

Information, Memory and Capacity in Dynamic Resource Allocation

Kuang Xu

Stanford University, Graduate School of Business
kuangxu@stanford.edu

Yuan Zhong

University of Chicago, Booth School of Business
yuan.zhong@chicagobooth.edu

ABSTRACT

We propose a general framework, dubbed *Stochastic Processing under Imperfect Information (SPII)*, to study the impact of information constraints and memories on dynamic resource allocation. The framework involves a Stochastic Processing Network (SPN) scheduling problem in which the decision maker may access the system state only through a noisy *channel*, and resource allocation decisions must be carried out through the interaction between an encoding policy (who observes the state) and allocation policy (who chooses the allocation) situated on the two ends of the channel. Applications in the management of large-scale data centers and human-in-the-loop service systems are among our chief motivations.

We quantify the degree to which information constraints *reduce* the size of the *capacity region* in general SPNs, and how such reduction depends on the amount of *memories* available to the encoding and allocation policies. Using a novel metric, *capacity factor*, our main theorem characterizes the reduction in capacity region (under “optimal” policies) for *all non-degenerate channels*, and across *almost all combinations of memory sizes*. Notably, the theorem demonstrates, in substantial generality, that (1) the presence of a noisy channel *always reduce* capacity, (2) more memories for the allocation policy *always improve* capacity, and (3) more memories for the encoding policy have *little to no effect* on capacity. Finally, all of our positive (achievability) results are established through constructive, implementable policies.

Our proof program involves the development of a host of new techniques, largely from first principles, by combining ideas from information theory, learning and queueing theory. As a sub-module of one of the policies proposed, we create a simple yet powerful generalization of the Max-Weight policy, in which individual Markov chains are selected dynamically, in a manner analogous to how schedules are used in a conventional Max-Weight policy.¹

1 INTRODUCTION

In many modern large-scale resource allocation systems, such as data centers, call centers and hospitals, getting reliable access to accurate system state information usually requires expensive investment in monitoring or machine learning infrastructures. Furthermore, such information can often be subject to noise, loss or misinterpretation. It is therefore crucial to understand how *imperfect* and *noisy* information affects system performance. Insights along this direction can provide crucial architectural guidelines on how to design efficient information-driven scheduling policies,

and also assist with infrastructure planning by quantifying the performance benefits from better information access, and hence the tradeoffs between such benefits and the investment costs.

As a first step towards this direction, we propose in this paper a new, general framework for quantifying the performance impact of information constraints on an underlying dynamic resource allocation problem. Specifically, we will focus on Stochastic Processing Networks (SPN) ([8, 14–16]) – a widely used paradigm for modeling resource allocation problems in diverse sectors, including information technology ([9, 23, 27, 35]), manufacturing ([6]), call centers ([12, 17, 21]), as well as other service industries – and address *how information, or the lack of, alters the capacity region of an SPN*.

We begin with an overview of the model; the formal description will be presented in Section 2. The framework, dubbed the **Stochastic Processing under Imperfect Information (SPII)** model, is illustrated in Figure 1, and consists of three main elements: an underlying dynamic resource allocation problem, a model of imperfect information, and memories.

I. Stochastic Processing Network Scheduling. The underlying dynamic resource allocation problem is that of scheduling in a discrete-time SPN, sometimes also known as a *switched network* ([29, 32]), where a finite set of processing resources is employed to serve incoming tasks of N different *types*. In each time slot t , new tasks arrive to the system in a stochastic manner, where unprocessed tasks of type i are buffered in a queue i , and the queue lengths are denoted by $Q(t) = (Q_1(t), \dots, Q_N(t))$. The number of arrivals of type i jobs, $A_i(t)$, has an expected value $\mathbb{E}[A_i(t)] = \lambda_i$, and we will refer to the vector $\lambda = (\lambda_1, \dots, \lambda_N)$ as the *arrival rate vector*. The decision maker is to select from a finite *schedule set*, Π , an *allocation vector*, $D(t) = (D_1(t), \dots, D_N(t)) \in \Pi$, where $D_i(t)$ corresponds to the number of tasks in queue i that can be processed during the present time slot. The queue lengths evolve according to the following dynamics²:

$$Q(t) = (Q(t-1) - D(t))^+ + A(t), \quad t \in \mathbb{N}, \quad (1.1)$$

As an example, the scheduling problem involving one server and two parallel queues falls under this framework, as is illustrated in Figure 2, where the schedule set contains two elements, $\Pi = \{(1, 0), (0, 1)\}$, corresponding to the server processing a job from the first and second queue, respectively.

II. Scheduling with Imperfect Information. As a major departure from conventional SPNs, the decision maker in our model *does not* have access to the full queue-length state information when making scheduling decisions. In contrast, she obtains information concerning the queues only through a **(noisy) channel**. A channel consists of a pair of finite input and output alphabets, \mathcal{X} and \mathcal{Y} , and

¹Technical report. This version: January 2019. The authors would like to thank the review team at ACM SIGMETRICS for their detailed, constructive feedback.

²For a vector $\mathbf{x} = (x_1, \dots, x_N)$, we use the notation $(\mathbf{x})^+$ to mean the vector $(\max\{x_1, 0\}, \dots, \max\{x_N, 0\})$.

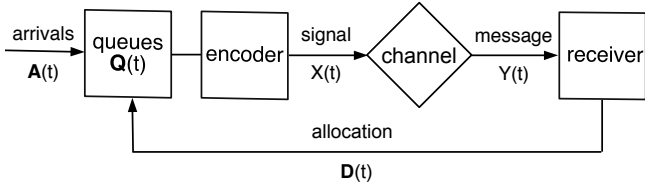


Figure 1: System diagram. The encoder is able to fully observe the state of the queues, $Q(t)$. The encoder and receiver are equipped with memories of size k and ℓ bits, respectively.

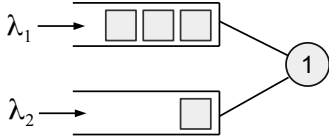


Figure 2: An example of a stochastic processing network with one server and two parallel queues. Two types of tasks arrive to the systems at rates λ_1 and λ_2 respectively, and wait in two separate queues. The schedule set consist of two possible allocation vectors, $(1, 0)$ and $(0, 1)$, representing the server's choice of processing one job from queue 1 or queue 2, respectively.

a family of probability distributions $\{\mathbb{P}_x^C\}_{x \in \mathcal{X}}$ over \mathcal{Y} . When an input *signal*, $x \in \mathcal{X}$, is sent through the channel, it results in a random output *message*, $Y \in \mathcal{Y}$, drawn from the probability distribution $\mathbb{P}_x^C(Y = \cdot)$. For instance, one simple channel is that of an ε -noisy binary symmetric channel with $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, where the input signal is correctly received with probability $1 - \varepsilon$, and is perturbed to the opposite symbol with probability ε , i.e., $\mathbb{P}_a^C(Y = a) = 1 - \varepsilon$ and $\mathbb{P}_a^C(Y = b) = \varepsilon$ for $a, b \in \{0, 1\}$, $a \neq b$. By using different alphabets and distributions \mathbb{P}^C , the channel is able to capture a variety of partial and/or lossy information models, such as controlling a data center over band-limited communication constraints, where only noisy and compressed signals of the full system state can be obtained. In this paper, we will impose little restriction on the form of the channel, in order to allow for a maximum degree of generality and modeling flexibility.

In the presence of a channel, the allocation decisions in our model are carried out by a pair of **encoding** and **allocation policies**, situated on opposite ends of the channel (Figure 1). An *encoder* is co-located with the queues and has complete knowledge of the queue lengths at all times. In time slot t , the encoder employs an *encoding policy*, ϕ , to send an *input symbol* $X(t) \in \mathcal{X}$ through the channel. The resulting random *output message*, $Y(t)$, arrives at the *receiver*, who then employs an *allocation policy*, ψ , to choose the allocation vector $D(t)$. Note that the receiver does not observe the queue lengths, and in most settings considered in this paper, the information contained in $Y(t)$ is severely limited; for instance, the output alphabet \mathcal{Y} can be substantially smaller than the number of possible allocation vectors in Π .

III. Memory. The last crucial element of the system is memory. In a classical SPN, external memories are rarely needed by a scheduling policy, because the queue lengths are fully observable and already contain all the relevant information for policies such as

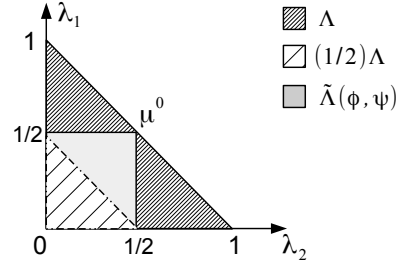


Figure 3: Illustration of the reduction of capacity region due to lossy information, for the two-queue-one-server SPN in Figure 2. The triangle with vertices $(0, 0)$, $(0, 1)$ and $(1, 0)$ represents the maximum capacity region, Λ , when the queue lengths are fully observable, Λ . The rectangle dominated by μ^0 represents the capacity region $\tilde{\Lambda}(\phi, \psi)$ under an uninformative channel and an allocation policy that induces an average service rate of μ^0 . The triangle with vertices $(0, 0)$, $(0, 1/2)$ and $(1/2, 0)$ represent the scaled (“shrunk”) version of the maximal capacity region that is contained in $\tilde{\Lambda}(\phi, \psi)$.

Max-Weight (see e.g., [8, 29, 32, 35]) to make scheduling decisions. However, when the system state information becomes constrained and imperfect, it is conceivable that memory can play a crucial role in determining the system performance. For instance, with memory, the allocation policy could aggregate past messages across multiple time slots to better estimate the system state and inform its scheduling decisions, and similarly, the encoding policy could benefit by remembering past transmissions in order to better tailor future input signals.

For this reason, we will allow for the possibility that the encoder and receiver have at their disposal a finite *memory* of k and ℓ bits, respectively, in which information can be recorded and subsequently retrieved in the next time slot. The encoding policy may generate the input signal $X(t)$ based on the queue lengths as well as the state of the encoder memory, and similarly, the allocation policy may choose the allocation vector $D(t)$ based on the received output message $Y(t)$ along with the state of the receiver memory. If k or ℓ is equal to 0, we say that the corresponding encoding or allocation policy is *memoryless*.

Finally, we say that the system has **memory-feedback** if the state of the receiver memory is accessible by the encoding policy. The existence of memory-feedback would correspond to applications in which the *backward* communication from the receiver to the encoder is lossless and relatively cheap. Arguably, this is not a very restrictive assumption within the SPII framework, considering that the receiver must, in any event, communicate the allocation decision, $D(t)$, back to the queues, and hence sending along the state of the receiver memory in the mean time should not incur too much additional overhead. A number of main results in the present paper will rely on this assumption, and we expect to relax it in future work.

1.1 Performance Metric: Capacity Factor

The main quantity of interest is the *capacity region* of an SPII: the set of arrival rate vectors λ under which the queue lengths, $Q(\cdot)$,

remain stable³, for a given pair of encoding and allocation policies. As we shall see shortly, harsh information constraints tend to *reduce* the capacity region. Therefore, at a high level, the main question that we aim to address can be stated as follows:

How does the information constraint, in the form of a noisy channel, reduce the capacity region of an SPII architecture, and how does the degree of reduction vary with respect to the sizes of encoder and receiver memories?

In order to *measure* the magnitude of the reduction in capacity region, we introduce a key (scalar) metric, dubbed the **capacity factor**, which captures the *fraction* of capacity region lost due to imperfect information. Informally, the notion of capacity factor can be described as follows (the formal definition will be given in Section 2). Fix a channel $C = (\mathcal{X}, \mathcal{Y}, \{\mathbb{P}_X^C\})$. For a pair of encoding-allocation policies (ϕ, ψ) , denote by $\tilde{\Lambda}(\phi, \psi)$ the set of all arrival rate vectors λ that the system is able to stabilize under the pair (ϕ, ψ) and channel C . Denote by Λ the *maximum capacity region* of the SPII, defined by the set of arrival rate vectors that are dominated by the convex hull of the schedule set, Π . In particular, Λ corresponds to the maximal capacity region of the system under *full information* and a maximally stable scheduling policy, such as Max-Weight.

Definition 1.1 (Capacity Factor (Informal)). Fix $k, \ell \in \mathbb{Z}_+$. The (k, ℓ) -*capacity factor* of the channel C , denoted by $\rho_{k, \ell}^*(C)$, is defined by the supremum of $\rho \in (0, 1)$, such that

$$\rho\Lambda \subset \tilde{\Lambda}(\phi, \psi), \quad (1.2)$$

for *some* pair of encoding and allocation policies (ϕ, ψ) , whose memory sizes are k and ℓ bits, respectively.

Intuitively, the capacity factor can be interpreted as the “largest” fraction of the full-information maximum capacity region that can be *preserved* under information constraints (using an “optimal” encoding-allocation policy pair). A capacity factor of 1 means that the information constraints result in *no loss* of the maximum capacity region, and on the other extreme, a capacity factor of 0 indicates that *most* of the maximum capacity region has been lost due to lossy information.

Example 1.2 (Scheduling with Parallel Queues). We now present a simple example to illustrate the drastic impact that lossy information can have on the capacity region, consequently resulting in a small capacity factor. Consider the SPN in Figure 2 with one server and two parallel queues. It is not difficult to see that the maximum capacity region under full information, Λ , consists of all arrival rates where $\lambda_1 + \lambda_2 < 1$, achieved by always serving a queue that is non-empty.

Now suppose the channel of the SPII is completely uninformative, so that the output message $Y(t)$ is independent of the input signal $X(t)$. Then, it can be shown that any allocation policy will induce a *fixed* average service rate vector, μ^0 , independent of the actual arrival rates. The resulting capacity region is represented by the rectangle in Figure 3 (dominated by μ^0), which is *always* substantially smaller than the maximum capacity region, regardless of the choice of μ^0 . It is not difficult to show that the capacity factor

³Roughly speaking, the process $Q(\cdot)$ is stable if queue lengths do not grow to infinity as time goes. The precise notion of stability that we consider in this paper will be defined in Section 2.1.

$\rho_{k, \ell}^*(C)$	$k = 0$	$k \in \mathbb{N}$	$k = \infty$
$\ell = 0$	$\rho_{0,0}^* = \rho_{\infty,0}^*$	$\Rightarrow \rho_{k,0}^* = \rho_{\infty,0}^*$	$\rho_{\infty,0}^* < 1^\dagger$
$\ell \in \mathbb{N}$	–	$\rho_{K(\Pi, \mathcal{X}), \ell}^* = \rho_{\infty, \ell}^*$	$\rho_{\infty, \ell}^* < 1^\dagger$
$\ell = \infty$	1	$\Rightarrow 1$	$\Rightarrow 1$

Table 1: Summary of our main results in Theorem 3.3, where we omit the dependence on C for simplicity of notation. The system is assumed to have memory-feedback. The parameters k and ℓ denote the size of the memory for the encoder and receiver, respectively. The symbol “ \Rightarrow ” indicates that the result of the cell is readily implied by that of the cell to the left. The symbol “ \dagger ” means that the result holds if in addition to being informative, the channel is also “noisy” (more precisely, ε -majorizing for some $\varepsilon > 0$; see Definition 3.2). Finally, the symbol “–” means that we do not yet know the exact value of $\rho_{k, \ell}^*(C)$ when $k < K(\Pi, \mathcal{X})$, beyond the fact that it’s not greater than $\rho_{K(\Pi, \mathcal{X}), \ell}^*(C)$. For instance, we do not know whether $\rho_{k, \ell}^*(C)$ is in fact equal to $\rho_{K(\Pi, \mathcal{X}), \ell}^*(C)$ for all $k < \rho_{k, \ell}^*(C)$.

under such an uninformative channel is equal to 1/2, which corresponds to the case where the allocation policy induces a service rate vector of $\mu^0 = (1/2, 1/2)$ by choosing the two schedules with equal frequency. Therefore, depending on the quality of the channel and the memories available, the fractional loss of capacity region in SPII can range from 1/2 (uninformative) to 1 (full-information).

