}(\Pi)$, both of which are maximal as well. Write $\boldsymbol{\lambda}^{(j)}=(1-\varepsilon')\boldsymbol{d}^{(j)}$ for $j=1,2$. Then, $\boldsymbol{\lambda}^{(j)}\in\tilde{\Lambda}(\phi,\psi)$ for both $j=1,2$. Consider $\mathbb{P}_{\boldsymbol{\lambda}^{(1)}}(\boldsymbol{D}(t)=\boldsymbol{d}^{(1)})$. By choosing $\varepsilon'>0$ sufficiently small, we have $\mathbb{P}_{\boldsymbol{\lambda}^{(1)}}\left(\boldsymbol{D}(t)=\boldsymbol{d}^{(1)}\right)>0$. Thus, there exist $y\in\mathcal{Y}$ and $m\in\mathcal{M}_r$ such that

$$\mathbb{P}_{\boldsymbol{\lambda}^{(1)}}\left(\boldsymbol{D}(t)=\boldsymbol{d}^{(1)} \,\middle|\, Y(t)=y, M_r(t-1)=m\right)>0. \tag{79}$$

Note that $\boldsymbol{D}(t)$ is generated by the policy ψ, only based on $Y(t)$ and $M_r(t-1)$, so whenever $\boldsymbol{\lambda} \in \tilde{\Lambda}(\phi, \psi)$,

$$\mathbb{P}_\psi \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \,\middle|\, Y(t) = y, M_r(t-1) = m \right) = \mathbb{P}_{\boldsymbol{\lambda}} \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \,\middle|\, Y(t) = y, M_r(t-1) = m \right). \tag{80}$$

This implies that

$$\begin{aligned} &\mathbb{P}_\psi \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \,\middle|\, Y(t) = y, M_r(t-1) = m \right) \\ =\ &\mathbb{P}_{\boldsymbol{\lambda}^{(1)}} \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \,\middle|\, Y(t) = y, M_r(t-1) = m \right) \\ =\ &\mathbb{P}_{\boldsymbol{\lambda}^{(2)}} \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \,\middle|\, Y(t) = y, M_r(t-1) = m \right). \end{aligned} \tag{81}$$

Thus,

$$\begin{aligned} &\mathbb{P}_{\boldsymbol{\lambda}^{(2)}} \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \right) \\ \geqslant\ &\mathbb{P}_{\boldsymbol{\lambda}^{(2)}} \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \,\middle|\, Y(t) = y, M_r(t-1) = m \right) \times \mathbb{P}_{\boldsymbol{\lambda}^{(2)}} \left(Y(t) = y, M_r(t-1) = m \right) \\ =\ &\mathbb{P}_\psi \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \,\middle|\, Y(t) = y, M_r(t-1) = m \right) \times \\ &\mathbb{P}_{\boldsymbol{\lambda}^{(2)}} \left(Y(t) = y \,\middle|\, Y(t-1) = y, M_r(t-1) = m \right) \times \mathbb{P}_{\boldsymbol{\lambda}^{(2)}} \left(Y(t-1) = y, M_r(t-1) = m \right) \\ \geqslant\ &\mathbb{P}_\psi \left(\boldsymbol{D}(t) = \boldsymbol{d}^{(1)} \,\middle|\, Y(t) = y, M_r(t-1) = m \right) \varepsilon \delta' \delta \triangleq \delta_1, \end{aligned} \tag{82}$$

where the last inequality uses the facts that (a) for an ε-majorizing channel, $\mathbb{P}(Y(t) = y \,|\, Y(t-1) = y, M_r(t-1) = m) \geqslant \varepsilon\delta'$ for any $y \in \mathcal{Y}$ (recall the definition of δ' in the proof of Lemma 5), independent of the conditioning event in the earlier time slot, and $\mathbb{P}_{\boldsymbol{\lambda}^{(2)}} (Y(t-1) = y, M_r(t-1) = m) \geqslant \delta$ with δ in (68), by Lemma 5.

Ineq. (82) implies that by choosing ε' sufficiently small, the policy pair (ϕ, ψ) cannot stabilize the system under the arrival rate vector $\boldsymbol{\lambda}^{(2)}$. This contradicts the supposition that $\boldsymbol{\lambda}^{(2)} \in \tilde{\Lambda}(\phi, \psi)$. Thus, for any $k', v \in \mathbb{Z}_+$, $\rho_{k',v}(\mathcal{C}) < 1$.

To show that $\rho_{\infty,v}(\mathcal{C}) < 1$ as well, simply note that the preceding argument in fact shows that there exists some $\delta_2 > 0$ such that $\rho_{k',v}(\mathcal{C}) \leqslant 1 - \delta_2$ for any $k', v \in \mathbb{Z}_+$. The theorem is hence established. Q.E.D.

A.3. Proof of Theorem 4

Proof. We begin by establishing the first statement of the theorem using a lifting argument. First, note that in Section 8, we will prove that $\rho^*_{k,0} = \rho^*_{0,0}$ for all $k \geqslant 0$ (see Theorem 5), so it suffices to show that

$$\rho^*_{0,0}(\mathcal{C}) \leqslant 1 - \varepsilon(1 - N^{-1}) \tag{83}$$

for any ε-majorizing channel.

Fix an ε-majorizing channel, $\mathcal{C} = (\mathcal{X}, \mathcal{Y}, C)$, and write

$$C = \varepsilon C^0 + (1-\varepsilon) C^1, \tag{84}$$

according to the decomposition in Definition 4. Use $\mathbb{P}_x^C$ to denote the probability distribution over $\mathcal{Y}$ when the input signal is x, for the channel C; use the notation $\mathbb{P}_x^{C^0}$ and $\mathbb{P}_x^{C^1}$ in a similarly manner. Since C^0 has identical rows, we drop the subscript x and use $\mathbb{P}^{C^0}$ to denote a row of C^0. Then, for each x, the probability distribution $\mathbb{P}_x^C$ is a mixture of $\mathbb{P}^{C^0}$ and $\mathbb{P}_x^{C^1}$. Thus, to generate the message $Y(t)$ from $X(t)$ through the channel $\mathcal{C}$, it is equivalent to (a) first generate a "switching" Bernoulli random variable $\Xi(t)$, with $\mathbb{P}(\Xi(t) = 0) = \varepsilon$, independent of everything else; (b) second, if $\Xi(t) = 1$, $Y(t)$ is set to be $Y^1(t)$, the message generated from $X(t)$ through the channel $(\mathcal{X}, \mathcal{Y}, C^1)$, and if $\Xi(t) = 0$, $Y(t)$ is set to be $Y^0(t)$, generated from the probability distribution $\mathbb{P}^{C^0}$. See Figure 4 for a pictorial illustration of this equivalent interpretation of the channel $\mathcal{C}$.

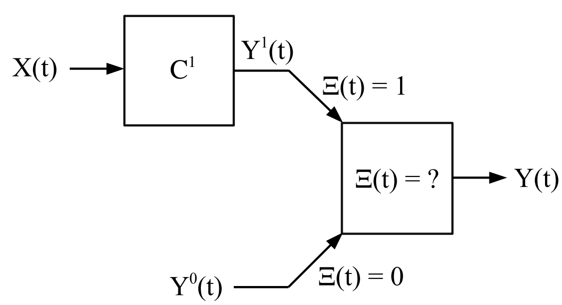

Figure 4 Equivalent construction of an ε-majorizing channel.

We shall adopt the preceding interpretation of $\mathcal{C}$ for the remainder of the proof. Consider a channel $\mathcal{C}^+$ constructed by augmenting the output of $\mathcal{C}$ with the value of the switching variable, $\Xi(t)$. Specifically, the output message of $\mathcal{C}^+$ is equal to $(Y(t), \Xi(t))$, taking values in the alphabet set $\mathcal{Y}^+ = \mathcal{Y} \times \{0, 1\}$. Note that $\mathcal{Y}^+$ is still a finite set and hence $\mathcal{C}^+$ is a valid channel. It is easy to see that for any policy pair (ϕ, ψ) operating under the channel $\mathcal{C}$, we can construct a corresponding pair (ϕ, ψ^+) operating under the channel $\mathcal{C}^+$ by simply ignoring the variable $\Xi(t)$, so that $\tilde{\Lambda}(\phi, \psi) \subseteq \tilde{\Lambda}(\phi, \psi^+)$. We hence conclude that $\rho^*_{0,0}(\mathcal{C}) \leqslant \rho^*_{0,0}(\mathcal{C}^+)$. Thus, to establish Ineq. (83), it suffices to show that

$$\rho^*_{0,0}(\mathcal{C}^+) \leqslant 1 - \varepsilon(1 - N^{-1}). \tag{85}$$

Let (ϕ, ψ^+) be a pair of memoryless encoding and allocation policies that works with the channel $\mathcal{C}^+$, and suppose that (ϕ, ψ^+) stabilizes the arrival rate vector $\boldsymbol{\lambda}$. Denote by $\boldsymbol{\mu}^0$ the N-dimensional vector where μ_i^0 is the probability under ψ^+ that the server chooses to serve queue i, conditioning on $\Xi(t)=0$, i.e.,

$$\mu_i^0 = \mathbb{P}_{\psi^+}\left(\boldsymbol{D}(t)=\boldsymbol{e}^{(i)} \,\middle|\, \Xi(t)=0\right), \quad t\in\mathbb{Z}_+. \tag{86}$$

Note that conditioning on $\Xi(t)=0$, the message $Y(t)$ is equal to $Y^0(t)$, which is independent from the signal $X(t)$. Therefore, $\boldsymbol{\mu}^0$ does not change as the arrival rate vector $\boldsymbol{\lambda}$ varies. Since the entries of $\boldsymbol{\mu}^0$ sum to no larger than 1, it follows that

$$\min_{i=1,2,\dots,N} \mu_i^0 \leqslant \frac{1}{N}. \tag{87}$$

Let $\boldsymbol{\mu}=\boldsymbol{\mu}(\phi,\psi^+,\boldsymbol{\lambda})$ be the vector of stationary service rates defined in Eq. (37) of Section 6.2. Since $\mathbb{P}(\Xi(t)=0)=\varepsilon$, we conclude that the server will choose queue i with probability at least $\mu_i^0\varepsilon$. Fix $i^*\in\arg\min_i \mu_i^0$. Then,

$$\mu_{i^*} \leqslant 1-\varepsilon\left(\sum_{i\neq i^*}\mu_i^0\right) \leqslant 1-\varepsilon(1-N^{-1}). \tag{88}$$

We claim that the system cannot be stable if $\boldsymbol{\lambda}$ admits the following form:

$$\lambda_{i^*} > 1-\varepsilon(1-N^{-1}) \quad \text{and} \quad \lambda_i = 0, \quad \forall i\neq i^*. \tag{89}$$

for if (89) were to hold, then

$$\lambda_{i^*} > 1-\varepsilon(1-N^{-1}) \geqslant \mu_{i^*}, \tag{90}$$

and queue i^* wouldn't be stable. Since ϕ and ψ^+ are arbitrary, we conclude that

$$\rho^*_{0,0}(\mathcal{C}^+) \leqslant 1-\varepsilon(1-N^{-1}), \tag{91}$$

which, by Eq. (85), proves the first statement of the theorem.

We now prove the second statement of the theorem by considering the following example. Let $\mathcal{X}=\mathcal{Y}=[N]$, so that $c_{\mathcal{X}}=c_{\mathcal{Y}}=N$. Define $I(t)$ to be the smallest index corresponding to a non-empty queue at time t, i.e.,

$$I(t) \triangleq \min\{i\in[N] : Q_i(t)>0\}, \tag{92}$$

and let the encoder simply send the signal $I(t)$ at time t, i.e., $X(t)=I(t)$. Consider the $N\times N$ channel matrix where $C_{x,x}=1-\varepsilon(1-N^{-1})$ for all $x\in[N]$, and $C_{x,y}=\varepsilon/N$ for all $x\neq y$. It is easy to verify that C is in fact ε-majorizing, where its corresponding C^0 has all entries equal to $1/N$, and C^1 is the N-by-N identity matrix. Finally, the allocation policy is simply to choose the queue whose index is equal to the message: $\boldsymbol{D}(t)=\boldsymbol{e}^{(Y(t))}$ for all t.

Let $\boldsymbol{\lambda}\in(1-\varepsilon(1-N^{-1}))\Lambda$, then $\sum_{i=1}^N \lambda_i < (1-\varepsilon(1-N^{-1}))$. Consider the aggregate queue length process $\|\boldsymbol{Q}(t)\|_1 \triangleq \sum_{i=1}^N Q_i(t)$. Note that with probability $1-\varepsilon(1-N^{-1})$, $Y(t)=X(t)$, in which case exactly 1 job would depart from the system if and only if $\|\boldsymbol{Q}(t)\|_1 > 0$. Otherwise, $Y(t) \neq X(t)$, and a job may or may not depart, depending on whether the chosen queue is empty or not. Therefore, the evolution of the process $\|\boldsymbol{Q}(t)\|_1$ is stochastically dominated by the queue length process in a single-server-single-queue system with i.i.d. arrivals, where the number of arrivals at time t is equal to $\sum_{i=1}^{N} A_i(t)$, and the number of jobs to depart from the queue at time t is a Bernoulli random variable with mean $1-\varepsilon(1-N^{-1})$, if the queue is non-empty, and zero, otherwise. A simple Lyapunov function argument would show that in this system with arrival rate $\sum_{i=1}^{N}\lambda_i < 1-\varepsilon(1-N^{-1})$ and service rate $1-\varepsilon(1-N^{-1})$ when the queue is non-empty, the queue length process is positive recurrent. This shows that the channel $\mathcal{C} = (\mathcal{X}, \mathcal{Y}, C)$ satisfies

$$\rho^*_{0,0}(\mathcal{C}) \geqslant 1-\varepsilon(1-N^{-1}), \tag{93}$$

which, in light of the first claim of the theorem, implies that $\rho^*_{0,0}(\mathcal{C})$ is in fact equal to $1-\varepsilon(1-N^{-1})$. This concludes the proof of Theorem 4.

A.4. Proof of Theorem 3

In this section, we prove Theorem 3. First, we proceed with the proof of the first statement. Let $\varepsilon \in (0,1)$, and let $\mathcal{C} = (\mathcal{X}, \mathcal{Y}, C)$ be an ε-majorizing channel. Let $\mathcal{C}^+ = (\mathcal{X}, \mathcal{Y}^+, C^+)$ be the augmented channel, which we used at the beginning of the proof of Theorem 4. We will provide a constructive characterization of the constant $\rho(\varepsilon, \Pi)$, and similar to Theorem 4, it will be sufficient to show that

$$\rho^*_{0,0}(\mathcal{C}^+) \leqslant \rho(\varepsilon, \Pi). \tag{94}$$

Let (ϕ, ψ^+) be a pair of memoryless encoding and allocation policies. Also let $p(\boldsymbol{d}) = \mathbb{P}_{\psi^+}(\boldsymbol{D}(t) = \boldsymbol{d} \,|\, \Xi(t) = 0)$ be the probability that the allocation policy chooses the schedule $\boldsymbol{d} \in \Pi$ at time t, conditioning on $\Xi(t) = 0$, and let $\boldsymbol{\mu}^0 = \sum_{\boldsymbol{d}\in\Pi} p(\boldsymbol{d})\boldsymbol{d}$ be the vector of stationary service rates, conditioning on $\Xi(t) = 0$. Note that the probabilities $p(\boldsymbol{d})$ are well-defined, independent of the arrival rate vector $\boldsymbol{\lambda}$, since for any $y \in \mathcal{Y}$, the probabilities $\mathbb{P}_{\psi^+}(\boldsymbol{D}(t) = \boldsymbol{d} \,|\, Y(t) = y)$ depend only on the policy ψ^+, $\mathbb{P}(Y(t) = y \,|\, \Xi(t) = 0)$ depend only on the channel $\mathcal{C}^+$, so

$$p(\boldsymbol{d}) = \mathbb{P}_{\psi^+}(\boldsymbol{D}(t) = \boldsymbol{d} \,|\, \Xi(t) = 0) = \sum_{y} \mathbb{P}_{\psi^+}(\boldsymbol{D}(t) = \boldsymbol{d} \,|\, Y(t) = y)\mathbb{P}(Y(t) = y \,|\, \Xi(t) = 0) \tag{95}$$

does not depend on $\boldsymbol{\lambda}$, for any $\boldsymbol{d} \in \Pi$.

Let Γ be the set of achievable vectors of stationary service rates under the policy pair (ϕ, ψ^+). Then, it is easy to see that $\Gamma \subseteq (1-\varepsilon)\mathrm{conv}(\Pi) + \varepsilon\boldsymbol{\mu}^0$. We also let $\Gamma^+ = \{\boldsymbol{\lambda} \in \mathbb{R}^N_+ : \boldsymbol{\lambda} \leqslant (1-\varepsilon)\mathrm{conv}(\Pi) + \varepsilon\boldsymbol{\mu}^0\}$. Then, we must have $\tilde{\Lambda}(\phi, \psi^+) \subseteq \Gamma^+$. Finally, let $\rho(\varepsilon, \boldsymbol{\mu}^0) = \sup\{\rho > 0 : \rho\Lambda \subseteq \Gamma^+\}$. Then, $\rho^*(\phi, \psi^+, \mathcal{C}^+) \leqslant \rho(\varepsilon, \boldsymbol{\mu}^0)$, where we recall the definition of $\rho^*(\phi, \psi^+, \mathcal{C}^+)$ in Eq. (19).