1.2 Preview of Main Results: Characterization of Capacity Factor

We now give an informal preview of our main results (Theorem 3.3), which are also summarized in Table 1. Fix any channel that is *informative*, such that the input signal $X(t)$ and output message $Y(t)$ are *not* independent. Suppose that the system has memory-feedback. We show the following:

- (1) When the receiver is memoryless ($\ell = 0$), the capacity factor $\rho_{k,0}^*(C)$ is independent of k , the number of bits in the encoder memory:

$$\rho_{k,0}^*(C) = \rho_{0,0}^*(C), \quad \forall k \geq 0. \quad (1.3)$$

- (2) When the receiver has a finite memory ($0 < \ell < \infty$), the capacity factor $\rho_{k, \ell}^*(C)$ may depend on k , but only up to a constant, $K(\Pi, \mathcal{X})$, whose value does not depend on ℓ or the channel C :

$$\rho_{k, \ell}^*(C) = \rho_{K(\Pi, \mathcal{X}), \ell}^*(C), \quad \forall k \geq K(\Pi, \mathcal{X}). \quad (1.4)$$

- (3) In the limit where the size of the receiver memory tends to infinity ($\ell = \infty$), the (limiting) capacity factor $\rho_{k, \infty}^*(C)$ is always 1, regardless of the magnitude of k :

$$\rho_{k, \infty}^*(C) = 1, \quad \forall k \geq 0. \quad (1.5)$$

In particular, there will be *no loss* of capacity region in this limiting regime.

- (4) Finally, if the channel is “noisy” (to be precisely defined), and the receiver does not have infinite memory ($\ell < \infty$), then the

capacity factor $\rho_{k,\ell}^*(C)$ is *always* strictly less than 1:

$$\rho_{k,\ell}^*(C) < 1, \quad \forall \ell < \infty. \quad (1.6)$$

Moreover, as a by-product of the proof, all of our achievability results (Items 1 through 3) are established through constructive and implementable encoding-allocation policy pairs. The capacity factor can also be explicitly computed in closed-form for restricted families of channels and schedule sets, which we will demonstrate in Section 7.1. Appendix D discusses how to compute the capacity factor and identify optimal policies that achieve the capacity factor for the general case.

Key implications of the main results: Our results have a number of architectural implications; we highlight some of them below.

(1) *Impact of imperfect information is non-trivial.* The result in Item 4 show that the presence of a noisy channel *always* reduces capacity (i.e., $\rho_{k,\ell}^*(C) < 1$), regardless of the amount of memories available to the encoding or allocation policies, so long as the allocation policy does not have infinite memory.

(2) *Memories play an intricate, asymmetric role.* The results in Items 1 through 3 demonstrate that memories are crucial in mitigating the reduction in capacity region due to imperfect information. However, memories are *substantially more important* for the allocation policy than they are for the encoding policy: increase in memory for the allocation policy *always* improves capacity, while the benefit of having additional memory for the encoding policy becomes *zero* beyond some finite threshold. This asymmetry suggests design principles that favor simple, low-memory communication modules for the encoding policy, while the allocation policy must aggregate sufficient amount of past observations in order to achieve a maximum capacity region.

1.2.1 Motivating Examples of SPII. While the SPII model is highly stylized, it is motivated by a range of dynamic resource allocation systems with information constraints. We discuss in this sub-section two motivating examples:

(1) *Example 1 - Scheduling in Large-Scale Data Centers.* It has long been recognized that obtaining reliable system-wide state information is challenging in modern-day data centers, where many tasks are processed in a massively parallel manner, servers may fail, and bandwidths may be limited (see e.g., [1, 2, 7, 41]). For a concrete illustration, consider a manager who is operating a large-scale data center via a limited communication channel. Here, the channel is broadly construed as incorporating all aspects of information obfuscations, such as those due to server failures, and limiting bandwidths on the network and server I/O. For example, due to possible server failures or malfunctions, incomplete information concerning the state of the data center may be transmitted to the manager across a network link. Furthermore, because of communication constraints on the link and server I/O, the information that the manager obtains may be condensed and corrupted by randomly dropped packets. In these scenarios, our model will speak to the design of the communication protocol from the data center to the manager, as well as how she should translate the received messages into resource allocation decisions.

(2) *Example 2 - Human-in-the-Loop Resource Allocation.* Our model can also be applied to resource allocation problems where communications amongst human operators may be subject to errors or misinterpretations (cf. [5, 22, 26]). For instance, one may consider a hospital setting where a physician (“encoder”), who observes the state of the ward, must communicate the relevant information and instructions to nurses (“decoder”), who will carry out the actual interventions.

1.3 Organization

The remainder of the paper is organized as follows. We formally describe our model in Section 2, and present our main results in Section 3. Section 4 discusses the related literature. The remainder of the paper is devoted to the proof of our main results, with a proof overview in Section 5 that summarizes the key techniques. We conclude the paper with some discussion in Section 10.

1.4 Notation

We reserve boldface letters for vectors and plain letters for scalars. For any scalar x , $\lceil x \rceil$ denotes the smallest integer greater than or equal to x , and $(x)^+ \triangleq \max\{x, 0\}$ denotes the nonnegative part of x . For any $d \in \mathbb{N}$, and vector $\mathbf{x} \in \mathbb{R}^d$, we use x_i to denote the i th coordinate of \mathbf{x} . We use $(\mathbf{x})^+$ to denote the vector $((x_1)^+, \dots, (x_d)^+)$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^d$, we write $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for all $i = 1, 2, \dots, d$, and we write $\mathbf{x} < \mathbf{y}$ if $x_i < y_i$ for all $i = 1, 2, \dots, d$. For a set $\mathcal{X} \subset \mathbb{R}_+^d$, we write $\mathbf{x} \leq \mathcal{X}$ ($\mathbf{x} < \mathcal{X}$, respectively) if there exists $\mathbf{y} \in \mathcal{X}$ such that $\mathbf{x} \leq \mathbf{y}$ ($\mathbf{x} < \mathbf{y}$, respectively), and say that the vector \mathbf{x} is *dominated* (*strictly dominated*, respectively) by the set \mathcal{X} . For $N \in \mathbb{N}$, we will use the short-hand $[N]$ to denote the set $\{1, 2, \dots, N\}$ of consecutive integers.

For a set \mathcal{S} and function f , we will use $f(\mathcal{S})$ to denote the set $\{f(s) : s \in \mathcal{S}\}$. For any set $\mathcal{S} \in \mathbb{R}_+^d$, we use $\text{conv}(\mathcal{S})$ to denote the convex hull of \mathcal{S} , $\text{conv}^-(\mathcal{S}) \triangleq \{\mathbf{x} \in \mathbb{R}_+^d : \mathbf{x} < \text{conv}(\mathcal{S})\}$ to denote the set of vectors $\mathbf{x} \in \mathbb{R}_+^d$ *strictly* dominated by $\text{conv}(\mathcal{S})$, and $\text{cl}(\mathcal{S})$ to denote the closure of the set \mathcal{S} .

For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, the inner product of \mathbf{x} and \mathbf{y} is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle \triangleq \sum_{i=1}^d x_i y_i$. The ℓ_2 -norm of vector \mathbf{x} is denoted by $\|\mathbf{x}\| \triangleq \sqrt{\sum_{i=1}^d x_i^2}$, and the ℓ_∞ -norm denoted by $\|\mathbf{x}\|_\infty \triangleq \max\{|x_1|, \dots, |x_d|\}$.

The i th standard unit vector in \mathbb{R}^d is denoted by $\mathbf{e}^{(i)}$, whose i th coordinate equals 1 and all other coordinates equal zero. The vector with all components being 1 is denoted by $\mathbf{1}$.

We use the shorthand “w.p.1” to mean “with probability 1,” “i.i.d” to mean “identically and independently distributed,” and “WLOG” to mean “without loss of generality.” We use $\text{Uniform}(0, 1)$ to describe random variables that are uniformly distributed over the interval $(0, 1)$. The indicator function of an event A is denoted by \mathbb{I}_A .

2 THE MODEL

We formally present the model, Stochastic Processing with Imperfect Information (SPII), in this section.

Stochastic Processing Network model. We consider a dynamic Stochastic Processing Network (SPN) evolving in discrete time $t \in \mathbb{Z}_+$ (Figure 1). The system consists of N queues, whose lengths at the end of the t^{th} time slot are represented by $\mathbf{Q}(t) \in \mathbb{Z}_+^N$.

The evolution of the queues is captured by the following equation:

$$\mathbf{Q}(t) = (\mathbf{Q}(t-1) - \mathbf{D}(t))^+ + \mathbf{A}(t), \quad t \in \mathbb{N}, \quad (2.1)$$

where $\mathbf{A}(t)$ and $\mathbf{D}(t)$ are the arrival and allocation vectors during slot t , respectively. We use $\Pi \subset \mathbb{Z}_+^N$ to denote the finite set of all allowable schedules. The arrivals are i.i.d. Bernoulli random variables, where

$$\mathbb{E}(A_i(1)) = \lambda_i, \quad i = 1, 2, \dots, N. \quad (2.2)$$

We will refer to $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_N)$ as the *arrival rate vector*. WLOG we suppose that the schedule set Π satisfies the following three assumptions.

Assumption 1. Π is *monotone*: if $\mathbf{d} \in \Pi$, then for any $\mathbf{d}' \in \mathbb{Z}_+^N$ with $\mathbf{d}' \leq \mathbf{d}$, $\mathbf{d}' \in \Pi$ as well.

Assumption 2. For each $i \in [N]$, $\mathbf{e}^{(i)} \in \Pi$.

Assumption 3. The schedule set Π has at least two distinct maximal elements. More specifically, there exist $\mathbf{d}^{(1)}, \mathbf{d}^{(2)} \in \Pi$ with $\mathbf{d}^{(1)} \neq \mathbf{d}^{(2)}$, and for $j = 1, 2$, $\mathbf{d}^{(j)} \leq \mathbf{d}$ with $\mathbf{d} \in \Pi$ implies that $\mathbf{d} = \mathbf{d}^{(j)}$.

Assumption 1 allows for more flexible allocation decisions without impacting system performance. Assumption 2 guarantees that each queue can receive a positive service rate, and Assumption 3 rules out the possibility of Π having only one maximal element, in which case the single maximal element would dominate all other schedules, and the trivial decision of always choosing that maximal element would be optimal for a wide range of performance objectives, such as maximizing throughput or minimizing queue lengths.

Signals, channels, and messages. A *channel* is a triple $C = (\mathcal{X}, \mathcal{Y}, C)$, where \mathcal{X} and \mathcal{Y} are finite sets representing the input and output alphabets, with cardinalities $c_{\mathcal{X}}$ and $c_{\mathcal{Y}}$, respectively, and C is an $c_{\mathcal{X}}$ -by- $c_{\mathcal{Y}}$ row stochastic matrix, which we will refer to as the *channel matrix*. Since C is row stochastic, each row corresponds to a probability distribution over \mathcal{Y} , the set of output alphabets, and we denote the probability distribution corresponding to the x th row as \mathbb{P}_x^C . When an input *signal*, $X \in \mathcal{X}$ is sent through the channel, it leads to a (possibly random) output *message*, $Y \in \mathcal{Y}$, drawn from the probability distribution $\mathbb{P}_x^C(Y = \cdot)$. Thus, the matrix C captures the stochastic distortion introduced by the channel, where the entry $C_{x,y}$ represents the probability that the output message of the channel is y when the input signal is x :

$$C_{x,y} = \mathbb{P}_x^C(Y = y), \quad x \in \mathcal{X}, y \in \mathcal{Y}. \quad (2.3)$$

We assume the channel is (1) memoryless, so that each output message only depends on the input signal of the present time slot, and is independent from the system's past history, and (2) stationary, so that for any $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, the probability that the output is y when the input is x does not depend on time.

Encoder and encoding policies. During each time slot, an *encoder* situated at the queues sends a signal, $X(t) \in \mathcal{X}$, over the channel. The encoder is equipped with a finite-sized lossless *memory* represented by a k -bit binary sequence, whose value at time t is denoted by $M_e(t)$. The signal $X(t)$ can depend on the most recent state of the queues $\mathbf{Q}(t-1)$, the most recent arrivals $\mathbf{A}(t-1)$, the content of the memory $M_e(t-1)$, and possibly some idiosyncratic

randomness. Formally, let ϕ_e be a deterministic *encoding policy*. Then,⁴

$$X(t) = \phi_e(\mathbf{Q}(t-1), \mathbf{A}(t-1), M_e(t-1), Z_e(t-1)), \quad t \in \mathbb{N}, \quad (2.4)$$

where $\{Z_e(t)\}_{t \in \mathbb{Z}_+}$ is a string of i.i.d. Uniform(0, 1) random variables, which are all independent from the rest of the system. In each time slot t , the content of the memory, $M_e(t)$, is also updated based on $\mathbf{Q}(t-1)$, $\mathbf{A}(t-1)$, $M_e(t-1)$, and $Z_e(t-1)$, and we can formally write

$$M_e(t) = \phi_m(\mathbf{Q}(t-1), \mathbf{A}(t-1), M_e(t-1), Z_e(t-1)), \quad t \in \mathbb{N}, \quad (2.5)$$

for some deterministic function ϕ_m . If we write $\phi = (\phi_e, \phi_m)$, then $(X(t), M_e(t)) = \phi(\mathbf{Q}(t-1), \mathbf{A}(t-1), M_e(t-1), Z_e(t-1))$, $t \in \mathbb{N}$.^(2.6)

With a slight abuse of notation, we also call ϕ the *encoding policy*. We will denote by Φ_k the set of all encoding policies with k bits of memory. We use $\mathcal{M}_e(k)$ to denote the set of possible values for the encoder memory when k bits are allowed. When the context is clear, we often suppress the dependence on k and simply write \mathcal{M}_e .

Roughly speaking, the size of the memory, k , serves as a measure of “complexity” of an encoding function. A special case is when $k = 0$, where the encoder is equipped with no memory and the signal depends only on the current state of the queues. We will refer to a policy $\phi \in \Phi_0$ as a *memoryless encoding policy*.

Receiver and allocation policies. The signal $X(t)$ passes through the channel and results in a message, $Y(t) \in \mathcal{Y}$ at a *receiver*. The responsibility of the receiver is to choose, at each time slot, the allocation vector $\mathbf{D}(t)$. However, the receiver is not able to observe the state of the queues directly, so the allocation decisions can only rely on information provided by the encoder through the channel. Similar to the encoder, the receiver is equipped with a memory of ℓ bits, whose state in slot t is denoted by $M_r(t)$. Let ψ_a be a deterministic *allocation policy*, such that

$$\mathbf{D}(t) = \psi_a(Y(t), M_r(t-1), Z_r(t-1)), \quad t \in \mathbb{N}, \quad (2.7)$$

where $\{Z_r(t)\}_{t \in \mathbb{Z}_+}$ is a string of i.i.d. Uniform(0, 1) random variables, which are all independent from the rest of the system. In each time slot t , the content of the receiver memory, $M_r(t)$, is also updated. However, different from the allocation decisions, $M_r(t)$ is updated with a *time lag*, and it depends on $M_r(t-1)$, $Z_r(t-1)$ and $Y(t-1)$, the message from an earlier time slot, instead of $Y(t)$, the most recent message.⁵ Formally,

$$M_r(t) = \psi_m(Y(t-1), M_r(t-1), Z_r(t-1)), \quad t \in \mathbb{N}, \quad (2.8)$$

for some deterministic function ψ_m . Similar to the encoder side, we also call $\psi \triangleq (\psi_a, \psi_m)$ the *allocation policy*.

We will denote by Ψ_ℓ the set of all allocation policies with ℓ bits of memory. Analogous to the encoding policies, an allocation policy with no memory generates the allocation decision using

⁴ In the sequel, we will see that our achievability results for the case of finite receiver memory (Items 1 and 2 of Theorem 3.3; also Sections 8 and 9) are established using encoding policies that do not depend on $\mathbf{A}(t-1)$, the most recent arrivals. However, to prove the result on infinite receiver memory (Item 3 of Theorem 3.3; also Appendix C), the encoding policy that we constructed makes crucial use of $\mathbf{A}(t-1)$, which is therefore included in Eq. (2.6).

⁵ The one-step lag can be removed without substantially changing the results; it serves the purpose of simplifying the notation and proof.

only the current message, $Y(t)$. We will refer to a policy $\psi \in \Psi_0$ as a *memoryless allocation policy*. The set of possible values for the receiver memory is denoted by $\mathcal{M}_r(\ell)$ when ℓ bits are allowed. Similar to \mathcal{M}_e , when the context is clear, we suppress dependence of $\mathcal{M}_r(\ell)$ on ℓ and simply write \mathcal{M}_r .

In this paper, we are primarily interested in the dynamics of the tuple

$$\mathbf{W}(t) \triangleq (\mathbf{Q}(t), \mathbf{A}(t), M_e(t), X(t), M_r(t)), \quad t \in \mathbb{Z}_+. \quad (2.9)$$

It is not difficult to verify that under any well-defined encoding-allocation policy pair (ϕ, ψ) , $\mathbf{W}(\cdot)$ is a countable-state Markov chain. Therefore, for any time t , we call $\mathbf{W}(t)$ *the system state at time t* , and from now on, we restrict our attention to pairs of encoding and allocation policies (ϕ, ψ) under which the Markov chain $\mathbf{W}(\cdot)$ is irreducible.

Memory-feedback. We say that the system has **memory-feedback** if the state of the receiver memory, $M_r(t-1)$, is accessible by the encoding policy in time slot t for generating the message $X(t)$ and updating the encoder memory state $M_e(t)$. That is, under this assumption, Eq. (2.6) would become:

$$\begin{aligned} & (X(t), M_e(t)) \\ & = \phi(\mathbf{Q}(t-1), \mathbf{A}(t-1), M_e(t-1), M_r(t-1), Z_e(t-1)), \end{aligned} \quad (2.10)$$

for $t \in \mathbb{N}$. As alluded to in the Introduction, note that the model already assumes a mode of feedback: the scheduling decision, $D(t)$, can be sent to the queues without obstruction, implying that the backward communication from the receiver to the encoder is lossless. Therefore, under the memory-feedback assumption, in addition to sending $D(t)$, the allocation policy also includes the state of its own memory in the backward communication.

2.1 Main Performance Metric: Capacity Factor

We define in this subsection the main performance metric of this paper, the capacity factor. Fix a pair of encoding and allocation policies, ϕ and ψ , respectively. We say the system is *stable* if $\mathbf{W}(\cdot)$ is positive recurrent. Define the *maximum capacity region*, Λ , to be the set of all vectors strictly dominated by the convex hull of the schedule set Π :

$$\Lambda \triangleq \text{conv}^-(\Pi) = \{\boldsymbol{\lambda} \in \mathbb{R}_+^N : \boldsymbol{\lambda} < \text{conv}(\Pi)\}. \quad (2.11)$$

Note that because the schedule set Π satisfies Assumptions 1 and 2, it is not difficult to see that $\text{cl}(\boldsymbol{\lambda}) = \text{conv}(\Pi)$.

We now define our main performance metric.

Definition 2.1 (Capacity Factor). Fix a channel C , and $k, \ell \in \mathbb{Z}_+$.

- (a) Consider encoding policy $\phi \in \Phi_k$ and allocation policy $\psi \in \Psi_\ell$. Define $\tilde{\Lambda}(\phi, \psi)$ to be the *capacity region* under the policy pair (ϕ, ψ) :

$$\tilde{\Lambda}(\phi, \psi) \triangleq \{\boldsymbol{\lambda} \in \mathbb{R}_+^N : (\phi, \psi) \text{ stabilizes the system under the arrival rate vector } \boldsymbol{\lambda}\}. \quad (2.12)$$

We also define the *capacity factor of the channel C under the policy pair (ϕ, ψ)* , denoted by $\rho^*(\phi, \psi, C)$, to be

$$\rho^*(\phi, \psi, C) = \sup\{\rho > 0 : \rho\Lambda \subseteq \tilde{\Lambda}(\phi, \psi)\}. \quad (2.13)$$

- (b) The (k, ℓ) -*capacity factor of channel C* , denoted by $\rho_{k, \ell}^*(C)$, is defined to be

$$\rho_{k, \ell}^*(C) = \sup\{\rho^*(\phi, \psi, C) : \phi \in \Phi_k, \psi \in \Psi_\ell\}. \quad (2.14)$$

When the context is clear, sometimes we just write the ‘‘capacity factor’’ to mean $\rho_{k, \ell}^*(C)$.

Some elementary properties of capacity factor. Before we proceed, we state some elementary properties of capacity factor. First, for any channel, C , $\rho_{k, \ell}^*(C)$ must be non-decreasing in both k and ℓ , because the capacity region can never decrease with more memory. Furthermore, since $\rho_{k, \ell}^*(C)$ is upper-bounded by 1 by definition, by the Monotone Convergence Theorem, we have the following: $\rho_{k, \infty}^*(C) \triangleq \lim_{\ell \rightarrow \infty} \rho_{k, \ell}^*(C)$, and $\rho_{k, \infty}^*(C) \geq \rho_{k, \ell}^*(C)$ for all $\ell \in \mathbb{Z}_+$; $\rho_{\infty, \ell}^*(C) \triangleq \lim_{k \rightarrow \infty} \rho_{k, \ell}^*(C)$, and $\rho_{\infty, \ell}^*(C) \geq \rho_{k, \ell}^*(C)$ for all $k \in \mathbb{Z}_+$; $\rho_{\infty, \infty}^*(C) \triangleq \lim_{k \rightarrow \infty} \lim_{\ell \rightarrow \infty} \rho_{k, \ell}^*(C) = \lim_{\ell \rightarrow \infty} \lim_{k \rightarrow \infty} \rho_{k, \ell}^*(C)$. In particular, the limits in which we take k , ℓ , or both to ∞ are well defined.

3 MAIN RESULTS

We formally state the main results in this section. We begin with two definitions.

Definition 3.1 (Informative Channels). A channel C is said to be *informative* if the corresponding channel matrix, C , admits at least two distinct rows. A channel C whose channel matrix has identical rows is called *uninformative*.

The purpose of Definition 3.1 is to rule out degenerate channels: simply put, a channel is informative if and only if its output is not independent of the input. The next definition speaks to the other extreme by describing channels that are sufficiently noisy.

Definition 3.2 (ε -Majorizing Channels). Fix $\varepsilon \in (0, 1)$. We say that a channel C is ε -*majorizing* if its corresponding channel matrix, C , can be written as

$$C = \varepsilon C^0 + (1 - \varepsilon) C^1, \quad (3.1)$$

where C^0 and C^1 are two row-stochastic matrices, such that (a) the rows of C^0 are identical, and (b) every column of C^1 has at least one zero entry.

Roughly speaking, an ε -majorizing channel can be interpreted as having at most ε -portion of the channel being ‘‘completely uninformative.’’ For technical reasons, we will also assume that the ‘‘uninformative portion,’’ C^0 , of an ε -majorizing channel, is everywhere positive.⁶

Assumption 4. Let $C = (X, \mathcal{Y}, C)$ be an ε -majorizing channel, and let C^0 be as in (3.1). Then, for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, $C_{x, y}^0 > 0$.

The following theorem is the main result of this paper. The same results are summarized in Table 1, where the rows of the table correspond to Items 1 through 3 in the theorem, respectively. We will assume that the SPII architecture *has memory-feedback* (See Appendix C.3 for a discussion of a scenario without memory-feedback.)

⁶Intuitively, since C^0 is completely uninformative, whether Assumption 4 is satisfied or not should have little impact on performance. Assumption 4 is used in some of the subsequent proofs to ensure that states of the chain $\mathbf{W}(\cdot)$ are ‘‘easily reachable’’ from each other. For more details, see Appendix B.3.

THEOREM 3.3 (CHARACTERIZATION OF CAPACITY FACTOR). Fix the number of queues, $N \in \mathbb{N}$ and a finite schedule set, Π . Let C be an informative channel (Definition 3.1). Suppose the system has memory-feedback. The capacity factor, $\rho_{k,\ell}^*(C)$, satisfies the following:⁷

(1) Memory-less receiver: when $\ell = 0$, we have that

$$\rho_{k,0}^*(C) = \rho_{0,0}^*(C), \quad \forall k \geq 0. \quad (3.2)$$

(2) Finite-memory receiver: when $\ell \in \mathbb{N}$, we have that

$$\rho_{k,\ell}^*(C) = \rho_{K(\Pi,\mathcal{X})}^*(C), \quad \forall k \geq K(\Pi,\mathcal{X}). \quad (3.3)$$

where $K(\Pi,\mathcal{X}) \in \mathbb{N}$ depends only on the structure of the schedule set, Π , and the channel's input alphabet, \mathcal{X} .

(3) Infinite-memory receiver: when $\ell = \infty$, we have that

$$\rho_{k,\infty}^*(C) = \rho_{0,\infty}^*(C) = 1, \quad \forall k \geq 0. \quad (3.4)$$

(4) Suppose, in addition, that the channel is also ε -majorizing for some $\varepsilon > 0$ (Definition 3.2). Then, for all $\ell < \infty$, we have

$$\rho_{k,\ell}^*(C) < 1, \quad \forall k \geq 0. \quad (3.5)$$

Appendix D discusses how to compute the capacity factor and identify optimal policies that achieve the capacity factor.

4 RELATED LITERATURE

The challenges in obtaining reliable and timely access to state information have long been recognized in large-scale dynamic resource allocation problems. One prominent example is the celebrated ‘‘power-of-two-choices’’ (PoT) routing algorithm ([24, 38]) for load-balancing. Designed to address the lack of full queue-length information in a system with a large number of parallel queues, the PoT algorithm routes an incoming job to the shorter one between two randomly sampled queues. The same design consideration underlies pull-based variants of PoT ([3, 20, 33, 34]), and the partially centralized scheduling policy by [37] that has access to queue-length information only a small fraction of the time. Beyond the realm of computer networks, information constraints are also prominent in systems with humans in the loop. For instance, communication failures and misunderstanding between physicians and nurses have been cited as a leading cause of adverse events in healthcare, and specialized messaging and decision protocols have been developed to minimize the impact of errors ([22]; see [5, 26] for other examples of information loss among healthcare providers). While information constraints play a central role in the aforementioned models and applications, in contrast to our work, they often serve as an implicit motivation behind a chosen design, rather than an explicit constraint with respect to which an optimal policy is to be identified. As a result, there has been little understanding as to what policies are optimal for a given level of information availability, and what the fundamental impact information has on system performance.

Taking a more principled approach to policy design, several recent papers have aimed at rigorously *quantifying* the performance impact of information in dynamic resource allocation. [11] characterize how the average delay in a load-balancing system scales depending on the rate of messaging between the dispatcher and the servers, as well as the size of the dispatcher memory. In the

context of queueing admission control, [30] and [39] quantify how the system's optimal heavy-traffic delay scales as a function of the amount of future information available. In contrast to our approach, however, these papers largely focus on a *specific* model of information constraint, e.g., captured by the rate of messaging or length of the lookahead window, while our framework allows for a substantially more general family of information models, achieved by using different channels. To the best of our knowledge, the present paper is one of the first attempts at rigorously establishing the link between information and the performance of a resource allocation system at this level of generality.

At a high-level, our framework is inspired by information theory, and more specifically, the research on feedback control under communication constraints ([28, 36]; see [40] for a survey). This literature studies the problem of stabilizing (i.e., minimizing the magnitude of the state) a linear dynamical system of the form: $X_{t+1} = AX_t + N_t + U_t$, where X_t is the state, N_t a noise disturbance, U_t the control action, and A a gain matrix, and the decision maker has access to the state X_t only through a rate-limited communication channel, similar to the scenario depicted in Figure 1. While our framework also admits a feedback loop over a communication channel, the dynamics in our problem differ fundamentally from those in a linear dynamical system, and consequently, so do the design approaches and analysis. The difference stems from the fact that the state process in a linear dynamical system is driven *multiplicatively* by the gain matrix, A , whereas in our system, it is updated in an *additive* manner (see Eq. (2.1)). Consequently, in the control setting, even when all parameters are known (e.g., gain matrix, noise distribution, etc) an informative channel is still *necessary* for stabilization (cf. [36]). In sharp contrast, if all parameters are known in our SPII, it becomes trivial to stabilize the queues without any feedback: the decision maker can simply choose a stationary, randomized allocation policy that matches the arrival rate vector λ .

A major theme on the dynamic control of SPNs concerns the setting where the decision maker does not have complete information about the underlying system. For example, the classical Max-Weight policy ([8, 21, 35]) is oblivious to detailed statistics of the arrival process. There is also a literature that addresses the setting where the service rates (or the service time distributions) are not known completely, and the decision maker needs to either learn these parameters or develop scheduling policies that do not depend on service rates ([4, 10, 18, 19, 25, 31]). Let us note that in all this literature, even though the decision maker has partial or no information on *system parameters*, she has full information on *system states*. In contrast, our model only assumes (often severely) noisy observations of system states. This fundamental difference requires us to take a very different approach in designing policies, and to develop new tools for analyzing them.

On the methodological front, our program involves the development of a host of new techniques, largely from first principles by combining ideas from areas such as information theory, learning and queueing theory. As a sub-module of one of the policies we propose, we also create a simple yet powerful generalization of the Max-Weight policy, in which individual Markov chains are selected dynamically, in a manner analogous to how schedules are used in a conventional Max-Weight policy.

⁷In this theorem, the notation $k \geq 0$ should be interpreted as a short-hand for k belonging to the extended positive integers: $k \in \mathbb{Z}_+ \cup \{\infty\}$.

5 OVERVIEW OF PROOF TECHNIQUES

The remainder of the paper is devoted to the proof of Theorem 3.3, and we provide in this section an overview of the key ideas.

Item 4 of Theorem 3.3 (Section 7) - Capacity factor is less than one for ε -majorizing channels. We show that any level of noise in the form of an ε -majorizing channel always reduces the capacity region. The intuition is that the channel noise causes any output symbol to appear with sufficiently positive probability, and thus limits the allocation policy's ability to adapt to different arrival rates with sufficient precision. The proof employs a *lifting* argument, whereby we carry out the analysis in a higher-dimensional product space for the output alphabet, with the added dimension capturing the realization of the channel noise. A coupling argument is then used to show that noise reduces capacity region.

Items 1 and 2 of Theorem 3.3 (Sections 8 and 9) - Memoryless and Finite-Memory Receivers. This is the most technically challenging part of our program. We begin with the simple case of a memoryless receiver ($\ell = 0$), where the size of the encoder memory k has *no effect* on the capacity factor. The key intuition is a *change of perspective*: instead of viewing the receiver (allocation policy) as the one making scheduling decisions based on noisy information, it turns out that the correct way to design the system is to treat the *encoder* as the more "intelligent" of the two policies that conducts Max-Weight-like scheduling, where the encoding policy treats the set of input symbols, \mathcal{X} , as its "scheduling actions." By formulating a transformed, but equivalent, scheduling problem from the perspective of the encoder, we then use a version of the Max-Weight policy to establish stability.