We claim that $\rho(\varepsilon, \boldsymbol{\mu}^0) < 1$ for all $\varepsilon \in (0,1)$ and $\boldsymbol{\mu}^0 \in \mathrm{conv}(\Pi)$. To prove the claim, we will show that

(a) $\rho(\varepsilon, \boldsymbol{\mu}^0)$ is achievable, in the sense that $\rho(\varepsilon, \boldsymbol{\mu}^0)\Lambda \subseteq \Gamma^+$; and

(b) $\Gamma^+ \neq \text{conv}(\Pi)$.

To prove part (a), suppose, on the contrary, that $\rho(\varepsilon, \boldsymbol{\mu}^0)\Lambda \subsetneq \Gamma^+$. Then, there exists $\boldsymbol{\mu} \in \Lambda$ such that $\rho(\varepsilon, \boldsymbol{\mu}^0)\boldsymbol{\mu} \notin \Gamma^+$. Since Γ^+ is a compact set, its complement is open, and there exists $\delta > 0$ such that $(\rho(\varepsilon, \boldsymbol{\mu}^0) - \delta)\,\boldsymbol{\mu} \notin \Gamma^+$. But this contradicts the definition of $\rho(\varepsilon, \boldsymbol{\mu}^0)$. This proves part (a).

To prove part (b), suppose on the contrary that $\Gamma^+ = \text{conv}(\Pi)$. Then, Γ^+ has two distinct extreme points $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ that are also maximal. By definition, there exists $\boldsymbol{\mu}^1 \in \text{conv}(\Pi)$ with $\boldsymbol{d}^{(1)} \leqslant (1-\varepsilon)\boldsymbol{\mu}^1 + \varepsilon\boldsymbol{\mu}^0$. By the maximality of $\boldsymbol{d}^{(1)}$, $\boldsymbol{d}^{(1)} = (1-\varepsilon)\boldsymbol{\mu}^1 + \varepsilon\boldsymbol{\mu}^0$. Since $\boldsymbol{d}^{(1)}$ is an extreme point, $\boldsymbol{d}^{(1)} = \boldsymbol{\mu}^1 = \boldsymbol{\mu}^0$. Using a similar argument, we can show that $\boldsymbol{d}^{(2)} = \boldsymbol{\mu}^0$. Thus, $\boldsymbol{d}^{(1)} = \boldsymbol{d}^{(2)}$, but this contradicts the supposition that $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ are distinct. This proves part (b).

By part (a), since $\rho(\varepsilon, \boldsymbol{\mu}^0)\Lambda \subseteq \Gamma^+$ and $\text{cl}(\Lambda) = \text{conv}(\Pi)$, we have $\rho(\varepsilon, \boldsymbol{\mu}^0)\text{conv}(\Pi) \subseteq \Gamma^+$. Note also that $\Gamma^+ \subseteq \text{conv}(\Pi)$. If $\rho(\varepsilon, \boldsymbol{\mu}^0) = 1$, then $\text{conv}(\Pi) \subseteq \Gamma^+ \subseteq \text{conv}(\Pi)$, which implies that $\text{conv}(\Pi) = \Gamma^+$, contradicting part (b). Thus, $\rho(\varepsilon, \boldsymbol{\mu}^0) < 1$, which establishes the claim.

Now, define

$$\rho(\varepsilon, \Pi) \triangleq \sup_{\boldsymbol{\mu}^0 \in \text{conv}(\Pi)} \rho(\varepsilon, \boldsymbol{\mu}^0). \tag{96}$$

It is not difficult to see that for any given $\varepsilon \in (0,1)$, $\rho(\varepsilon, \boldsymbol{\mu}^0)$ is a continuous function of $\boldsymbol{\mu}^0$. $\text{conv}(\Pi)$ is a compact set, so $\rho(\varepsilon, \Pi) = \rho(\varepsilon, \boldsymbol{\mu}^0)$ for some $\boldsymbol{\mu}^0 \in \text{conv}(\Pi)$. This implies that $\rho(\varepsilon, \Pi) = \rho(\varepsilon, \boldsymbol{\mu}^0) < 1$, establishing the first statement of the theorem.

To prove the second statement of the theorem, consider the following example. First, let $\boldsymbol{\mu}^0 \in \text{conv}(\Pi)$ be such that $\rho(\varepsilon, \Pi) = \rho(\varepsilon, \boldsymbol{\mu}^0)$, and let $(p(\boldsymbol{d}))_{\boldsymbol{d}\in\Pi}$ be a probability distribution with $\boldsymbol{\mu}^0 = \sum_{\boldsymbol{d}\in\Pi} p(\boldsymbol{d})\boldsymbol{d}$. We now turn to the construction of the channel $\mathcal{C} = (\mathcal{X}, \mathcal{Y}, C)$. Let $\mathcal{X} = \mathcal{Y} = \Pi$. Let the channel matrix C be given by: $C = \varepsilon C^0 + (1-\varepsilon)C^1$, where C^0 is the matrix with rows equal to $(p(\boldsymbol{d}))_{\boldsymbol{d}\in\Pi}$, and C^1 is the identity matrix. The encoding policy ϕ is defined as follows. If $\boldsymbol{d} \in \arg\max_{\boldsymbol{d}'\in\Pi} \langle \boldsymbol{Q}(t-1), \boldsymbol{d}' \rangle$, then $X(t) = \boldsymbol{d}$. The allocation policy ψ simply chooses the schedule $\boldsymbol{d}$ that equals the received message $Y(t)$.

Let $\boldsymbol{\lambda} < (1-\varepsilon)\text{conv}(\Pi) + \varepsilon\boldsymbol{\mu}^0$. Then, by a simple Lyapunov function argument, we can show that the policy pair (ϕ, ψ) stabilizes the arrival rate vector $\boldsymbol{\lambda}$. This implies that the capacity factor $\rho^*(\phi, \psi, \mathcal{C}) \geqslant \rho(\varepsilon, \boldsymbol{\mu}^0) = \rho(\varepsilon, \Pi)$. By the first statement, $\rho^*(\phi, \psi, \mathcal{C}) \leqslant \rho(\varepsilon, \Pi)$. Thus, $\rho^*(\phi, \psi, \mathcal{C}) = \rho(\varepsilon, \Pi)$, establishing the second statement of the theorem. Q.E.D.

A.5. Proof of Proposition 2

Proof. We will use the stationary distribution of the Markov chain $\boldsymbol{W}(\cdot)$ under the policy pair (ϕ, ψ) to design a simple encoding policy $\phi^0 \in \Phi^S$, under which we will show that

$$\boldsymbol{\mu}(\phi^0, \psi) = \boldsymbol{\mu}(\phi, \psi, \boldsymbol{\lambda}) \geqslant \boldsymbol{\lambda}. \tag{97}$$

More specifically, consider the system under the policy pair (ϕ,ψ) and arrival rate vector $\boldsymbol{\lambda}$. Because $\boldsymbol{\lambda}\in\tilde{\Lambda}(\phi,\psi)$ by assumption, the Markov chain $\boldsymbol{W}(\cdot)$ is positive recurrent, so it has a unique stationary distribution. Suppose that the chain $\boldsymbol{W}(\cdot)$ is initialized with this unique stationary distribution at time 0, and we use $\mathbb{P}^{\infty}_{\phi,\psi}(\cdot)$ to denote the corresponding probability, where the subscripts ϕ,ψ are used to highlight the encoding and allocation policies employed.

We now construct the simply encoding policy ϕ^0 for ψ. In each time slot t, the encoder memory $M_e(t)$, is set to equal the current signal $X(t)$, i.e.,

$$M_e(t)=X(t) \quad \text{under } \phi^0, \tag{98}$$

and the signals $X(\cdot)$ under ψ^0 are generated according to the conditional probabilities

$$\mathbb{P}\left(X(t+1)=x' \,\middle|\, X(t)=x, M_r(t)=m\right)=\mathbb{P}^{\infty}_{\phi,\psi}\left(X(1)=x \,\middle|\, X(0)=x, M_r(0)=m\right), \tag{99}$$

for all $x\in\mathcal{X}$ and $m\in\mathcal{M}$. That is, the signal $X(t)$ will be sampled with respect to the stationary marginal probabilities of $X(1)$ conditioned on $(X(0),M_r(0))$, under (ϕ,ψ) and $\boldsymbol{\lambda}$. Denote by $r(x,m)$ the stationary marginal probability

$$r(x,m)=\mathbb{P}^{\infty}_{\phi,\psi}(X(0)=x, M_r(0)=m), \quad x\in\mathcal{X}, m\in\mathcal{M}_r. \tag{100}$$

Since $\boldsymbol{W}(\cdot)$ is by assumption irreducible, $r(x,m)>0$ for all x and m. Thus, the conditional probabilities on the right-hand side in Eq. (99) are always well-defined.

Next, we show that under (ϕ^0,ψ), $(X(\cdot),M_r(\cdot))$ forms a homogeneous irreducible Markov chain. It is easy to check that $(X(\cdot),M_r(\cdot))$ is a homogeneous Markov chain, and we now show that it is irreducible. Define by $\mathcal{W}$ to be the product space corresponding to the state space of $W(\cdot)$: $\mathcal{W}\triangleq\mathcal{Q}\times\mathcal{M}_e\times\mathcal{X}\times\mathcal{M}_r$, and use $\mathbb{P}_{\phi^0,\psi}$ to denote probabilities under (ϕ^0,ψ). Fix $x,x'\in\mathcal{X}$ and $m,m'\in\mathcal{M}$. We have that

$$\begin{aligned}&\mathbb{P}_{\phi^0,\psi}\left(X(1)=x', M_r(1)=m' \,\middle|\, X(0)=x, M_r(0)=m\right)\\ &=\mathbb{P}^{\infty}_{\phi,\psi}\left(X(1)=x', M_r(1)=m' \,\middle|\, X(0)=x, M_r(0)=m\right).\end{aligned} \tag{101}$$

Eq. (101) can be derived as follows.

$$\begin{aligned}&\mathbb{P}^{\infty}_{\phi,\psi}(X(1)=x', M_r(1)=m' \mid X(0)=x, M_r(0)=m)\\ &\overset{(a)}{=}\mathbb{P}^{\infty}_{\phi,\psi}(X(1)=x' \mid X(0)=x, M_r(0)=m)\cdot\mathbb{P}^{\infty}_{\phi,\psi}(M_r(1)=m' \mid X(0)=x, M_r(0)=m)\\ &\overset{(b)}{=}\mathbb{P}^{\infty}_{\phi,\psi}(X(1)=x' \mid X(0)=x, M_r(0)=m)\cdot\mathbb{P}_{\psi}(M_r(1)=m' \mid X(0)=x, M_r(0)=m)\\ &\overset{(c)}{=}\mathbb{P}_{\phi^0,\psi}(X(1)=x' \mid X(0)=x, M_r(0)=m)\cdot\mathbb{P}_{\phi^0,\psi}(M_r(1)=m' \mid X(0)=x, M_r(0)=m)\\ &\overset{(d)}{=}\mathbb{P}_{\phi^0,\psi}\left(X(1)=x', M_r(1)=m' \,\middle|\, X(0)=x, M_r(0)=m\right).\end{aligned} \tag{102}$$

Step (a) follows from the fact that, under the pair (ϕ,ψ) and conditional on $X(0)$ and $M_r(0)$, the only randomness in generating the next receiver memory state $M_r(1)$ is from $U_r(0)$, so $X(1)$ is conditionally independent from $M_r(1)$; this conditional independence can also be seen from the dependencies of variables illustrated in Figure 5. In Step (b), we remove the superscript ∞ and the subscript ϕ in the second term, so as to emphasize the fact that the distribution of $M_r(1)$ is fully determined by the values of $X(0)$, $M_r(0)$ and the allocation policy ψ, regardless of how the overall chain $W(0)$ is initialized, or what encoding policy is used. Step (c) follows from the definition of the policy ϕ^0 in Eq. (99). Finally, step (d) is based on the same conditional independence property as that in Step (a) in Eq. (102), which holds under any encoding and allocation policies (Figure 5).

We proceed further to derive

$$\begin{aligned}
&\mathbb{P}_{\phi^0,\psi}\left(X(1)=x', M_r(1)=m' \,\middle|\, X(0)=x, M_r(0)=m\right)\\
&=\mathbb{P}^{\infty}_{\phi,\psi}\left(X(1)=x', M_r(1)=m' \,\middle|\, X(0)=x, M_r(0)=m\right)\\
&=r(x,m)^{-1}\mathbb{P}^{\infty}_{\phi,\psi}\left(X(1)=x', M_r(1)=m', X(0)=x, M_r(0)=m\right)\\
&=r(x,m)^{-1}\sum_{\substack{\boldsymbol{w}^1,\boldsymbol{w}^0\in\mathcal{W}:\\ x^1=x',m_r^1=m',x^0=x,m_r^0=m}}\mathbb{P}^{\infty}_{\phi,\psi}\left(\boldsymbol{W}(1)=\boldsymbol{w}^1,\boldsymbol{W}(0)=\boldsymbol{w}^0\right)\\
&\overset{(e)}{\geqslant}\sum_{\substack{\boldsymbol{w}^1,\boldsymbol{w}^0\in\mathcal{W}:\\ x^1=x',m_r^1=m',x^0=x,m_r^0=m}}\mathbb{P}^{\infty}_{\phi,\psi}\left(\boldsymbol{W}(1)=\boldsymbol{w}^1,\boldsymbol{W}(0)=\boldsymbol{w}^0\right)\\
&=\sum_{\substack{\boldsymbol{w}^1,\boldsymbol{w}^0\in\mathcal{W}:\\ x^1=x',m_r^1=m',x^0=x,m_r^0=m}}\mathbb{P}_{\phi,\psi}\left(\boldsymbol{W}(1)=\boldsymbol{w}^1 \,\middle|\, \boldsymbol{W}(0)=\boldsymbol{w}^0\right)\mathbb{P}^{\infty}_{\phi,\psi}(\boldsymbol{W}(0)=\boldsymbol{w}^0),
\end{aligned}\tag{103}$$

where we use the notation $\boldsymbol{w}^j=(\boldsymbol{q}^j,m_e^j,x^j,m_r^j)$, $j=0,1$. Here, step (e) follows from the fact that $0<r(x,m)\leqslant 1$.

The irreducibility of $\boldsymbol{W}(\cdot)$ implies that $\mathbb{P}^{\infty}_{\phi,\psi}(\boldsymbol{W}(0)=\boldsymbol{w}^0)>0$ for all $\boldsymbol{w}^0\in\mathcal{W}$. Eq. (103) thus shows that the one-step transition probability

$$\mathbb{P}_{\phi^0,\psi}\left(X(1)=x', M_r(1)=m' \,\middle|\, X(0)=x, M_r(0)=m\right)>0 \tag{104}$$

if any only if there exist two states $\boldsymbol{w}^1,\boldsymbol{w}^0\in\mathcal{W}$, with $x^1=x', m_r^1=m', x^0=x, m_r^0=m$, such that

$$\mathbb{P}_{\phi,\psi}(\boldsymbol{W}(1)=\boldsymbol{w}^1 \,|\, \boldsymbol{W}(0)=\boldsymbol{w}^0)>0. \tag{105}$$

Because $(X(\cdot),M_r(\cdot))$ is a time-homogenous Markov chain under (ϕ^0,ψ), it is not difficult to extend the same observation from Eq. (103) to over multiple time slots, and conclude that, for any fixed $T>0$,

$$\mathbb{P}_{\phi^0,\psi}\left(X(T)=x', M_r(T)=m' \,\middle|\, X(0)=x, M_r(0)=m\right)>0 \tag{106}$$

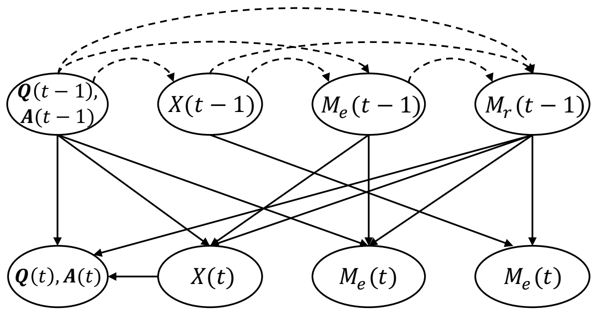

Figure 5 **A Bayesian network representation of the evolution of $\boldsymbol{W}(\cdot)$. The directed lines represent dependence relations; for example, there is a directed line from $M_r(t-1)$ to $M_r(t)$, so $M_r(t)$ is dependent on $M_r(t-1)$. The set of dashed lines is a general representation of arbitrary dependence among $\boldsymbol{Q}(t-1)$, $\boldsymbol{A}(t-1)$, $X(t-1)$, $M_e(t-1)$ and $M_r(t-1)$, whereas solid lines indicate dependence relations from the policy update equations (12), (13) and (14).**

if and only if there exist $\boldsymbol{w}^1, \boldsymbol{w}^0 \in \mathcal{W}$, with $x^1 = x', m_r^1 = m', x^0 = x, m_r^0 = m$, such that

$$\mathbb{P}_{\phi,\psi}(\boldsymbol{W}(T) = \boldsymbol{w}^1 \,|\, \boldsymbol{W}(0) = \boldsymbol{w}^0) > 0. \tag{107}$$

Since the chain $\boldsymbol{W}(\cdot)$ under (ϕ, ψ) is irreducible, $(X(\cdot), M_r(\cdot))$ under (ϕ^0, ψ) is also irreducible.