For the general case of $\ell > 0$, we obtain a slightly weaker result than that of $\ell = 0$, showing that the encoding memory size becomes irrelevant after a *finite threshold*, $K(\Pi, \mathcal{X})$. The proof builds on the same intuition of viewing the encoding policy as the main decision-maker. However, a non-trivial receiver memory will mandate a substantially more sophisticated argument. This is because in each time slot, the induced service action no longer depends solely on the input signal $X(t)$, as in the case of memoryless receiver; it will now also depend on the state of the receiver memory, which is by itself a stochastic process, and hence conventional Lyapunov arguments for Max-Weight cannot be applied. Instead, we will formulate a generalized version of the conventional Max-Weight policy, dubbed the *Episodic Max Weight (EMW)* policy, where the encoding policies switches between a family of *Markov chains*, as opposed to input symbols, in a manner that is analogous to how schedules are used in conventional Max-Weight. The stability proof heavily exploits a certain conditional independence properties among different elements of the overall process $\mathbf{W}(t)$, which in turn stems from the feedback structure of the SPII.

Item 3 of Theorem 3.3 (Appendix C) - Infinite-Memory Receiver. The last part of the proof shows that as the receiver memory size $\ell \rightarrow \infty$, the capacity factor always converges to 1, regardless of the size of encoder memory. The argument is relatively straightforward compared to the other parts. Since the receiver has abundant memory in this regime, the main idea will be shifting the burden of decision-making *back* to the receiver. We construct an *Episodic Greedy Learning (EGL)* policy, where the receiver first estimates the arrival rates from the noisy messages, and subsequently deploys

a randomized schedule that dominates the estimated arrival rate vector in expectation. With more memory, the receiver is able to estimate arrival rates more accurately, leading to capacity factors that are arbitrarily close to 1.

6 PRELIMINARIES

The main purpose of this section is to establish some results and conventions that will be used throughout the remainder of the paper. Section 6.1 introduces a generalized formulation of the Max-Weight policy, whose stability properties will be used as a sub-module in subsequent proofs. Since our primary focus is on stability, we will often be concerned with the question of whether the *long-run average* service rates dominate the arrival rates. Section 6.2 formalizes the notion of long-run average service rates for our model, which is used extensively in later sections.

6.1 A Generic Max-Weight Stability Theorem

In this section, we present a simple generalization of the stability result of the celebrated Maximum Weight (Max-Weight) policy ([35]) to a class of systems that is more general than those typically seen in prior literature (e.g., [8, 29, 32]). This result, Proposition 6.1, will be used as a basic building block in our subsequent proofs. We first describe the setup, and then present the stability result. The proof is a simple modification of the standard stability proof of the Max-Weight policy, which we include for completeness.

Consider a discrete-time, irreducible Markov chain $\{\mathbf{Z}(t)\}_{t \in \mathbb{Z}_+}$ with two components $Q(\cdot)$ and $G(\cdot)$, so that for any time t , $\mathbf{Z}(t) = (Q(t), G(t))$. Here $G(\cdot)$ takes value in a finite set \mathcal{G} , and $Q(\cdot) \in \mathbb{Z}_+^N$ evolves according to the following dynamics:

$$Q(t) = (Q(t-1) - D(t))^+ + A(t) - R(t), \quad t \in \mathbb{N}. \quad (6.1)$$

Here, $A(t)$, $D(t)$ and $R(t)$ are all random vectors taking values in \mathbb{Z}_+^N . It is useful to think of the system dynamics in the following way. During each time slot t , upon observing the current queue size vector $Q(t-1)$, the system makes the service allocation decision $D(t)$, which is used twice in the current slot. First, it is used as much as possible to reduce the queue sizes $Q(t-1)$, as represented by the term $(Q(t-1) - D(t))^+$. There may be residual services left from this first use of $D(t)$. Then, arrivals $A(t)$ take place, and some of the residual services may be used to serve the arrivals $A(t)$; this portion of the residual services is denoted by $R(t)$.⁸

More formally, we require the random vectors $A(t)$ to be i.i.d. with finite second moment, and independent from the rest of the system, with $\mathbb{E}(A(1)) = \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N) \in \mathbb{R}_+^N$. The variables $D(t)$ are also required to have finite second moments, which are uniformly bounded above by a constant that does not depend on the time index t . Furthermore, given $\mathbf{Z}(t)$, $D(t+1)$ is conditionally independent from the past history $\{\mathbf{Z}(s) : s < t\}$. $R(t)$ are required to satisfy the following properties w.p.1:

⁸Let us provide some remarks about the differences between the process $\mathbf{Z}(\cdot)$ and those in prior literature, as well as how these differences are used in the present paper. Compared to other works on SPNs (e.g., [29]), $\mathbf{Z}(\cdot)$ has the additional component $G(\cdot)$, which is included to model the signals $X(t)$ and the memory contents $M_e(t)$ and $M_r(t)$. The "residual service" term $R(t)$, also absent from related prior literature, will be used in Section 9.3 to describe the dynamics of $\mathbf{W}(\cdot)$ (recall (2.9)) under a so-called *episodic Max-Weight policy*.

- (a) $R(t) \leq \min\{D(t), A(t)\}$, and $R(t)$ guarantees that $Q(t+1) \in \mathbb{Z}_+^N$;
(b) given $Z(t)$, $R(t+1)$ is conditionally independent from the past history $\{Z(s) : s < t\}$.

We are now ready to state the stability result. The proof is given in Appendix B.1.

Proposition 6.1. Let \mathcal{D} be a non-empty, finite subset of \mathbb{R}_+^N that satisfies Assumptions 1 and 2. Define the sets

$$\mathcal{D}^*(t) \triangleq \arg \max_{\mathbf{d} \in \mathcal{D}} \langle Q(t), \mathbf{d} \rangle, \quad t \in \mathbb{Z}_+. \quad (6.2)$$

Suppose that $\boldsymbol{\alpha} \in \text{conv}^-(\mathcal{D})$, and for all $t \in \mathbb{Z}_+$,

$$\mathbb{E}(D(t+1) | Z(t)) \geq \mathbf{d}^*, \quad \text{for some } \mathbf{d}^* \in \mathcal{D}^*(t) \text{ almost surely.} \quad (6.3)$$

Then $Z(\cdot)$ is positive recurrent.

An immediate corollary of Proposition 6.1 is when the residual services $R(t)$ are always zero.

COROLLARY 6.2. Suppose that the residual services $R(t)$ are zero w.p.1, so that $Q(\cdot)$ follows the dynamics given by (2.1). Let \mathcal{D} , $\mathcal{D}^*(\cdot)$ and $\boldsymbol{\alpha}$ be exactly as in Proposition 6.1. Then $Z(\cdot)$ is positive recurrent.

6.2 Stationary Service Rates

In conventional SPN systems, it is a simple fact that a stabilizing scheduling policy induces stationary service rates that dominate corresponding arrival rates. The presence of channel, encoder and receiver in the SPII architecture slightly complicates the notion of stationary service rates, for which we provide some discussion in this subsection.

Consider a channel $C = (\mathcal{X}, \mathcal{Y}, C)$, constants $k, \ell \in \mathbb{Z}_+$, and a policy pair (ϕ, ψ) with $\phi \in \Phi_k$ and $\psi \in \Psi_\ell$. Suppose that when the arrival rate vector is $\boldsymbol{\lambda}$, the process $W(\cdot)$ is positive recurrent under the policy pair (ϕ, ψ) . Then, $W(\cdot)$ has a unique stationary distribution, whose corresponding probability we denote by $\mathbb{P}^\infty(\cdot)$.

Next, we define some further notation. First, for each $x \in \mathcal{X}$, define

$$\gamma_x \triangleq \mathbb{P}^\infty(X(t) = x) \quad (6.4)$$

to be the stationary probability of sending signal x . By the ergodicity of the Markov chain $W(\cdot)$, we also know that w.p.1,

$$\gamma_x = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{I}_{\{X(t)=x\}}, \quad (6.5)$$

the long-run average fraction of time that the sent signal is x . Second, for each $y \in \mathcal{Y}$ and $\mathbf{d} \in \Pi$, the following conditional probability is well-defined:

$$M_{y, \mathbf{d}} = \mathbb{P}^\infty(D(t) = \mathbf{d} | Y(t) = y). \quad (6.6)$$

We use M to denote the $c_{\mathcal{Y}} \times |\Pi|$ -matrix $(M_{y, \mathbf{d}})_{y \in \mathcal{Y}, \mathbf{d} \in \Pi}$, and call it the *rate allocation matrix*. Finally, we also represent the set of allowable schedules in a matrix form. Let $S = (S_{\mathbf{d}, i})$ be the $|\Pi| \times N$ *schedule matrix* where

$$S_{\mathbf{d}, i} = \text{number of services offered to queue } i \text{ when schedule } \mathbf{d} \text{ is chosen.} \quad (6.7)$$

Let us note that in general, $\boldsymbol{\gamma}$ and M may depend on $\boldsymbol{\lambda}$, whereas S does not. With the preceding notation in mind, we define the *stationary service rate vector* $\boldsymbol{\mu} = \boldsymbol{\mu}(\phi, \psi, \boldsymbol{\lambda})$ to be

$$\boldsymbol{\mu} = \boldsymbol{\mu}(\phi, \psi, \boldsymbol{\lambda}) = \boldsymbol{\gamma}CMS, \quad (6.8)$$

so that for each i , μ_i represents the long-run average service rate offered to queue i . The following lemma is a simple but useful fact. The proof follows simply from the fact that $W(\cdot)$ is positive recurrent.

LEMMA 6.3. Let $C = (\mathcal{X}, \mathcal{Y}, C)$, $\boldsymbol{\lambda}$, and $W(\cdot)$ be described as above, and $\boldsymbol{\mu}$ be defined by (6.8). Then, $\boldsymbol{\mu} \geq \boldsymbol{\lambda}$.

7 CAPACITY FACTORS ARE NON-TRIVIAL FOR ε -MAJORIZING CHANNELS

In this section, we will establish Item 4 of Theorem 3.3, that capacity factors are non-trivial for ε -majorizing channels (recall Definition 3.2), i.e., that they are in general less than 1. Towards this end, we prove the following

THEOREM 7.1. Consider a system with schedule set Π that satisfies Assumptions 1, 2 and 3. Fix any finite $\ell \in \mathbb{Z}_+$, $k \geq 0$ (recall that $k \geq 0$ means $k \in \mathbb{Z}_+ \cup \{\infty\}$), and $\varepsilon \in (0, 1)$. Then, for any ε -majorizing channel $C = (\mathcal{X}, \mathcal{Y}, C)$,

$$\rho_{k, \ell}^*(C) < 1. \quad (7.1)$$

The main idea behind the proof of Theorem 7.1 is that because the channel is ε -majorizing, each combination of output symbol $Y(t)$ and receiver memory content $M_r(t)$ appears with sufficiently positive probability. As a result, no allocation policy (with finite memory) is *adaptive* enough to stabilize all arrival rate vectors in Λ . The proof is given in Appendix B.3.

The next theorem states a stronger result, for the special case of memoryless receiver, i.e., $\ell = 0$. It gives a tight characterization of the $(k, 0)$ -capacity factor of any ε -majorizing channel, by providing an achievable upper bound, $\rho(\varepsilon, \Pi)$, on the capacity factors.

THEOREM 7.2. Consider a system with schedule set Π that satisfies Assumptions 1, 2 and 3. Then, for any $\varepsilon \in (0, 1)$, there exists a constant $\rho(\varepsilon, \Pi) < 1$, which only depends on ε and Π , such that:

- (1) For any ε -majorizing channel $C = (\mathcal{X}, \mathcal{Y}, C)$, and for any $k \geq 0$,

$$\rho_{k, 0}^*(C) \leq \rho(\varepsilon, \Pi). \quad (7.2)$$

- (2) Conversely, there exists an ε -majorizing channel, C , such that

$$\rho_{0, 0}^*(C) = \rho(\varepsilon, \Pi). \quad (7.3)$$

The next subsection proves Theorem 7.3, a special case of Theorem 7.2, for systems of parallel queues with a single server, where we derive a simple, explicit expression for the bound $\rho(\varepsilon, \Pi)$. The proof of Theorem 7.2 is more abstract than that of Theorem 7.3 and involves using general properties of convex polyhedron, but it follows a similar argument, and is provided in Appendix B.4.

7.1 An Example: Single Server Scheduling of Parallel Queues

In this subsection, we prove Theorem 7.3, a special case of Theorem 7.2. While more restricted in scope, it best illustrates the key intuition present in the more general version of the result. Consider the

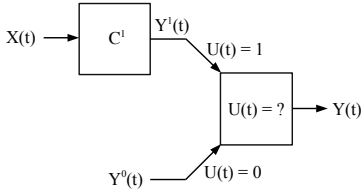


Figure 4: Equivalent construction of an ε -majorizing channel.

following system that consists of one server and N parallel queues, where the queues are indexed by $[N]$. In each time slot, the server picks one queue, so that exactly one job departs from the chosen queue if it is non-empty, and no departure occurs in the system, otherwise. Thus, the schedule set Π is given by

$$\Pi = \{\mathbf{0}\} \cup \{\mathbf{e}^{(i)}\}_{i=1}^N, \quad (7.4)$$

where $\mathbf{0}$ is the zero vector, and $\mathbf{e}^{(i)}$ is the standard i th unit vector, $i \in [N]$. It is easy to see that the capacity region of this system is given by

$$\Lambda = \left\{ \boldsymbol{\lambda} \geq \mathbf{0} : \sum_{i=1}^N \lambda_i < 1 \right\}. \quad (7.5)$$

The following theorem provides a tight characterization on the capacity factors of ε -majorizing channels, when the allocation policy is memoryless.

THEOREM 7.3. Fix $\varepsilon \in (0, 1)$. Consider a single-server system with N parallel queues. Then, the following hold.

(1) For any ε -majorizing channel, $C = (\mathcal{X}, \mathcal{Y}, C)$, and any $k \geq 0$,

$$\rho_{k,0}^*(C) \leq 1 - \varepsilon(1 - N^{-1}). \quad (7.6)$$

(2) Conversely, there exists an ε -majorizing channel, C , such that

$$\rho_{0,0}^*(C) = 1 - \varepsilon(1 - N^{-1}). \quad (7.7)$$

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 $k \geq 0$ (see Theorem 8.1), so it suffices to show that $\rho_{0,0}^*(C) \leq 1 - \varepsilon(1 - N^{-1})$ for any ε -majorizing channel.

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

$$C = \varepsilon C^0 + (1 - \varepsilon)C^1, \quad (7.8)$$

according to the decomposition in Definition 3.2. 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 C , it is equivalent to (a) first generate a *switching* Bernoulli random variable $U(t)$, with $\mathbb{P}(U(t) = 0) = \varepsilon$, independent of everything else; (b) second, if $U(t) = 1$, $Y(t)$ is set to be $Y^1(t)$, the message generated from $X(t)$ through the channel $(\mathcal{X}, \mathcal{Y}, C^1)$, and if $U(t) = 0$, $Y(t)$ 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 C .