We now proceed to show that

$$\boldsymbol{\mu}(\phi^0, \psi) = \boldsymbol{\mu}(\phi, \psi, \boldsymbol{\lambda}).$$

By the irreducibility of the finite-state-space chain $(X(\cdot), M_r(\cdot))$ under (ϕ^0, ψ), it has a unique stationary distribution, which we denote by $(\bar{r}(x,m))_{x\in\mathcal{X}, m\in\mathcal{M}_r}$:

$$\bar{r}(x,m) = \mathbb{P}^{\infty}_{\phi^0,\psi}(X(0) = x, M_r(0) = m), \quad x \in \mathcal{X}, m \in \mathcal{M}_r. \tag{108}$$

Here $\mathbb{P}^{\infty}_{\phi^0,\psi}$ denotes the probability associated with this stationary distribution. By definition, for each $x' \in \mathcal{X}$ and $m' \in \mathcal{M}_r$, $\bar{r}(x', m')$ satisfies the balance equation

$$\bar{r}(x', m') = \sum_{x\in\mathcal{X}, m\in\mathcal{M}_r} \mathbb{P}_{\phi^0,\psi}(X(1) = x', M_r(1) = m' \,|\, X(0) = x, M_r(0) = m)\bar{r}(x,m). \tag{109}$$

By definition, we also have

$$r(x', m') = \sum_{x\in\mathcal{X}, m\in\mathcal{M}_r} \mathbb{P}^{\infty}_{\phi,\psi}(X(1) = x', M_r(1) = m' \,|\, X(0) = x, M_r(0) = m) r(x,m). \tag{110}$$

Thus, by (101), $r(x', m')$ also satisfies the balance equation (109), and since $(\bar{r}(x,m))_{x,m}$ is unique, we must have

$$\bar{r}(x,m) = r(x,m), \quad \forall x \in \mathcal{X}, m \in \mathcal{M}_r. \tag{111}$$

This implies that $\boldsymbol{\mu}(\phi^0, \psi) = \boldsymbol{\mu}(\phi, \psi, \boldsymbol{\lambda})$. By Lemma 1 and the positive recurrence of $\boldsymbol{W}(\cdot)$ under (ϕ, ψ) and $\boldsymbol{\lambda}$, We have $\boldsymbol{\mu}(\phi, \psi, \boldsymbol{\lambda}) \geqslant \boldsymbol{\lambda}$. Therefore, $\boldsymbol{\mu}(\phi^0, \psi) \geqslant \boldsymbol{\lambda}$ and this completes the proof of Proposition 2. Q.E.D.

A.6. Proof of Lemma 2

Proof. By the definition of $\rho^*_{k,v}$, for all $\varepsilon \in (0, \rho^*_{k,v})$, there exist $\phi \in \Phi_k$ and $\psi \in \Psi_l$, such that

$$\left(\rho^*_{k,v} - \frac{\varepsilon}{3}\right)\Lambda \subseteq \tilde{\Lambda}(\phi, \psi). \tag{112}$$

By Ineq. (112) and Proposition 2, we have that for all $i \in [|\mathcal{F}|]$, there exists $\phi_i \in \Pi^S$, such that

$$\boldsymbol{\mu}(\phi_i, \psi) \geqslant \left(\rho^*_{k,v} - \frac{\varepsilon}{3}\right)\boldsymbol{\mu}_i^{\mathcal{F}}, \tag{113}$$

and the chain $(X(\cdot), M_r(\cdot))$ is irreducible under (ϕ_i, ψ).

Let $\delta > 0$, and consider perturbations ψ^δ and ϕ_i^δ of the policies ψ and ϕ_i for $i \in [|\mathcal{F}|]$. Given the current signal $X(t)$ and receiver memory $M_r(t)$, ψ^δ updates $M_r(t+1) = M_r(t)$ with probability $\delta > 0$, and it updates $M_r(t+1)$ according to ψ with probability $1-\delta$; independently, ϕ_i^δ updates $X(t+1) = X(t)$ with probability δ, and updates $X(t+1)$ according to ϕ_i with probability $1-\delta$. Then, $(X(\cdot), M_r(\cdot))$ is clearly aperiodic under $(\phi_i^\delta, \psi^\delta)$. Furthermore, by choosing $\delta = \delta(\varepsilon)$ sufficiently small, we can also make $(X(\cdot), M_r(\cdot))$ irreducible under $\left(\phi_i^{\delta(\varepsilon)}, \psi^{\delta(\varepsilon)}\right)$, and have

$$\boldsymbol{\mu}\left(\phi_i^{\delta(\varepsilon)}, \psi^{\delta(\varepsilon)}\right) \geqslant \frac{\rho^*_{k,v} - 2\varepsilon/3}{\rho^*_{k,v} - \varepsilon/3} \cdot \boldsymbol{\mu}(\phi_i, \psi). \tag{114}$$

Combining Ineqs. (113) and (114), we have

$$\boldsymbol{\mu}\left(\phi_i^{\delta(\varepsilon)}, \psi^{\delta(\varepsilon)}\right) \geqslant \left(\rho^*_{k,v} - \frac{2}{3}\varepsilon\right) \cdot \boldsymbol{\mu}(\phi_i, \psi). \tag{115}$$

With a slight abuse of notation, we write ϕ_i^ε for $\phi_i^{\delta(\varepsilon)}$, $i \in [|\mathcal{F}|]$, and ψ^ε for $\psi^{\delta(\varepsilon)}$. Then, Ineq. (115) implies that $(\rho^*_{k,v} - \varepsilon)\Lambda \subseteq \operatorname{conv}^-(\boldsymbol{\mu}(\Phi^\varepsilon, \psi^\varepsilon))$. Q.E.D.

Appendix B: Infinite-Memory Receiver

In this section, we consider the case where the size of the receiver memory is infinite, i.e., $v = \infty$. The main result of this section is Theorem 7 (Item 3 of Theorem 1), which states that $\rho^*_{0,\infty}(\mathcal{C}) = 1$, for any informative channel $\mathcal{C}$. In other words, we do not need to maintain any encoder memory in order to recover the maximal capacity region, when the receiver is equipped with infinite memory.

THEOREM 7. $\rho^*_{0,\infty}(\mathcal{C}) = 1$.

The rest of this section is organized as follows. In Section B.1, we provide a precise description of the encoding-allocation policy pair – *Episodic Greedy Learning* (EGL) – that will be used to prove Theorem 7. Then, in Section B.2, we prove the main result proper. Finally, Appendix B.3 contains some discussion on a simple modification to EGL to the case of no memory-feedback, where we can show that $\rho^*_{\log N,\infty} = 1$, when the encoder does not have memory-feedback from the receiver.

B.1. Episodic Greedy Learning

To prove Theorem 7, we first describe the encoding-allocation policy pair (ϕ, ψ) that will be used, which we call *Episodic Greedy Learning* (EGL), where $\phi \in \Phi_0$ and $\psi \in \Psi_\infty$. At a high level, the encoding policy ϕ is designed in such a way that the encoded signals are only used for the receiver to *estimate* the arrival rate vector, from the corresponding messages received. The allocation policy ψ operates in *episodes* – these are fixed-length blocks of time slots – each one of which consists of two phases:

1. Phase 1: *Learning*. The allocation policy learns an estimator of the arrival rate vector.
2. Phase 2: *Deployment*. The allocation policy chooses a randomized schedule in each time slot, whose expectation strictly dominates the estimator produced in Phase 1.

The idea behind the allocation policy is that if (a) the length of the deployment phase is substantially longer than that of the learning phase, and (b) the arrival rate estimator generated from the learning phase is reasonably accurate, then the amount of service dedicated to each queue should be greater than the arrivals during each episode, in expectation.

Before we proceed to a precise description of EGL, let us provide some remarks on the use of memory-feedback under the infinite-receiver-memory setting. On the one hand, with memory-feedback, the encoder potentially has access to a lot of past system information from the receiver, so it may seem intuitive that the encoder need not be equipped with any memory of her own. On the other hand, as we will see in the policy description, the memory-feedback is only used to *synchronize* the encoder and the receiver, so that the encoder knows the exact time in each episode. In fact, with some simple changes to the policy pair described in this section, it is possible to show that without memory-feedback, $\rho^*_{\log N,\infty}(\mathcal{C}) = 1$; see Appendix B.3 for details.

We now describe the EGL policy pair in detail. Let $B \in \mathbb{N}$ be the length of an episode, and B_1 and $B - B_1$ be the lengths of the learning and deployment phases, respectively. We also assume that B_1 is divisible by N. Note that, because the encoder has access to the receiver memory through memory-feedback, we may assume that by recording time in M_r, the encoder knows the exact time relative to the start of the episode.

Encoding Policy. Recall that the channel $\mathcal{C}$ is informative, which implies that we can find $x_1, x_2 \in \mathcal{X}$ and $y_1 \in \mathcal{Y}$ such that

$$q_1 \triangleq \mathbb{P}(Y(1) = y_1 \,|\, X(1) = x_1) < \mathbb{P}(Y(1) = y_1 \,|\, X(1) = x_2) \triangleq q_2. \tag{116}$$

During each time slot, the encoding policy first observes the current time t relative to the start of the episode. Then, for $i \in [N]$, if $i \equiv t \bmod N$, the encoder sets $X(t) = x_1$ if $A_i(t-1) = 1$, and sets $X(t) = x_2$ otherwise. Thus, the encoder observes the N queues in a round-robin manner, based on which the signals are decided. Note that the signal $X(t)$ only depends on $\boldsymbol{A}(t-1)$ and $M_r(t)$ (via memory-feedback), and the policy does not require any encoder memory. Let us also note that even though we have specified the encoder's decisions for the entire duration of the episode, this is not necessary, since the allocation policy will not be using the outputs of the channel during the deployment phase. The decisions of the encoding policy for the entire episode are provided for concreteness, and also for ease of reference in Appendix B.3, when we discuss the case of no memory-feedback.

Allocation policy: To define the allocation policy, we first introduce some notation. For $\boldsymbol{x} \in \mathbb{R}_+^N$ and a closed convex set $\mathcal{X} \subset \mathbb{R}_+^N$, define $\text{proj}(\boldsymbol{x}, \mathcal{X})$ to be *the scaled projection of* $\boldsymbol{x}$ *to the outer boundary of* $\mathcal{X}$:

$$\text{proj}(\boldsymbol{x}, \mathcal{X}) \triangleq a\boldsymbol{x}, \tag{117}$$

where $a = \sup\{\tilde{a} \in \mathbb{R}_+ : \tilde{a}\boldsymbol{x} \in \mathcal{X}\}$.

1. *Learning Phase.* The allocation policy generates an estimator for $\boldsymbol{\lambda}$, denoted by $\widehat{\boldsymbol{\lambda}}$, as follows. For each $i \in [N]$, denote by $\widehat{p}_i$ the empirical frequency that the symbol y_1 is observed, out of all time slots t for which $i = t \bmod N$, i.e.,

$$\widehat{p}_i \triangleq \frac{1}{B_1/N} \sum_{t: t = i \bmod N} \mathbb{I}_{\{Y(t) = y_1\}}. \tag{118}$$

   Then, set

$$\widehat{\lambda}_i = \frac{q_2 - \widehat{p}_i}{q_2 - q_1}, \quad i = 1, \ldots, N. \tag{119}$$

   We do not need to specify details of the allocation decisions in this phase, since the primary function of this phase is to learn an estimator of the arrival rate vector.

2. *Deployment Phase.* Let $\alpha > 0$ be a parameter, and let

$$\widehat{\boldsymbol{\lambda}}^+ \triangleq \left(\max\{0, \widehat{\lambda}_1\}, \ldots, \max\{0, \widehat{\lambda}_N\}\right). \tag{120}$$

   Consider the vector $\text{proj}\left(\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1}, \text{cl}(\Lambda)\right)$, the scaled projection of $\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1}$ to the boundary of $\text{cl}(\Lambda)$, where we recall that $\mathbf{1}$ is the vector of all ones, and $\text{cl}(\Lambda)$ is the closure of Λ. Since $\text{proj}\left(\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1}, \text{cl}(\Lambda)\right)$ is on the outer boundary of $\text{cl}(\Lambda)$, there exists a random schedule $\boldsymbol{D}$, distributed over the set Π of schedules, so that

$$\mathbb{E}\left[\boldsymbol{D} \,\middle|\, \widehat{\boldsymbol{\lambda}}^+\right] = \text{proj}\left(\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1}, \text{cl}(\Lambda)\right)^{8}. \tag{121}$$

   During the deployment phase, the allocation vector $\boldsymbol{D}(t)$ chosen by the allocation policy are i.i.d. samples of the random schedule $\boldsymbol{D}$.

[8] Note that in Eq. (121), we have written the expectation of $\boldsymbol{D}$ as being conditioned on $\widehat{\boldsymbol{\lambda}}^+$, which is itself random, and the distribution of $\boldsymbol{D}$ depends on the realization of $\widehat{\boldsymbol{\lambda}}^+$.

Memory Requirement. As we explained earlier, the encoder does not require any memory, due to the presence of memory-feedback. The receiver memories are used in the following ways:

1. *Time-keeping*: $\log(B)$ bits used to keep track of the relative time past since the start of an episode.
2. *Learning*: B_1 bits are used to store the messages received during the learning phase, one bit for each of the B_1 times slots.

Thus, altogether the allocation policy uses $B_1 + \log(B)$ bits of memory.

B.2. Proof of Theorem 7

Proof. The following simple fact will be used throughout the proof: since Λ is a bounded set, there exists $K > 0$ such that for all $\boldsymbol{\lambda}' \in \Lambda$, $\boldsymbol{\lambda} \leqslant K\mathbf{1}$.

Let $\varepsilon \in (0,1)$. Consider the EGL policy pair $(\phi, \psi) \in \Phi_0 \times \Psi_\infty$ described in Section B.1. We will find suitable values of B_1, B and α, for which the capacity factor $\rho^*(\phi, \psi, \mathcal{C}) \geqslant 1 - \varepsilon$.

Let the arrival rate vector be $\boldsymbol{\lambda} \in (1-\varepsilon)\Lambda$. By Assumption 2, for all $i \in [N]$, we have $\boldsymbol{e}^{(i)} \in \Pi$. This implies that

$$\frac{\varepsilon}{N}\mathbf{1} = \varepsilon \sum_{i=1}^{N} \frac{1}{N}\boldsymbol{e}^{(i)} \leqslant \varepsilon \mathrm{conv}(\Pi).$$

Recalling $\boldsymbol{\lambda} \in (1-\varepsilon)\Lambda$, and $\Lambda = \mathrm{conv}^-(\Pi)$, we have $\boldsymbol{\lambda} + \frac{\varepsilon}{N}\mathbf{1} \in \Lambda$.

By construction, the estimator $\widehat{p}_i$ defined in (118) is the empirical average from B_1/N i.i.d. samples of a Bernoulli distribution with mean $\lambda_i q_1 + (1-\lambda_i)q_2$. By the law of large numbers and the union bound, there exists B_1^* such that

$$\mathbb{P}\left(\|\widehat{\boldsymbol{\lambda}}^+ - \boldsymbol{\lambda}\|_\infty \leqslant \frac{\varepsilon}{3N}\right) \geqslant 1 - \frac{\varepsilon}{12KN}, \quad \forall B_1 \geqslant B_1^*. \tag{122}$$

For the rest of the proof, set

$$B_1 = B_1^* N, \quad B = \frac{8KN}{\varepsilon} \cdot B_1, \quad \text{and} \quad \alpha = \frac{2\varepsilon}{3N}. \tag{123}$$

Consider the event $E = \left\{\|\widehat{\boldsymbol{\lambda}}^+ - \boldsymbol{\lambda}\|_\infty \leqslant \frac{\varepsilon}{3N}\right\}$. We claim that under event E, we have

$$\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1} \in \Lambda, \quad \widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1} \geqslant \boldsymbol{\lambda} + \frac{\varepsilon}{3N}\mathbf{1}, \quad \text{and} \quad \mathrm{proj}(\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1}, \mathrm{cl}(\Lambda)) \geqslant \widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1}. \tag{124}$$

First, observe that under event E, $\widehat{\boldsymbol{\lambda}}^+ \leqslant \boldsymbol{\lambda} + \frac{\varepsilon}{3N}\mathbf{1}$, so $\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1} \leqslant \boldsymbol{\lambda} + \frac{\varepsilon}{N}\mathbf{1}$. But $\boldsymbol{\lambda} + \frac{\varepsilon}{N}\mathbf{1} \in \Lambda$, so $\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1} \in \Lambda$ as well. Second, under event E we have $\widehat{\boldsymbol{\lambda}}^+ \geqslant \boldsymbol{\lambda} - \frac{\varepsilon}{3N}\mathbf{1}$, so $\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1} \geqslant \boldsymbol{\lambda} + \frac{\varepsilon}{3N}\mathbf{1}$. Finally, for any $\boldsymbol{\lambda}' \in \Lambda$, $\mathrm{proj}(\boldsymbol{\lambda}', \mathrm{cl}(\Lambda)) \geqslant \boldsymbol{\lambda}'$, so if $\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1} \in \Lambda$, then $\mathrm{proj}(\widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1}, \mathrm{cl}(\Lambda)) \geqslant \widehat{\boldsymbol{\lambda}}^+ + \alpha\mathbf{1}$. This proves the claim.