We shall adopt the preceding interpretation of C for the remainder of the proof. Consider a channel C^+ constructed by augmenting

the output of C the value of the switching variable, $U(t)$. Specifically, the output message of C^+ is equal to $(Y(t), U(t))$, taking values in the alphabet set $\mathcal{Y}^+ = \mathcal{Y} \times \{0, 1\}$. Note that \mathcal{Y}^+ is still a finite set and hence C^+ is a valid channel. It is easy to see that for any policy pair (ϕ, ψ) operating under the channel C , we can construct a corresponding pair (ϕ, ψ^+) operating under the channel C^+ by simply ignoring the variable $U(t)$, so that $\tilde{\Lambda}(\phi, \psi) \subseteq \tilde{\Lambda}(\phi, \psi^+)$. We hence conclude that $\rho_{k,0}^*(C) \leq \rho_{k,0}^*(C^+)$, for any $k \geq 0$. In light of the fact that $\rho_{k,0}^*(C) = \rho_{0,0}^*(C)$, which will be proved in Section 8, it suffices to show that

$$\rho_{0,0}^*(C^+) \leq 1 - \varepsilon(1 - N^{-1}). \quad (7.9)$$

Let (ϕ, ψ^+) be a pair of memoryless encoding and allocation policies that works with the channel 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 $U(t) = 0$, i.e.,

$$\mu_i^0 = \mathbb{P}_{\psi^+} \left(D(t) = \mathbf{e}^{(i)} \mid U(t) = 0 \right), \quad t \in \mathbb{Z}_+. \quad (7.10)$$

Note that conditioning on $U(t) = 0$, 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 $\min_{i=1,2,\dots,N} \mu_i^0 \leq \frac{1}{N}$. Let $\boldsymbol{\mu} = \boldsymbol{\mu}(\phi, \psi^+, \boldsymbol{\lambda})$ be the vector of stationary service rates defined in Eq. (6.8) of Section 6.2. Since $\mathbb{P}(U(t) = 0) = \varepsilon$, 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, $\mu_{i^*} \leq 1 - \varepsilon \left(\sum_{i \neq i^*} \mu_i^0 \right) \leq 1 - \varepsilon(1 - N^{-1})$. We claim that $\boldsymbol{\lambda}$ cannot be of the following form:

$$\lambda_{i^*} > 1 - \varepsilon(1 - N^{-1}) \quad \text{and} \quad \lambda_i = 0, \quad \forall i \neq i^*. \quad (7.11)$$

This is because if (7.11) holds, then queue i^* cannot be stable, because $\lambda_{i^*} > 1 - \varepsilon(1 - N^{-1}) \geq \mu_{i^*}$. Since ϕ and ψ^+ are arbitrary, we conclude that

$$\rho_{0,0}^*(C^+) \leq 1 - \varepsilon(1 - N^{-1}), \quad (7.12)$$

which, by Eq. (7.9), proves the first statement of the theorem.

We now prove the second statement of the theorem by considering the following example. Let $\mathcal{X} = \mathcal{Y} = \{1, 2, \dots, N\}$, 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.,

$$I(t) \triangleq \min\{i = 1, 2, \dots, N : Q_i(t) > 0\}, \quad (7.13)$$

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: $D(t) = \mathbf{e}^{(Y(t))}$ 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 $\|\mathbf{Q}(t)\|_1 \triangleq \sum_{i=1}^N Q_i(t)$. 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 $\|\mathbf{Q}(t)\| > 0$. 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 $\|\mathbf{Q}(t)\|$ 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 simply 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 $C = (\mathcal{X}, \mathcal{Y}, C)$ satisfies

$$\rho_{0,0}^*(C) \geq 1 - \varepsilon(1 - N^{-1}), \quad (7.14)$$

which, in light of the first claim of the theorem, implies that $\rho_{0,0}^*(C)$ is in fact equal to $1 - \varepsilon(1 - N^{-1})$. This concludes the proof of Theorem 7.3. \square

8 MEMORYLESS RECEIVER

In this section, we show Item 1 of Theorem 3.3: if the allocation policy is memoryless (i.e., $\ell = 0$), then there exists some memoryless encoding policy $\phi \in \Phi_0$, which achieves the optimal capacity factor among all encoding policies, memoryless or not. In short, when the receiver does not have memory, then one will not benefit from any memory at the encoder, either.

THEOREM 8.1. *Fix any channel $C = (\mathcal{X}, \mathcal{Y}, C)$. We have that*

$$\rho_{0,0}^*(C) = \rho_{\infty,0}^*(C). \quad (8.1)$$

PROOF. It is clear that $\rho_{0,0}^*(C) \leq \rho_{\infty,0}^*(C)$, so it suffices to show that $\rho_{0,0}^*(C) \geq \rho_{\infty,0}^*(C)$. For the rest of the proof, we write ρ for $\rho_{\infty,0}^*(C)$ to simplify notation.

Let $\varepsilon > 0$. We will show that $\rho_{0,0}^*(C) \geq \rho - \varepsilon$ by constructing a policy pair (ϕ, ψ) with $\phi \in \Phi_0$ and $\psi \in \Psi_0$, under which $(\rho - \varepsilon)\Lambda \subseteq \Lambda(\phi, \psi)$. By the definition of $\rho_{\infty,0}^*(C)$, there exist $k \in \mathbb{N}$, encoding policy $\phi_k \in \Phi_k$ and allocation policy $\psi \in \Psi_0$, such that $\rho^*(\phi_k, \psi, C)$, the capacity factor of C under (ϕ_k, ψ) , satisfies $\rho^*(\phi_k, \psi, C) \geq \rho - \varepsilon/3$.

Let $\lambda \in (\rho - 2\varepsilon/3)\Lambda \subseteq \tilde{\Lambda}(\phi_k, \psi)$, and consider a system with arrival rate vector λ , operating under the policy pair (ϕ_k, ψ) . Let the stationary service rate vector $\mu = \gamma CMS$ be defined by (6.8), and γ, M and S defined by (6.4), (6.6) and (6.7) respectively. Then, by Lemma 6.3, $\mu \geq \lambda$. Furthermore, because the allocation policy ψ is memoryless, the matrix M depends only on the policy ψ , but not on λ . Thus, the matrix CMS is fixed, and $\lambda \leq \text{conv}(\widehat{S})$, where \widehat{S} denotes the set consisting of the rows of CMS : $\widehat{S} = \{(CMS)_x : x \in \mathcal{X}\}$. Here $(CMS)_x$ represents the x th row of CMS . Since λ is arbitrary except $\lambda \in (\rho - 2\varepsilon/3)\Lambda$, this implies that

$$(\rho - \varepsilon)\Lambda \subseteq \text{conv}^-(\widehat{S}). \quad (8.2)$$

We are now ready to define the memoryless encoding policy $\phi \in \Phi_0$ and prove that $\rho^*(\phi, \psi, C) \geq \rho - \varepsilon$. For all time $t \in \mathbb{N}$, the sent signal $X(t)$ is a deterministic function of $Q(t-1)$, defined to be

$$X(t) = \phi(Q(t-1)) \in \arg \max_{x \in \mathcal{X}} (Q(t-1), (CMS)_x), \quad (8.3)$$

where ties are broken arbitrarily. Note that $(CMS)_x$ is the vector of expected services offered when alphabet x is chosen to be the signal.

Let us summarize the encoding-allocation policy pair (ϕ, ψ) that we have constructed so far:

- (1) encoding policy ϕ : send a signal x^* according to Eq. (8.3). Note that the optimization in Eq. (8.3) only involves the current queue lengths, so ϕ is memoryless.
- (2) allocation policy ψ : one that corresponds to the allocation policy matrix M defined by Eq. (6.6).

Under the pair (ϕ, ψ) , the expected services offered at time t is given by $(CMS)_{X(t)}$. The set \widehat{S} satisfies Assumption 2 because of Eq. (8.2), and the fact that Λ satisfies Assumption 2. Thus, we can apply Corollary 6.2 to the process $W(\cdot)$, and conclude that for any $\lambda \in (\rho - \varepsilon)\Lambda \subseteq \text{conv}^-(\widehat{S})$, $W(\cdot)$ is positive recurrent under the pair (ϕ, ψ) . This implies that $\rho_{0,0}^*(C) \geq \rho^*(\phi, \psi, C) \geq \rho - \varepsilon$.

Since $\varepsilon > 0$ is arbitrary, we have that $\rho_{0,0}^*(C) \geq \rho = \rho_{\infty,0}^*(C)$. This concludes the proof. \square

9 FINITE-MEMORY RECEIVER

We consider in this section the scenario where the size of the receiver memory, ℓ , is finite. Recall that we assume the existence of *memory-feedback*, so that the encoder can fully observe $M_r(t)$, the receiver memory, and use it as input to the policy design: $(X(t), M_e(t)) = \phi(Q(t-1), A(t-1), M_e(t-1), M_r(t-1), Z_e(t-1))$.

The goal of this section is to prove Item 2 of Theorem 3.3. Towards this end, we define the notion of *simple (encoding) policies* in Section 9.1, and establish a key property of simple policies in Proposition 9.2. The *Episodic Max-Weight* policy, defined in Section 9.2, makes use of simple policies, and is itself used in Section 9.3 to prove Theorem 9.5 (Item 2 of Theorem 3.3). Along the way, we also provide an explicit expression for $K(\Pi, \mathcal{X})$, the threshold beyond which $\rho_{k,\ell}^*$ are all the same.

9.1 Projection to Simple Policies

Definition 9.1 (Simple Encoding Policies). Fix $\ell \in \mathbb{N}$ and $\psi \in \Psi_\ell$. An encoding policy ϕ^0 is called *simple* for the allocation policy ψ , if

- (1) $\phi^0 \in \Phi_{\lceil \log |\mathcal{X}| \rceil}$.
- (2) ϕ^0 takes as inputs only $M_e(t-1)$, $M_r(t-1)$ and $Z_e(t-1)$, and does not rely on $Q(t-1)$.
- (3) $\{(X(t), M_r(t))\}_{t \in \mathbb{N}}$ is a homogeneous irreducible Markov chain under the pair (ϕ^0, ψ) .

We use $\Phi^S(\psi)$ to denote the set of all simple encoding policies for ψ . When the context is clear, we sometimes write Φ^S for $\Phi^S(\psi)$, and call a policy simple if it is simple for ψ .

Recall $\mu(\phi, \psi, \lambda)$, the stationary service rate vector defined by (6.8) when the arrival rate vector is λ , and the system is operating under the policy pair (ϕ, ψ) . If the service rates do not change as a function of λ , then we write $\mu(\phi, \psi)$ and omit the dependence on λ . Let us note that for any simple policy ϕ^0 , even though $W(\cdot)$ may not be positive recurrent under the policy pair (ϕ^0, ψ) and arrival rate vector λ , the stationary service rate vector $\mu(\phi^0, \psi, \lambda)$ is still well-defined according to (6.8), because $\{(X(t), M_r(t))\}_{t \in \mathbb{N}}$ is irreducible with a finite state space. Furthermore, because ϕ^0 does not rely on queue lengths when generating signals, $\mu(\phi^0, \psi, \lambda)$ does not depend on λ , in which case we can just write $\mu(\phi^0, \psi)$.

The following proposition is the main result of this subsection.

Proposition 9.2. Fix $k, \ell \in \mathbb{N}$, $\phi \in \Phi_k$, $\psi \in \Psi_\ell$, and $\lambda \in \tilde{\Lambda}(\phi, \psi)$. There exists a simple encoding policy, $\phi^0 \in \Phi^S$, such that $\mu(\phi^0, \psi) \geq \lambda$.

Proposition 9.2 provides the technical foundation for the design of the EMW policy in Section 9.2. With the proposition on hand, we can “approximately” cover the entire schedule set Π (see Lemma 9.3 for details), and the EMW policy will switch between simple policies in this covering, similar to how schedules are chosen in conventional Max-Weight. The proof of Proposition 9.2 heavily exploits the conditional independence of $X(t)$ and $M_r(t)$ given $X(t-1)$ and $M_r(t-1)$, for any t . See Appendix B.2 for the detailed proof.

9.2 Episodic Max Weight

We present in this section our main policy, Episodic Max Weight, for the regime of $0 < \ell < \infty$, which will employ the simple policies (Definition 9.1) as basic building blocks. Recall from Eq. (2.11) that $\Lambda = \text{conv}^-(\Pi)$, the set of points strictly dominated by the convex hull of Π . Let \mathcal{F} be the set of all maximal elements of Π , which we denote by $\mathcal{F} = \{\mu_1^{\mathcal{F}}, \dots, \mu_{|\mathcal{F}|}^{\mathcal{F}}\}$. Then, it is easy to see that we also have $\Lambda = \text{conv}^-(\mathcal{F})$. The next lemma is a useful structural result concerning the simple policies; the proof is given in Appendix B.5.

LEMMA 9.3. Fix $k, \ell \in \mathbb{N}$ and $\varepsilon \in (0, \rho_{k,\ell}^*)$. There exist an allocation policy $\psi^\varepsilon \in \Psi_\ell$, and a set Φ^ε of $|\mathcal{F}|$ simple encoding policies, $\Phi^\varepsilon \triangleq \{\phi_1^\varepsilon, \dots, \phi_{|\mathcal{F}|}^\varepsilon\} \subseteq \Phi^S$, such that for each $i \in [|\mathcal{F}|]$, the chain $(X(\cdot), M_r(\cdot))$ is irreducible and aperiodic under $(\phi_i^\varepsilon, \psi^\varepsilon)$, and

$$(\rho_{k,\ell}^* - \varepsilon)\Lambda \subseteq \text{conv}^-(\mu(\Phi^\varepsilon, \psi^\varepsilon)). \quad (9.1)$$

The Episodic Max Weight Policy. Fix $k, \ell \in \mathbb{N}$ and $\varepsilon \in (0, \rho_{k,\ell}^*)$. Let Φ^ε and $\psi^\varepsilon \in \Psi_\ell$ be defined as in Lemma 9.3. The Episodic Max Weight (EMW) encoding policy is parameterized by a expected episode length parameter, $B \in \mathbb{N}$, whose memory requirement is

$$K(\Pi, \mathcal{X}) = \log(|\mathcal{F}|) + \log(|\mathcal{X}|). \quad (9.2)$$

In every time slot, the EMW policy applies a simple encoding policy from the set Φ^ε to generate the signal, where Φ^ε is defined as in Lemma 9.3. The encoder memory stores only two types of information:

- (1) The index of the simple encoding policy in Φ^ε that is currently used to generate the signal $X(t+1)$. Since Φ^ε contains $|\mathcal{F}|$ simple policies, this can be implemented using $\log(|\mathcal{F}|)$ bits of memory.
- (2) The current signal $X(t)$. This can be implemented using $\log(|\mathcal{X}|)$ bits of memory.

We denote the first $\log(|\mathcal{F}|)$ bits and last $\log(|\mathcal{X}|)$ bits of $M_e(t)$ by $M_e^{\mathcal{F}}(t)$ and $M_e^{\mathcal{X}}(t)$, respectively.

Definition 9.4 (*B*-Episodic Max Weight (*B*-EMW)). Fix $B \in \mathbb{N}$. Let $\{Z_t^P\}_{t \in \mathbb{N}}$ be a sequence of i.i.d. Bernoulli random variables with $\mathbb{P}(Z_1^P = 1) = \frac{1}{B}$, independently from everything else. For each $t \in \mathbb{N}$:

- (1) If $Z_t^P = 1$, let $i^* \in \arg \max_{i=1, \dots, |\mathcal{F}|} \langle Q(t), \mu(\phi_i^\varepsilon, \psi^\varepsilon) \rangle$, with ties broken arbitrarily. Set $M_e^{\mathcal{F}}(t) = i^*$.
- (2) If $Z_t^P = 0$, set $M_e^{\mathcal{F}}(t) = M_e^{\mathcal{F}}(t-1)$.

Apply the simple encoding policy $\phi_{M_e^{\mathcal{F}}(t)}^\varepsilon$ to generate $X(t)$. Set $M_e^{\mathcal{X}}(t) = X(t)$.

9.3 Asymptotic Feasible Region under EMW

We now present the main result of this section, which will imply Item 3 of Theorem 3.3.