Let us now consider the randomized schedule $\boldsymbol{D}$ chosen during the deployment phase. By construction, $\mathbb{E}\left[\boldsymbol{D}\,\middle|\,\widehat{\boldsymbol{\lambda}}^{+}\right]=\operatorname{proj}\left(\widehat{\boldsymbol{\lambda}}^{+}+\alpha\mathbf{1},\operatorname{cl}(\Lambda)\right)$. By Eqs. (122) and (124), we have that

$$\begin{aligned}
&\mathbb{E}\left[\boldsymbol{D}\right]\\
&=\mathbb{E}\left[\mathbb{E}\left[\boldsymbol{D}\,\middle|\,\widehat{\boldsymbol{\lambda}}^{+}\right]\right]=\mathbb{E}\left[\operatorname{proj}\left(\widehat{\boldsymbol{\lambda}}^{+}+\alpha\mathbf{1},\operatorname{cl}(\Lambda)\right)\right]\\
&\geqslant\mathbb{P}\left(\operatorname{proj}(\widehat{\boldsymbol{\lambda}}^{+}+\alpha\mathbf{1},\operatorname{cl}(\Lambda))\geqslant\widehat{\boldsymbol{\lambda}}^{+}+\alpha\mathbf{1},\ \widehat{\boldsymbol{\lambda}}^{+}+\alpha\mathbf{1}\geqslant\boldsymbol{\lambda}+\frac{\varepsilon}{3N}\mathbf{1}\right)\left(\boldsymbol{\lambda}+\frac{\varepsilon}{3N}\mathbf{1}\right)\\
&\geqslant\mathbb{P}\left(E\right)\left(\boldsymbol{\lambda}+\frac{\varepsilon}{3N}\mathbf{1}\right)\geqslant\left(1-\frac{\varepsilon}{12KN}\right)\left(\boldsymbol{\lambda}+\frac{\varepsilon}{3N}\mathbf{1}\right)\\
&\geqslant\boldsymbol{\lambda}+\frac{\varepsilon}{3N}\mathbf{1}-\frac{\varepsilon}{12KN}\cdot K\mathbf{1}\\
&=\boldsymbol{\lambda}+\frac{\varepsilon}{4N}\mathbf{1},
\end{aligned}\tag{125}$$

where, in the last inequality, we have used the fact that $\boldsymbol{\lambda}+\frac{\varepsilon}{3N}\mathbf{1}\in\Lambda$, so $\boldsymbol{\lambda}+\frac{\varepsilon}{3N}\mathbf{1}\leqslant K\mathbf{1}$.

Denote by $\boldsymbol{A}[n]$ and $\boldsymbol{D}[n]$ the vector representing the total number of jobs that arrive and the total amount of service dedicated during the nth episode, respectively. We have that for all $n\in\mathbb{N}$

$$\mathbb{E}(\boldsymbol{A}[n])=B\boldsymbol{\lambda},\quad\text{and}\quad\mathbb{E}[\boldsymbol{D}[n]]=(B-B_1)\mathbb{E}[\boldsymbol{D}].\tag{126}$$

Using (123), (125) and (126), we have

$$\begin{aligned}
\mathbb{E}[\boldsymbol{D}[n]]&\geqslant\left(1-\frac{\varepsilon}{8KN}\right)B\left(\boldsymbol{\lambda}+\frac{\varepsilon}{4N}\mathbf{1}\right)\\
&\geqslant B\left(\boldsymbol{\lambda}+\frac{\varepsilon}{4N}\mathbf{1}\right)-\frac{\varepsilon}{8KN}\cdot BK\mathbf{1}\\
&=B\left(\boldsymbol{\lambda}+\frac{\varepsilon}{8N}\mathbf{1}\right),
\end{aligned}\tag{127}$$

where in the last inequality we have used the fact that $\boldsymbol{\lambda}+\frac{\varepsilon}{4N}\mathbf{1}\in\Lambda$, so $\boldsymbol{\lambda}+\frac{\varepsilon}{4N}\mathbf{1}\leqslant K\mathbf{1}$.

Similar to the proof of Theorem 6, we consider the Markov chain $\boldsymbol{W}[\cdot]$ sampled at the beginnings of the episodes. Noting that $\{(\boldsymbol{A}[n],\boldsymbol{D}[n])\}$ are i.i.d. sequences, and using (126), (127), and Proposition 1, the sampled chain $\boldsymbol{W}[\cdot]$ is positive recurrent under the EGL pair (ϕ,ψ) with parameters given in (123), whenever $\boldsymbol{\lambda}\in(1-\varepsilon)\Lambda$. This implies that $\rho^*(\phi,\psi,\mathcal{C})\geqslant 1-\varepsilon$. But $\varepsilon\in(0,1)$ is arbitrary, so we must have $\rho^*_{0,\infty}(\mathcal{C})=1$. Q.E.D.

B.3. Infinite Receiver Memory: No Memory-Feedback

In this subsection, we consider the case where the encoder does not have memory-feedback from the receiver; in other words, other than the allocation vector $\boldsymbol{D}(t)$ that the encoder can observe from the receiver in each time slot, the encoder does not have access to any content of the receiver memory. We still assume that the receiver is equipped with infinite memory. Then, we can prove the following

THEOREM 8. *Under no memory-feedback,* $\rho^*_{\log N,\infty}(\mathcal{C})=1$.

Proof Sketch. The proof of the theorem uses a simple, modified version of EGL, and is then essentially that of Theorem 7 verbatim. Hence, we only provide a proof sketch here.

Recall the EGL policy pair described in Section B.1. Consider the following changes to EGL:

1. The encoding policy still observes the queues in a round-robin manner, and sends signals in the same way as EGL. However, instead of keeping track of the time relative to the start of the current episode, the encoder memory maintains the index of the queue currently observed. More specifically, if the current time is t and $t \equiv i \bmod N$, then $M_e(t) = i$. The memory content is updated as $M_e(t+1) = M_e(t) + 1$ if $M_e(t) < N$, and $M_e(t+1) = 1$ if $M_e(t) = N$. Note that the encoder only needs $\log N$ bits of memory to keep track of the queue indices, and does not require memory-feedback from the receiver to know which queue to observe.
2. At time 0, the encoder and the receiver synchronize for the receiver to know that at time 1, the encoder sends a signal based on the state of the first queue. The encoder and the receiver do not need to synchronize after time 0.

Under the preceding changes to the EGL policy pair, essentially the same argument from the proof of Theorem 7 can be used to show that under no memory-feedback, $\rho^*_{\log N, \infty}(\mathcal{C}) = 1$. One key point to note is that without the synchronization between the encoder and the receiver at time 0, the receiver may only obtain estimates of the arrival rates up to a cyclic permutation on the queue indices. Synchronization resolves this issue, and lets the receiver know which estimate correspond to which queue. Q.E.D.

Appendix C: Finding the Capacity Factor and Optimal Policies

We demonstrate in this section how to calculate the capacity factor $\rho^*_{\infty,v}(\mathcal{C})$ as the optimal value of a polynomial optimization problem over finite-dimensional matrices. Note that, by Theorem 1, this will cover all $\rho^*_{k,v}(\mathcal{C})$ for $k \geqslant K(\Pi, \mathcal{X})$, and serve as an upper bound on $\rho^*_{k,v}$ for $k < K(\Pi, \mathcal{X})$. Moreover, the optimal solutions to this optimization problem lead to the allocation policy (ψ^ϵ) and simple encoding policies (Φ^ϵ) that will be used by the Episodic Max-Weight policy (Lemma 2, Section 9.2) to achieve the capacity factor.