THEOREM 9.5. Let $k, \ell \in \mathbb{N}$, and $\varepsilon \in (0, \rho_{k,\ell}^*)$. Let ϕ^ε be defined as in Lemma 9.3, and denote by ϕ^B the *B*-EMW policy defined in Definition 9.4. Then, there exists $B^* > 0$ such that

$$(\rho_{k,\ell}^* - 2\varepsilon)\Lambda \subseteq \tilde{\Lambda}(\phi^B, \psi^\varepsilon), \quad \text{for all } B > B^*. \quad (9.3)$$

Because $\phi^B \in \Phi_{K(\Pi, \mathcal{X})}$ and Eq. (9.3) holds for all k and ε , we have that $\rho_{k,\ell}^* \leq \rho_{K(\Pi, \mathcal{X}), \ell}^*$ for all k , so

$$\rho_{K(\Pi, \mathcal{X}), \ell}^*(C) = \rho_{\infty, \ell}^*(C). \quad (9.4)$$

PROOF. Denote by T_n the n th time that $Z_t^P = 1$, and call the time slots between T_n and $T_{n+1} - 1$ (including both T_n and $T_{n+1} - 1$), the n th episode. It is easy to see that the episode lengths are i.i.d geometric random variables with mean B , and that within each episode, the same simple policy is employed by ϕ^B . Define $A[n] = \sum_{t=T_{n-1}}^{T_n-1} A(t)$, $D[n] = \sum_{t=T_{n-1}}^{T_n-1} D(t)$, $Q[n] = Q(T_n)$, and $W[n] = W(T_n)$. Because the Z_t^P 's are i.i.d Bernoulli random variables which are independent from everything else, it is easy to see that $W[\cdot]$ is a Markov chain. Furthermore, $W[\cdot]$ is of the form $W[\cdot] = (Q[\cdot], G[\cdot])$, where $G[n] = (M_e(T_n), X(T_n), M_r(T_n))$ taking values in a finite-state space, and $Q[\cdot]$ evolves according to the dynamics

$$Q[n] = (Q[n-1] - D[n])^+ + A[n] - R[n], \quad (9.5)$$

with $R[n]$ satisfying properties (a) and (b) in Section 6.1. To verify Eq. (9.5), we can think of the allocations $D[n]$ as first being used to serve $Q[n-1]$, the queue lengths at the beginning of the episode, and any residual allocations as being used to serve subsequent arrivals in the episode. These residual allocations are captured by the term $R[n]$.

Denote by $\phi[n]$ the simple policy used in the n th episode. Recall the set $\Phi^\varepsilon = \{\phi_i^\varepsilon : i \in [|\mathcal{F}|]\}$ of simple policies defined in Lemma 9.3, and that for each $i \in [|\mathcal{F}|]$, $(X(\cdot), M_r(\cdot))$ is irreducible and aperiodic under $(\phi_i^\varepsilon, \psi^\varepsilon)$. Since Φ^ε is a finite set, the convergence speeds to the steady-state distribution of $(X(\cdot), M_r(\cdot))$ under different $(\phi_i^\varepsilon, \psi^\varepsilon)$, $i \in [|\mathcal{F}|]$, can be uniformly upper bounded. Also recall that the episode lengths are i.i.d geometric random variables with mean B , which are independent from everything else. Thus, there exists $B_\varepsilon^* > 0$, such that if $B > B_\varepsilon^*$, then

$$\mathbb{E}(D[n+1] | W[n]) \geq B(1-\varepsilon)\mu(\phi[n], \psi^\varepsilon), \quad \forall n \in \mathbb{N}, \text{ a.s.} \quad (9.6)$$

Fix $B > B_\varepsilon^*$, and let $\lambda \in (1-\varepsilon)(\rho_{k,\ell}^* - \varepsilon)\Lambda$. By Lemma 9.3, $(\rho_{k,\ell}^* - \varepsilon)\Lambda \subseteq \text{conv}^-(\mu(\Phi^\varepsilon, \psi^\varepsilon))$, so

$$B\lambda \in \text{conv}^-(B(1-\varepsilon)\mu(\Phi^\varepsilon, \psi^\varepsilon)). \quad (9.7)$$

Also note that

$$\mathbb{E}(A[n]) = B\lambda, \quad \forall n \in \mathbb{N}. \quad (9.8)$$

By Eqs. (9.6), (9.7) and (9.8), and Proposition 6.1, the sampled chain $W[\cdot]$ is positive recurrent under *B*-EMW whenever $\lambda \in (1 -$

$\varepsilon)(\rho_{k,\ell}^* e^{-\varepsilon})\Lambda$. Using the above fact, combined with $(1-\varepsilon)(\rho_{k,\ell}^* e^{-\varepsilon}) \geq \rho_{k,\ell}^* - 2\varepsilon$, we establish (9.3) and, hence the theorem as well. \square

10 CONCLUSION AND FUTURE WORK

We proposed in this paper the *Stochastic Processing under Imperfect Information (SPII)* framework to study the impact of information constraints and memories on the capacity of general Stochastic Processing Networks. Using a novel metric, capacity factor, our main theorem characterizes the reduction in capacity region (under “optimal” policies) under all non-degenerate channels, and across almost all memory sizes.

A few interesting questions remain open. Firstly, our results leave open the question of whether the encoder memory has *any* impact on the capacity factor when the receiver memory is finite but *non zero* ($\ell > 0$), i.e., whether we can reduce the threshold $K(\Pi, \mathcal{X})$ down to 0. Secondly, our results currently rely on the assumption of memory-feedback for the regime of finite receiver memory ℓ , which we hope can be removed in the future. Finally, we believe the SPII framework put forth in this paper can be extended to studying the impact of imperfect information and memory on other, more refined metrics, such as average delay, and more broadly, to other sequential decision-making problems.

REFERENCES

- [1] G. Ananthanarayanan, M. C.-C. Hung, X. Ren, I. Stoica, A. Wierman, and M. Yu. Grass: Trimming stragglers in approximation analytics. In *USENIX NSDI*, 2014.
- [2] M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. Above the clouds: A view of cloud computing. *Communications of the ACM*, 53(4):50–58, 2010.
- [3] R. Badonnel and M. Burgess. Dynamic pull-based load balancing for autonomic servers. In *Network Operations and Management Symposium, 2008. NOMS 2008. IEEE*, pages 751–754. IEEE, 2008.
- [4] G. Baharian and T. Tezcan. Stability analysis of parallel server systems under longest queue first. *Mathematical Methods of Operations Research*, 74:257–279, 2011.
- [5] A. J. Carter, K. A. Davis, L. V. Evans, and D. C. Cone. Information loss in emergency medical services handover of trauma patients. *Prehospital Emergency Care*, 13(3):280–285, 2009.
- [6] H. Chen, J. M. Harrison, A. Mandelbaum, A. van Ackere, and L. Wein. Empirical evaluation of a queueing network model for semiconductor wafer fabrication. *Operations Research*, 36(2):202–215, 1988.
- [7] M. Chowdhury and I. Stoica. Efficient coflow scheduling without prior knowledge. In *Proceedings of the ACM SIGCOMM*, 2015.
- [8] J. G. Dai and W. Lin. Maximum pressure policies in stochastic processing networks. 53(2):197–218, 2005.
- [9] J. G. Dai and B. Prabhakar. The throughput of switches with and without speed-up. In *Proceedings of IEEE Infocom*, pages 556–564. IEEE, 2000.
- [10] A. Dimakis and J. Walrand. Sufficient conditions for stability of longest-queue-first scheduling: Second-order properties using fluid limits. *Advances in Applied Probability*, 38(2):505–521, 2006.
- [11] D. Gamarnik, J. N. Tsitsiklis, and M. Zubeldia. Delay, memory, and messaging tradeoffs in distributed service systems. *Stochastic Systems*, 8(1):45–74, 2018.
- [12] N. Gans, G. Koole, and A. Mandelbaum. Telephone call centers: Tutorial, review, and research prospects. *Manufacturing & Service Operations Management*.
- [13] B. Hajek. *Random processes for engineers*. Cambridge university press, 2015.
- [14] J. M. Harrison. Brownian models of open processing networks: Canonical representation of workload. *Annals of Applied Probability*, 10(1):75–103, 2000.
- [15] J. M. Harrison. Stochastic networks and activity analysis. *Translations of the American Mathematical Society-Series 2*, 207:53–76, 2002.
- [16] J. M. Harrison. Correction: “brownian models of open processing networks: Canonical representation of workload”. *Annals of Applied Probability*, 16(3):1703–1732, 2003.
- [17] J. M. Harrison and A. Zeevi. A method for staffing large call centers based on stochastic fluid models. *Manufacturing & Service Operations Management*, 7(1):20–36, 2005.
- [18] S. Krishnasamy, A. Arapostathis, R. Johari, and S. Shakkottai. On learning the $c\mu$ rule in single and parallel server networks. *Preprint*, 2018.
- [19] S. Krishnasamy, R. Sen, R. Johari, and S. Shakkottai. Regret of queueing bandits. In *Advances in Neural Information Processing Systems*, 2016.
- [20] Y. Lu, Q. Xie, G. Kliot, A. Geller, J. R. Larus, and A. Greenberg. Join-idle-queue: A novel load balancing algorithm for dynamically scalable web services. *Performance Evaluation*, 68(11):1056–1071, 2011.
- [21] A. Mandelbaum and A. L. Stolyar. Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized $c\mu$ -rule. *Operations Research*, 52(6):836–855, 2004.
- [22] M. Manojlovich, J. Adler-Milstein, M. Harrod, A. Sales, T. P. Hofer, S. Saint, and S. L. Krein. The effect of health information technology on health care provider communication: a mixed-method protocol. *JMIR research protocols*, 4(2), 2015.
- [23] N. McKeown, V. Anantharam, and J. Walrand. Achieving 100% throughput in an input-queued switch. In *Proceedings of IEEE Infocom*, pages 296–302. IEEE, 1996.
- [24] M. Mitzenmacher. The power of two choices in randomized load balancing. *IEEE Transactions on Parallel and Distributed Systems*, 12(10):1094–1104, 2001.
- [25] R. Pedarsani, J. Walrand, and Y. Zhong. Robust scheduling for flexible processing networks. *Advances in Applied Probability*, 49(2):603–628, 2017.
- [26] D. Pothier, P. Monteiro, M. Mooktiar, and A. Shaw. Pilot study to show the loss of important data in nursing handover. *British Journal of Nursing*, 14(20):1090–1093, 2005.
- [27] J. Roberts and Massoulié. Bandwidth sharing and admission control for elastic traffic. *Telecommunication Systems*, 15:185–201, 2000.
- [28] A. Sahai and S. Mitter. The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link: Part I: Scalar systems. *IEEE Transactions on Information Theory*, 8(52):3369–3395, 2006.
- [29] D. Shah, D. Wischik, et al. Switched networks with maximum weight policies: Fluid approximation and multiplicative state space collapse. *The Annals of Applied Probability*, 22(1):70–127, 2012.
- [30] J. Spencer, M. Sudan, K. Xu, et al. Queueing with future information. *The Annals of Applied Probability*, 24(5):2091–2142, 2014.
- [31] A. Stolyar and E. Yudovina. Tightness of invariant distributions of a large-scale flexible service system under a priority discipline. *Stochastic Systems*, 2(2):381–408, 2012.
- [32] A. L. Stolyar. Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *The Annals of Applied Probability*, 14(1):1–53, 2004.
- [33] A. L. Stolyar. Pull-based load distribution in large-scale heterogeneous service systems. *Queueing Systems*, 80(4):341–361, 2015.
- [34] A. L. Stolyar. Pull-based load distribution among heterogeneous parallel servers: the case of multiple routers. *Queueing Systems*, 85(1-2):31–65, 2017.
- [35] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE transactions on automatic control*, 37(12):1936–1948, 1992.
- [36] S. Tatikonda and S. Mitter. Control-under communication constraints. *IEEE Transactions on Automatic Control*, 49(7):1056–1068, 2004.
- [37] J. N. Tsitsiklis and K. Xu. On the power of (even a little) resource pooling. *Stochastic Systems*, 2(1):1–66, 2012.
- [38] N. D. Vvedenskaya, R. L. Dobrushin, and F. I. Karpelevich. Queueing system with selection of the shortest of two queues: An asymptotic approach. *Problemy Peredachi Informatsii*, 32(1):20–34, 1996.
- [39] K. Xu. Necessity of future information in admission control. *Operations Research*, 63(5):1213–1226, 2015.
- [40] S. Yüksel and T. Başar. *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Springer Science & Business Media, 2013.
- [41] D. Zats, A. P. Iyer, G. Ananthanarayanan, R. Agarwal, R. Katz, I. Stoica, and A. Vahdat. Fastlane: Making short flows shorter with agile drop notification. In *ACM Symposium on Cloud Computing*, 2015.

A GLOSSARY OF FREQUENTLY USED SYMBOLS

$A(t)$	arrival vector in time slot t
C, C	channel, channel matrix
$D(t)$	chosen allocation / schedule in time slot t
k	encoder memory size
ℓ	receiver memory size
Λ	maximum capacity region
N	system size / number of queues
Π	set of allowable allocations / schedules
$Q(t)$	queue state in time slot t
$\rho_{k,\ell}^*(C)$	(k, ℓ) -capacity factor
$X(t), \mathcal{X}$	input symbol in time slot t , input alphabet
$Y(t), \mathcal{Y}$	output symbol in time slot t , output alphabet

B ADDITIONAL PROOFS

B.1 Proof of Proof of Proposition 6.1

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

$$\begin{aligned} & L(Z(t+1)) - L(Z(t)) \\ &= \|Q(t+1)\|^2 - \|Q(t)\|^2 \\ &= \sum_{i=1}^N \left\{ \left[(Q_i(t) - D_i(t+1))^+ + A_i(t+1) - R_i(t+1) \right]^2 - Q_i^2(t) \right\} \\ &\leq \sum_{i=1}^N \left\{ \left[(Q_i(t) - D_i(t+1))^+ + A_i(t+1) \right]^2 - Q_i^2(t) \right\} \end{aligned} \quad (\text{B.1})$$

$$\begin{aligned} &= \sum_{i=1}^N \left\{ \left[(Q_i(t) - D_i(t+1))^+ \right]^2 + A_i^2(t+1) \right. \\ &\quad \left. + 2A_i(t+1)(Q_i(t) - D_i(t+1))^+ - Q_i^2(t) \right\} \\ &\leq \sum_{i=1}^N \left\{ (Q_i(t) - D_i(t+1))^2 + A_i^2(t+1) \right. \\ &\quad \left. + 2A_i(t+1)Q_i(t) - Q_i^2(t) \right\} \end{aligned} \quad (\text{B.2})$$

$$= \sum_{i=1}^N \left\{ -2Q_i(t)D_i(t+1) + 2A_i(t+1)Q_i(t) + D_i^2(t+1) + A_i^2(t+1) \right\}. \quad (\text{B.3})$$

Inequality (B.1) follows from properties (a) and (c) of the ‘‘residual service’’ $R(\cdot)$, and Inequality (B.2) follows from the facts that $[(x)^+]^2 \leq x^2$ for any $x \in \mathbb{R}$, and that for any $x, y \in \mathbb{R}_+$, $0 \leq (x-y)^+ \leq x$.

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

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

Conditioning on $Z(t)$ and taking expectations on both sides of (B.3) gives

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

for some $\mathbf{d}^* \in \mathcal{D}^*(t+1)$. Since $\alpha \in \text{rel}(\text{conv}^-(\mathcal{D}))$, there exist constants $\delta \in (0, 1)$, and $p_{\mathbf{d}} \geq 0$ for each $\mathbf{d} \in \mathcal{D}$, such that $\sum_{\mathbf{d} \in \mathcal{D}} p_{\mathbf{d}} \leq 1 - \delta$ and $\alpha \leq \sum_{\mathbf{d} \in \mathcal{D}} p_{\mathbf{d}} \mathbf{d}$. Therefore,

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

Thus,

$$\begin{aligned} & \mathbb{E} \left[L(Z(t+1)) - L(Z(t)) \mid Z(t) \right] \\ &\leq -2\delta \langle \mathbf{d}^*, Q(t) \rangle + K \\ &\leq -2\delta \|Q(t)\|_{\infty} + K, \end{aligned} \quad (\text{B.4})$$

where the last inequality follows from Assumption 2.

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

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

The positive recurrence of $Z(\cdot)$ then follows from a standard application of the Foster-Lyapunov criteria (cf. [13, 35]). \square

B.2 Proof of Proposition 9.2

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

$$\mu(\phi^0, \psi) = \mu(\phi, \psi, \lambda) \geq \lambda. \quad (\text{B.5})$$

More specifically, consider the system under the policy pair (ϕ, ψ) and arrival rate vector λ . Because $\lambda \in \tilde{\Lambda}(\phi, \psi)$ by assumption, the Markov chain $\mathbf{W}(\cdot)$ is positive recurrent, so it has a unique stationary distribution. Suppose that the chain $\mathbf{W}(\cdot)$ is initialized with this unique stationary distribution at time 0, and we use $\mathbb{P}_{\phi, \psi}^{\infty}(\cdot)$ 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.,

$$M_e(t) = X(t) \quad \text{under } \phi^0, \quad (\text{B.6})$$

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

$$\begin{aligned} & \mathbb{P} \left(X(t+1) = x' \mid X(t) = x, M_r(t) = m \right) \\ &= \mathbb{P}_{\phi, \psi}^{\infty} \left(X(1) = x \mid X(0) = x, M_r(0) = m \right), \end{aligned} \quad (\text{B.7})$$

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 λ . Denote by $r(x, m)$ the stationary marginal probability

$$r(x, m) = \mathbb{P}_{\phi, \psi}^{\infty} (X(0) = x, M_r(0) = m), \quad x \in \mathcal{X}, m \in \mathcal{M}. \quad (\text{B.8})$$

Since $\mathbf{W}(\cdot)$ is by assumption irreducible, $r(x, m) > 0$ for all x and m . Thus, the conditional probabilities on the right-hand side in Eq. (B.7) 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 $\mathbf{W}(\cdot)$: $\mathcal{W} \triangleq \mathcal{Q} \times \mathcal{M}_e \times \mathcal{X} \times \mathcal{M}_r$, and use $\mathbb{P}_{\phi^0, \psi}^\infty$ to denote probabilities under (ϕ^0, ψ) . Fix $x, x' \in \mathcal{X}$ and $m, m' \in \mathcal{M}$. We have that

$$\begin{aligned} & \mathbb{P}_{\phi^0, \psi}^\infty (X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m) \\ &= \mathbb{P}_{\phi^0, \psi}^\infty (X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m). \end{aligned} \quad (\text{B.9})$$

Eq. (B.9) can be derived as follows.

$$\begin{aligned} & \mathbb{P}_{\phi^0, \psi}^\infty (X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m) \\ & \stackrel{(a)}{=} \mathbb{P}_{\phi^0, \psi}^\infty (X(1) = x' \mid X(0) = x, M_r(0) = m) \\ & \quad \cdot \mathbb{P}_{\phi^0, \psi}^\infty (M_r(1) = m' \mid X(0) = x, M_r(0) = m) \\ & \stackrel{(b)}{=} \mathbb{P}_{\phi^0, \psi}^\infty (X(1) = x' \mid X(0) = x, M_r(0) = m) \\ & \quad \cdot \mathbb{P}_\psi (M_r(1) = m' \mid X(0) = x, M_r(0) = m) \\ & \stackrel{(c)}{=} \mathbb{P}_{\phi^0, \psi} (X(1) = x' \mid X(0) = x, M_r(0) = m) \\ & \quad \cdot \mathbb{P}_{\phi^0, \psi} (M_r(1) = m' \mid X(0) = x, M_r(0) = m) \\ & \stackrel{(d)}{=} \mathbb{P}_{\phi^0, \psi} (X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m). \end{aligned} \quad (\text{B.10})$$

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 $Z_r(0)$, 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 $\mathbf{W}(0)$ is initialized, or what encoding policy is used. Step (c) follows from the definition of the policy ϕ^0 in Eq. (B.7). Finally, step (d) is based on the same conditional independence property as that in Step (a) in Eq. (B.10), which holds under any encoding and allocation policies (Figure 5).

We proceed further to derive

$$\begin{aligned} & \mathbb{P}_{\phi^0, \psi} (X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m) \\ &= \mathbb{P}_{\phi^0, \psi}^\infty (X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m) \\ &= r(x, m)^{-1} \mathbb{P}_{\phi^0, \psi}^\infty (X(1) = x', M_r(1) = m', X(0) = x, M_r(0) = m) \\ &= r(x, m)^{-1} \sum_{\substack{\mathbf{w}^1, \mathbf{w}^0 \in \mathcal{W}: \\ x^1 = x', m_r^1 = m', x^0 = x, m_r^0 = m}} \mathbb{P}_{\phi^0, \psi}^\infty (\mathbf{W}(1) = \mathbf{w}^1, \mathbf{W}(0) = \mathbf{w}^0) \\ & \stackrel{(e)}{\geq} \sum_{\substack{\mathbf{w}^1, \mathbf{w}^0 \in \mathcal{W}: \\ x^1 = x', m_r^1 = m', x^0 = x, m_r^0 = m}} \mathbb{P}_{\phi^0, \psi}^\infty (\mathbf{W}(1) = \mathbf{w}^1, \mathbf{W}(0) = \mathbf{w}^0) \\ &= \sum_{\substack{\mathbf{w}^1, \mathbf{w}^0 \in \mathcal{W}: \\ x^1 = x', m_r^1 = m', x^0 = x, m_r^0 = m}} \mathbb{P}_{\phi^0, \psi} (\mathbf{W}(1) = \mathbf{w}^1 \mid \mathbf{W}(0) = \mathbf{w}^0) \\ & \quad \cdot \mathbb{P}_{\phi^0, \psi}^\infty (\mathbf{W}(0) = \mathbf{w}^0), \end{aligned} \quad (\text{B.11})$$

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

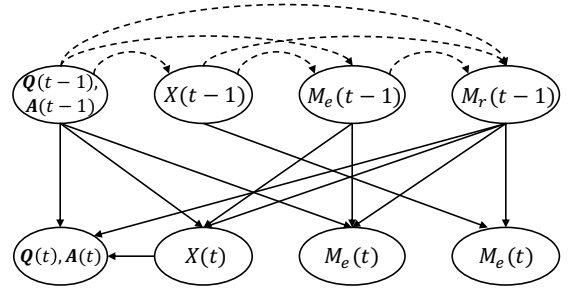


Figure 5: A Bayesian network representation of the evolution of $\mathbf{W}(\cdot)$. The directed lines represent dependence relations; for example, there is a directed line from $M_r(t-1)$ to $M_r(t)$, so $M_r(t)$ is dependent on $M_r(t-1)$. The set of dashed lines is a general representation of arbitrary dependence among $Q(t-1)$, $A(t-1)$, $X(t-1)$, $M_e(t-1)$ and $M_r(t-1)$, whereas solid lines indicate dependence relations from the policy update equations (2.6), (2.7) and (2.8).

The irreducibility of $\mathbf{W}(\cdot)$ implies that $\mathbb{P}_{\phi^0, \psi}^\infty (\mathbf{W}(0) = \mathbf{w}^0) > 0$ for all $\mathbf{w}^0 \in \mathcal{W}$. Eq. (B.11) thus shows that the one-step transition probability

$$\mathbb{P}_{\phi^0, \psi} (X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m) > 0 \quad (\text{B.12})$$

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

$$\mathbb{P}_{\phi^0, \psi} (\mathbf{W}(1) = \mathbf{w}^1 \mid \mathbf{W}(0) = \mathbf{w}^0) > 0. \quad (\text{B.13})$$

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. (B.11) to over multiple time slots, and conclude that, for any fixed $T > 0$,

$$\mathbb{P}_{\phi^0, \psi} (X(T) = x', M_r(T) = m' \mid X(0) = x, M_r(0) = m) > 0 \quad (\text{B.14})$$

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

$$\mathbb{P}_{\phi, \psi}(\mathbf{W}(T) = \mathbf{w}^1 \mid \mathbf{W}(0) = \mathbf{w}^0) > 0. \quad (\text{B.15})$$

Since the chain $\mathbf{W}(\cdot)$ 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 $(\bar{r}(x, m))_{x \in \mathcal{X}, m \in \mathcal{M}_r}$:

$$\bar{r}(x, m) = \mathbb{P}_{\phi^0, \psi}^{\infty}(X(0) = x, M_r(0) = m), \quad x \in \mathcal{X}, m \in \mathcal{M}_r. \quad (\text{B.16})$$

Here $\mathbb{P}_{\phi^0, \psi}^{\infty}$ 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

$$\begin{aligned} & \bar{r}(x', m') \\ = & \sum_{x \in \mathcal{X}, m \in \mathcal{M}_r} \mathbb{P}_{\phi^0, \psi}(X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m) \bar{r}(x, m). \end{aligned} \quad (\text{B.17})$$

By definition, we also have

$$\begin{aligned} & r(x', m') \\ = & \sum_{x \in \mathcal{X}, m \in \mathcal{M}_r} \mathbb{P}_{\phi, \psi}^{\infty}(X(1) = x', M_r(1) = m' \mid X(0) = x, M_r(0) = m) r(x, m) \end{aligned} \quad (\text{B.18})$$

Thus, by (B.9), $r(x', m')$ also satisfies the balance equation (B.17), and since $(\bar{r}(x, m))_{x, m}$ is unique, we must have

$$\bar{r}(x, m) = r(x, m), \quad \forall x \in \mathcal{X}, m \in \mathcal{M}_r. \quad (\text{B.19})$$

This implies that $\boldsymbol{\mu}(\phi^0, \psi) = \boldsymbol{\mu}(\phi, \psi, \boldsymbol{\lambda})$. By Lemma 6.3 and the positive recurrence of $\mathbf{W}(\cdot)$ under (ϕ, ψ) and $\boldsymbol{\lambda}$, We have $\boldsymbol{\mu}(\phi, \psi, \boldsymbol{\lambda}) \geq \boldsymbol{\lambda}$. Therefore, $\boldsymbol{\mu}(\phi^0, \psi) \geq \boldsymbol{\lambda}$ and this completes the proof of Proposition 9.2. \square

B.3 Proof of Theorem 7.1

We prove Theorem 7.1 in this subsection, and begin with the following technical result on the geometry of the set $\text{conv}(\Pi)$.

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

PROOF. Let $\boldsymbol{\mu} \in \text{conv}(\Pi)$ be such that $\mathbf{d}^{(0)} \leq \boldsymbol{\mu}$ and $\boldsymbol{\mu} \neq \mathbf{d}^{(0)}$. Then, $\boldsymbol{\mu}$ can be written as a convex combination of extreme points; i.e., $\boldsymbol{\mu} = p_1 \mathbf{d}^1 + \dots + p_r \mathbf{d}^r$ with $p_s > 0$ for all $s = 1, 2, \dots, r$, $\sum_s p_s = 1$, $\mathbf{d}^s \in \mathcal{E}$ for all s and the \mathbf{d}^s are all distinct. WLOG we may assume that $\mathbf{d}^s \neq \mathbf{d}^{(0)}$ for all s , since, for example, if $\mathbf{d}^1 = \mathbf{d}^{(0)}$, then $\frac{\mu - p_1 \mathbf{d}^1}{1 - p_1} = \frac{p_2}{1 - p_1} \mathbf{d}^2 + \dots + \frac{p_r}{1 - p_1} \mathbf{d}^r \in \text{conv}(\Pi)$, $\frac{\mu - p_1 \mathbf{d}^1}{1 - p_1} \geq \mathbf{d}^{(0)}$ and $\frac{\mu - p_1 \mathbf{d}^1}{1 - p_1} \neq \mathbf{d}^{(0)}$, so we can re-choose $\boldsymbol{\mu}$ to be $\frac{\mu - p_1 \mathbf{d}^1}{1 - p_1}$.

Consider the set $\Pi' = \{\mathbf{d}' \in \mathbb{Z}_+^N : \mathbf{d}' \leq \mathbf{d}^s \text{ for some } s = 1, 2, \dots, r\}$. By Assumption 1, $\Pi' \subseteq \Pi$. Since $\mathbf{d}^{(0)} \leq \boldsymbol{\mu} = p_1 \mathbf{d}^1 + \dots + p_r \mathbf{d}^r$, it is easy to see that we can find $p'_{\mathbf{d}'} \geq 0$ for $\mathbf{d}' \in \Pi'$

such that $\sum_{\mathbf{d}' \in \Pi'} p'_{\mathbf{d}'} = 1$ and $\sum_{\mathbf{d}' \in \Pi'} p'_{\mathbf{d}'} \mathbf{d}' = \mathbf{d}^{(0)}$. Now $\mathbf{d}^{(0)}$ is an extreme point, so we must have $p'_{\mathbf{d}'} > 0 \Rightarrow \mathbf{d}' = \mathbf{d}^{(0)}$. Consider any $\mathbf{d}' \in \Pi'$ such that $p'_{\mathbf{d}'} > 0$. By definition, $\mathbf{d}' \leq \mathbf{d}^s$ for some s . Thus, $\mathbf{d}^{(0)} = \mathbf{d}' \leq \mathbf{d}^s$. However, $\mathbf{d}^s \neq \mathbf{d}^{(0)}$ by assumption, which contradicts the maximality of $\mathbf{d}^{(0)}$ in \mathcal{E} . This establishes the lemma. \square

LEMMA B.2. *Consider a schedule set Π that satisfies Assumptions 1 and 3, and let \mathcal{E} be the set of extreme points of $\text{conv}(\Pi)$. Then, $\text{conv}(\Pi)$ has two distinct extreme points that are maximal.*

PROOF. 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 $\mathbf{d}^{(0)}$. Then $\mathbf{d}^{(0)}$ 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 $\mathbf{d}^{(0)}$ 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 B.1, are also maximal in $\text{conv}(\Pi)$. \square

LEMMA B.3. *Fix any finite $k, \ell \in \mathbb{Z}_+$, and $\varepsilon \in (0, 1)$. Consider some $\phi \in \Phi_k$ and $\psi \in \Psi_{\ell}$, an ε -majorizing channel $C = (X, \mathcal{Y}, C)$ that satisfies Assumption 4, and recall $\tilde{\Lambda}(\phi, \psi)$, the capacity region under the policy pair (ϕ, ψ) , defined in (2.12). For any $\boldsymbol{\lambda} \in \tilde{\Lambda}(\phi, \psi)$, let $\mathbb{P}_{\boldsymbol{\lambda}}(\cdot)$ denote the stationary probability of the chain $\mathbf{W}(\cdot)$. Then, there exists some $\delta > 0$ such that the following is true. For any $y \in \mathcal{Y}$*

$$\mathbb{P}_{\boldsymbol{\lambda}}(Y(t) = y, M_r(t) = m) \geq \delta. \quad (\text{B.20})$$

PROOF. Write $C = \varepsilon C^0 + (1 - \varepsilon) C^1$ as in Eq. (3.1), with C^0 having identical rows. Let $\delta' = \min\{C_{x, y}^0 : x \in \mathcal{X}, y \in \mathcal{Y}\}$. 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

$$\mathbb{P}_{\boldsymbol{\lambda}}(Y(t+1) = y', M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m'') > 0. \quad (\text{B.21})$$

We claim that there exists $\tilde{\delta} > 0$ such that whenever (B.21) holds, then

$$\mathbb{P}_{\boldsymbol{\lambda}}(Y(t+1) = y', M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m'') \geq \tilde{\delta}. \quad (\text{B.22})$$