We first consider the case where $v \geqslant 1$ (i.e., the receiver is not memoryless). Fix $v \in \mathbb{N}$. The chain $\{(X(t), M_r(t), \boldsymbol{D}(t))\}_{t\in\mathbb{N}}$ is Markov with the following transition dynamics. Let a simple encoding policy (Definition 5) be parameterized by the $(|\mathcal{M}_r||\mathcal{X}|) \times |\mathcal{X}|$ row-stochastic matrix, G^E, with

$$G^E_{(m,x),x'} = \mathbb{P}(X(t+1) = x' \,|\, M_r(t) = m, X(t) = x). \tag{128}$$

Similarly, an allocation policy ψ can be parameterized by the pair (G^A, H^A), where G^A is an $(|\mathcal{M}_r||\mathcal{Y}|) \times |\mathcal{M}_r|$ matrix, with

$$G^A_{(m,y),m'} = \mathbb{P}(M_r(t+1) = m' \,|\, M_r(t) = m, Y(t) = y), \tag{129}$$

and H^A is an $(|\mathcal{M}_r||\mathcal{Y}|) \times |\Pi|$ matrix, with

$$H^A_{(m,y),\boldsymbol{d}} = \mathbb{P}(\boldsymbol{D}(t) = \boldsymbol{d} \,|\, M_r(t) = m, Y(t) = y). \tag{130}$$

Denote by G^S the transition matrix associated with the chain $\{(X(t), M_r(t), \boldsymbol{D}(t))\}_{t\in\mathbb{N}}$:

$$\begin{aligned} & G^S_{(m,x,\boldsymbol{d}),(m',x',\boldsymbol{d}')} \\ =& \mathbb{P}\big(X(t+1) = x', M_r(t+1) = m', \boldsymbol{D}(t+1) = \boldsymbol{d}' \\ & \quad \big| \, M_r(t) = m, X(t) = x, \boldsymbol{D}(t) = \boldsymbol{d}\big). \end{aligned} \tag{131}$$

We can write G^S as a function of G^E, G^A and G^E:

$$G^S_{(m,x,\boldsymbol{d}),(m',x',\boldsymbol{d}')} = G^E_{(m,x),x'} \left(\sum_{y\in\mathcal{Y}} C_{x',y} G^A_{(m,y),m'} H^A_{(m,y),\boldsymbol{d}} \right), \tag{132}$$

where C is the channel matrix, with $C_{x,y} = \mathbb{P}(Y(t) = y \,|\, X(t) = x)$. Note that G^E, G^E and H^A are row-stochastic matrices chosen by the system designer, while C is given. To ensure that the resulting G^S is irreducible, we may perturb the entries in C by a very small amount so that all entries are positive, and similarly, we may constrain the entries of the row-stochastic matrices to be bounded from below by a small constant. The irreducibility of G^S implies that it is associated with a unique stationary distribution, given by

$$\mathbf{p} = \left(I - G^S\right)^{-1}. \tag{133}$$

Denote by $\widehat{\mathbf{p}}$ the marginalized stationary distribution over the allocation vectors:

$$\widehat{\mathbf{p}}_{\boldsymbol{d}} = \sum_{m\in\mathcal{M}_r, x\in\mathcal{X}} \mathbf{p}^S_{(m,x,\boldsymbol{d})}, \quad \boldsymbol{d} \in \Pi. \tag{134}$$

The resulting stationary service rate is given by

$$\boldsymbol{\mu}(G^E, (G^A, H^A)) = \sum_{\boldsymbol{d}\in\Pi} \widehat{\mathbf{p}}_{\boldsymbol{d}} \cdot \boldsymbol{d}. \tag{135}$$

Using the above construction, we now formulate the optimization problem that will lead to the capacity factor. Recall from Lemma 2 that the capacity factor is given by the minimal shrinkage to the maximum capacity region such that it can be dominated by the set of service rates achievable through simple encoding policies. We can therefore compute $\rho^*_{\infty,v}(\mathcal{C})$ as follows. Let $\mathcal{E}'$ be the set of maximal schedules in $\mathcal{E}$, where $\mathcal{E}' = \{\boldsymbol{d}^{(i)}\}_{i=1,\ldots,|\mathcal{E}'|}$. Consider the following polynomial optimization problem:

$$\begin{aligned} \text{maximize} \quad & \rho \\ \text{subject to} \quad & \boldsymbol{\mu}\big(G^E(i), \big(G^A, H^A\big)\big) \geqslant \rho \boldsymbol{d}^{(i)}, \\ & i = 1, \ldots, |\mathcal{E}'|, \end{aligned} \tag{136}$$

where the variables to be optimized are the row-stochastic matrices G^A, H^A, and $\{G^E(i)\}_{i=1,\ldots,|\mathcal{E}'|}$. Denote by $\overline{\rho}$ and $\left(\{\overline{G}^E(i)\}_{i=1,\ldots,|\mathcal{E}'|}, (\overline{G}^A, \overline{H}^A)\right)$ the optimal value and an optimal solution of (136), respectively. We have that the optimal value corresponds to the capacity factor:

$$\rho^*_{\infty,v}(\mathcal{C}) = \overline{\rho}. \tag{137}$$

Furthermore, to construct the Episodic Max Weight policy that achieves the capacity factor, the allocation policy (ψ^ϵ in Lemma 2) and the set of simple encoding policies (Φ^ϵ in Lemma 2) are given by those associated with $(\overline{G}^A, \overline{H}^A)$ and $\{\overline{G}^E(i)\}_{i=1,\ldots,|\mathcal{E}'|}$, respectively.

Special case of memoryless receiver. When $v = 0$, the optimization problem in (136) can be further simplified. Recall the notation of rate allocation matrix Θ in Eq. (35), and schedule matrix S in (36). The capacity factor $\rho^*_{\infty,0}(\mathcal{C})$ is the optimal value of the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & \rho \\ \text{subject to} \quad & \mathbf{r}^{(i)} C \Theta S \geqslant \rho \boldsymbol{d}^{(i)}, \quad i = 1, \ldots, |\mathcal{E}'|, \\ & \mathbf{r}^{(i)} \geqslant 0, \sum_{x \in \mathcal{X}} \mathbf{r}^{(i)}_x = 1, \quad i = 1, \ldots, |\mathcal{E}'|, \end{aligned} \tag{138}$$

where the variables to be optimized are the probability vectors $\{\mathbf{r}^{(i)}\}_{i=1,\ldots,|\mathcal{E}'|}$ and row-stochastic matrix Θ. Here, $\mathbf{r}^{(i)}$ represents the probabilities over the set of input symbols, $\mathcal{X}$. The allocation policy that achieves the capacity factor corresponds to the matrix Θ^* in an optimal solution of (138).

Notably, enabled by our theoretical results, the above derivations show that $\rho^*_{\infty,v}(\mathcal{C})$ can be calculated by solving polynomial optimization problems over *finite-dimensional* matrices. In contrast, without Theorem 1, it would have not been clear *a priori* how to compute $\rho^*_{\infty,v}(\mathcal{C})$, since it is defined as the supremum over an unbounded family of encoding policies, who may take as input the entire queue lengths, as well as an encoding memory state ($M_e(t)$) with an unbounded size. In the same way, identifying the optimal encoding and allocation policies that achieve the capacity factor would have been quite difficult if one were to solve it via brute force. Admittedly, the optimization problems in (136) and (138) still have their drawbacks: they may be non-convex and could scale poorly as the size of the receiver memory, v, or that of the maximal schedules, $|\mathcal{E}'|$, becomes larges. It would be an interesting future direction to investigate how these optimization problems can be solved efficiently.

Appendix D: Glossary of Frequently Used Symbols

$\boldsymbol{A}(t)$	arrival vector in time slot t
$\mathcal{C}$, C	channel, channel matrix
$\boldsymbol{D}(t)$	chosen allocation / schedule in time slot t
k	encoder memory size (bits)
v	receiver memory size (bits)
$\boldsymbol{\Lambda}$	maximum capacity region
M_e, M_r	encoder / receiver memory
N	system size / number of queues
Π	set of allowable allocations / schedules
$\boldsymbol{Q}(t)$	queue state in time slot t
$\rho^*_{k,v}(\mathcal{C})$	(k,v)-capacity factor
Θ	rate allocation matrix for memoryless allocation policies
$X(t), \mathcal{X}$	input symbol in time slot t, input alphabet
$Y(t), \mathcal{Y}$	output symbol in time slot t, output alphabet