To prove the claim, note that $M_r(t+1)$ is generated from $Y(t)$ and $M_r(t)$ by ψ alone (see Eq. (2.8)). Thus, $M_r(t+1)$ is conditionally independent from $Y(t+1)$, given $Y(t)$ and $M_r(t)$, and we can write

$$\begin{aligned} & \mathbb{P}_{\boldsymbol{\lambda}}(Y(t+1) = y', M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m'') \\ = & \mathbb{P}_{\boldsymbol{\lambda}}(Y(t+1) = y' \mid Y(t) = y'', M_r(t) = m'') \\ & \cdot \mathbb{P}_{\psi}(M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m''), \end{aligned} \quad (\text{B.23})$$

where the notation \mathbb{P}_{ψ} is used to emphasize the fact the conditional probability $\mathbb{P}_{\psi}(M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m'')$ depends only on the policy ψ , but not on $\boldsymbol{\lambda}$. Since $C = \varepsilon C^0 + (1 - \varepsilon) C^1$ and $\delta' = \min_{x, y} C_{x, y}^0 > 0$, $\mathbb{P}_{\boldsymbol{\lambda}}(Y(t+1) = y' \mid Y(t) = y'', M_r(t) = m'') \geq \varepsilon \delta'$. Thus,

$$\begin{aligned} & \mathbb{P}_{\boldsymbol{\lambda}}(Y(t+1) = y', M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m'') \\ \geq & \varepsilon \delta' \mathbb{P}_{\psi}(M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m''). \end{aligned} \quad (\text{B.24})$$

Also, by (B.21) and (B.23), we have

$$\mathbb{P}_\psi \left(M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m'' \right) > 0. \quad (\text{B.25})$$

The preceding conditional probability only depends on ψ , so we can find a uniform lower bound $\delta'' > 0$ with

$$\mathbb{P}_\psi \left(M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m'' \right) \geq \delta''. \quad (\text{B.26})$$

Therefore, we have

$$\mathbb{P}_\lambda \left(Y(t+1) = y', M_r(t+1) = m' \mid Y(t) = y'', M_r(t) = m'' \right) \geq \varepsilon \delta' \delta''. \quad (\text{B.27})$$

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 irreducibility, $\mathbb{P}_\lambda(Y(T) = y, M_r(T) = m \mid Y(1) = y_0, M_r(1) = m_0) > 0$ for some $T \geq 1$. Using Eq. (B.21) in the claim above, it is easy to show that

$$\mathbb{P}_\lambda(Y(T) = y, M_r(T) = m \mid Y(1) = y_0, M_r(1) = m_0) \geq \tilde{\delta}^T. \quad (\text{B.28})$$

Since the set $\mathcal{Y} \times \mathcal{M}_r$ is finite, there exists some T_{\max} such that

$$\mathbb{P}_\lambda(Y(T) = y, M_r(T) = m \mid Y(1) = y_0, M_r(1) = m_0) \geq \tilde{\delta}^{T_{\max}} \quad (\text{B.29})$$

uniformly over all $y_0, y \in \mathcal{Y}$ and $m_0, m \in \mathcal{M}_r$. Thus, we must have

$$\mathbb{P}_\lambda(Y(T) = y, M_r(T) = m) \geq \tilde{\delta}^{T_{\max}}, \quad (\text{B.30})$$

which establishes the lemma. \square

PROOF OF THEOREM 7.1. We prove Theorem 7.1 by contradiction. Towards this end, suppose that there exists $\varepsilon > 0$ and an ε -majorizing channel $C = (\mathcal{X}, \mathcal{Y}, C)$, such that $\rho_{k, \ell}^*(C) = 1$ for some $k \geq 0$ and $\ell \in \mathbb{Z}_+$.

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

$$\left(1 - \frac{\varepsilon'}{2}\right) \Lambda \subseteq \tilde{\Lambda}(\phi, \psi). \quad (\text{B.31})$$

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

$$\mathbb{P}_{\lambda^{(1)}}(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m) > 0. \quad (\text{B.32})$$

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

$$\begin{aligned} & \mathbb{P}_\psi(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m) \\ &= \mathbb{P}_\lambda(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m). \end{aligned} \quad (\text{B.33})$$

This implies that

$$\begin{aligned} & \mathbb{P}_\psi(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m) \\ &= \mathbb{P}_{\lambda^{(1)}}(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m) \\ &= \mathbb{P}_{\lambda^{(2)}}(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m). \end{aligned} \quad (\text{B.34})$$

Thus,

$$\begin{aligned} & \mathbb{P}_{\lambda^{(2)}}(\mathbf{D}(t) = \mathbf{d}^{(1)}) \\ & \geq \mathbb{P}_{\lambda^{(2)}}(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m) \\ & \quad \cdot \mathbb{P}_{\lambda^{(2)}}(Y(t) = y, M_r(t-1) = m) \\ & = \mathbb{P}_\psi(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m) \\ & \quad \cdot \mathbb{P}_{\lambda^{(2)}}(Y(t) = y \mid Y(t-1) = y, M_r(t-1) = m) \\ & \quad \cdot \mathbb{P}_{\lambda^{(2)}}(Y(t-1) = y, M_r(t-1) = m) \\ & \geq \mathbb{P}_\psi(\mathbf{D}(t) = \mathbf{d}^{(1)} \mid Y(t) = y, M_r(t-1) = m) \varepsilon \delta' \delta \triangleq \delta_1, \end{aligned} \quad (\text{B.35})$$

where the last inequality uses the facts that (a) for an ε -majorizing channel, $\mathbb{P}(Y(t) = y \mid Y(t-1) = y, M_r(t-1) = m) \geq \varepsilon \delta'$ for any $y \in \mathcal{Y}$ (recall the definition of δ' in the proof of Lemma B.3), independent of the conditioning event in the earlier time slot, and $\mathbb{P}_{\lambda^{(2)}}(Y(t-1) = y, M_r(t-1) = m) \geq \delta$ with δ in (B.20), by Lemma B.3.

Ineq. (B.35) implies that by choosing ε' sufficiently small, the policy pair (ϕ, ψ) cannot stabilize the system under the arrival rate vector $\lambda^{(2)}$. This contradicts the supposition that $\lambda^{(2)} \in \tilde{\Lambda}(\phi, \psi)$. Thus, for any $k', \ell \in \mathbb{Z}_+$, $\rho_{k', \ell}(C) < 1$.

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

B.4 Proof of Theorem 7.2

In this section, we prove Theorem 7.2. First, we proceed with the proof of the first statement. Let $\varepsilon \in (0, 1)$, and let $C = (\mathcal{X}, \mathcal{Y}, C)$ be an ε -majorizing channel. Let $C^+ = (\mathcal{X}, \mathcal{Y}^+, C^+)$ be the augmented channel, which we used at the beginning of the proof of Theorem 7.3. We will provide a constructive characterization of the constant $\rho(\varepsilon, \Pi)$, and similar to Theorem 7.3, it will be sufficient to show that

$$\rho_{0,0}^*(C^+) \leq \rho(\varepsilon, \Pi). \quad (\text{B.36})$$

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

$$\begin{aligned} p(\mathbf{d}) &= \mathbb{P}_{\psi^+}(\mathbf{D}(t) = \mathbf{d} \mid U(t) = 0) \\ &= \sum_y \mathbb{P}_{\psi^+}(\mathbf{D}(t) = \mathbf{d} \mid Y(t) = y) \mathbb{P}(Y(t) = y \mid U(t) = 0) \end{aligned} \quad (\text{B.37})$$

does not depend on λ , for any $\mathbf{d} \in \Pi$.

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 $\Gamma \subseteq (1 - \varepsilon)\text{conv}(\Pi) + \varepsilon\boldsymbol{\mu}^0$. We also let $\Gamma^+ = \{\lambda \in \mathbb{R}_+^N : \lambda \leq (1 - \varepsilon)\text{conv}(\Pi) + \varepsilon\boldsymbol{\mu}^0\}$. 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, $\rho^*(\phi, \psi^+, C^+) \leq \rho(\varepsilon, \boldsymbol{\mu}^0)$, where we recall the definition of $\rho^*(\phi, \psi^+, C^+)$ in Eq. (2.13).

We claim that $\rho(\varepsilon, \boldsymbol{\mu}^0) < 1$ for all $\varepsilon \in (0, 1)$ and $\boldsymbol{\mu}^0 \in \text{conv}(\Pi)$. 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 $\rho(\varepsilon, \boldsymbol{\mu}^0)\Lambda \not\subseteq \Gamma^+$. 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 $\boldsymbol{d}^{(1)} \leq (1 - \varepsilon)\boldsymbol{\mu}^1 + \varepsilon\boldsymbol{\mu}^0$. 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

$$\rho(\varepsilon, \Pi) \triangleq \sup_{\boldsymbol{\mu}^0 \in \text{conv}(\Pi)} \rho(\varepsilon, \boldsymbol{\mu}^0). \quad (\text{B.38})$$

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 $C = (\mathcal{X}, \mathcal{Y}, C)$. 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 $\rho^*(\phi, \psi, C) \geq \rho(\varepsilon, \boldsymbol{\mu}^0) = \rho(\varepsilon, \Pi)$. By the first statement, $\rho^*(\phi, \psi, C) \leq \rho(\varepsilon, \Pi)$. Thus, $\rho^*(\phi, \psi, C) = \rho(\varepsilon, \Pi)$, establishing the second statement of the theorem.

B.5 Proof of Lemma 9.3

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

$$\left(\rho_{k,\ell}^* - \frac{\varepsilon}{3}\right)\Lambda \subseteq \tilde{\Lambda}(\phi, \psi). \quad (\text{B.39})$$

By Ineq. (B.39) and Proposition 9.2, we have that for all $i \in [|\mathcal{F}|]$, there exists $\phi_i \in \Pi^S$, such that

$$\boldsymbol{\mu}(\phi_i, \psi) \geq \left(\rho_{k,\ell}^* - \frac{\varepsilon}{3}\right)\boldsymbol{\mu}_i^{\mathcal{F}}, \quad (\text{B.40})$$

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 $(\phi_i^{\delta(\varepsilon)}, \psi^{\delta(\varepsilon)})$, and have

$$\boldsymbol{\mu}\left(\phi_i^{\delta(\varepsilon)}, \psi^{\delta(\varepsilon)}\right) \geq \frac{\rho_{k,\ell}^* - 2\varepsilon/3}{\rho_{k,\ell}^* - \varepsilon/3} \cdot \boldsymbol{\mu}(\phi_i, \psi). \quad (\text{B.41})$$

Combining Ineqs. (B.40) and (B.41), we have

$$\boldsymbol{\mu}\left(\phi_i^{\delta(\varepsilon)}, \psi^{\delta(\varepsilon)}\right) \geq \left(\rho_{k,\ell}^* - \frac{2}{3}\varepsilon\right) \cdot \boldsymbol{\mu}(\phi_i, \psi). \quad (\text{B.42})$$

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. (B.42) implies that $(\rho_{k,\ell}^* - \varepsilon)\Lambda \subseteq \text{conv}^-(\boldsymbol{\mu}(\Phi^\varepsilon, \Psi^\varepsilon))$. \square

C INFINITE-MEMORY RECEIVER

In this section, we consider the case where the size of the receiver memory is infinite, i.e., $\ell = \infty$. The main result of this section is Theorem C.1 (Item 3 of Theorem 3.3), which states that $\rho_{0,\infty}^*(C) = 1$, for any informative channel 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 C.1. $\rho_{0,\infty}^*(C) = 1$.

The rest of this section is organized as follows. In Appendix C.1, we provide a precise description of the encoding-allocation policy pair – *episodic greedy learning* (EGL) – that will be used to prove Theorem C.1. Then, in Appendix C.2, we prove the main result proper. Finally, Appendix C.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.

C.1 Episodic Greedy Learning

To prove Theorem C.1, 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 when (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, $\rho_{\log N, \infty}^*(C) = 1$; see Appendix C.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 C is informative, which implies that we can find $x_1, x_2 \in \mathcal{X}$ and $y_1 \in \mathcal{Y}$ such that

$$q_1 \triangleq \mathbb{P}(Y(1) = y_1 \mid X(1) = x_1) < \mathbb{P}(Y(1) = y_1 \mid X(1) = x_2) \triangleq q_2. \quad (\text{C.1})$$

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 $i \equiv t \pmod N$, 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 $A(t-1)$ 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 C.3, when we discuss the case of no memory-feedback.

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

$$\text{proj}(\mathbf{x}, \mathcal{X}) \triangleq a\mathbf{x}, \quad (\text{C.2})$$

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

- (1) *Learning Phase.* The allocation policy generates an estimator for λ , denoted by $\widehat{\lambda}$, 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 $i = t \pmod N$, i.e.,

$$\widehat{p}_i \triangleq \frac{1}{B_1/N} \sum_{t:t \equiv i \pmod N} \mathbb{I}_{\{Y(t)=y_1\}}. \quad (\text{C.3})$$

Then, set

$$\widehat{\lambda}_i = \frac{q_2 - \widehat{p}_i}{q_2 - q_1}, \quad i = 1, \dots, N. \quad (\text{C.4})$$

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 $\widehat{\lambda}^+ \triangleq (\max\{0, \widehat{\lambda}_1\}, \dots, \max\{0, \widehat{\lambda}_N\})$. Consider the vector

$$\text{proj}(\widehat{\lambda}^+ + \alpha \mathbf{1}, \text{cl}(\Lambda)), \quad (\text{C.5})$$

i.e., the scaled projection of $\widehat{\lambda}^+ + \alpha \mathbf{1}$ 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 $\text{proj}(\widehat{\lambda}^+ + \alpha \mathbf{1}, \text{cl}(\Lambda))$ is on the outer boundary of $\text{cl}(\Lambda)$, there exists a random schedule D , distributed over the set Π of schedules, so that

$$\mathbb{E}[D \mid \widehat{\lambda}^+] = \text{proj}(\widehat{\lambda}^+ + \alpha \mathbf{1}, \text{cl}(\Lambda))^9. \quad (\text{C.6})$$

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

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 timeslots.

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

C.2 Proof of Theorem C.1

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 $\lambda' \in \Lambda$, $\lambda \leq K\mathbf{1}$.

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

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

$$\frac{\varepsilon}{N} \mathbf{1} = \varepsilon \sum_{i=1}^N \frac{1}{N} e^{(i)} \leq \varepsilon \text{conv}(\Pi).$$

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

By construction, the estimator \widehat{p}_i defined in (C.3) 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

$$\mathbb{P}\left(\|\widehat{\lambda}^+ - \lambda\|_\infty \leq \frac{\varepsilon}{3N}\right) \geq 1 - \frac{\varepsilon}{12KN}, \quad \forall B_1 \geq B_1^*. \quad (\text{C.7})$$

⁹Note that in Eq. (C.6), we have written the expectation of D as being conditioned on $\widehat{\lambda}^+$, which is itself random, and the distribution of D depends on the realization of $\widehat{\lambda}^+$.

For the rest of the proof, set

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

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

$$\widehat{\lambda}^+ + \alpha \mathbf{1} \in \Lambda, \quad \widehat{\lambda}^+ + \alpha \mathbf{1} \geq \lambda + \frac{\varepsilon}{3N} \mathbf{1}, \quad \text{and} \quad \text{proj}(\widehat{\lambda}^+ + \alpha \mathbf{1}, \text{cl}(\Lambda)) \geq \widehat{\lambda}^+ + \alpha \mathbf{1}. \quad (\text{C.9})$$

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

Let us now consider the randomized schedule D chosen during the deployment phase. By construction, $\mathbb{E}[D | \widehat{\lambda}^+] = \text{proj}(\widehat{\lambda}^+ + \alpha \mathbf{1}, \text{cl}(\Lambda))$. By Eqs. (C.7) and (C.9), we have that

$$\begin{aligned} & \mathbb{E}[D] \\ &= \mathbb{E}[\mathbb{E}[D | \widehat{\lambda}^+]] = \mathbb{E}[\text{proj}(\widehat{\lambda}^+ + \alpha \mathbf{1}, \text{cl}(\Lambda))] \\ &\geq \mathbb{P}\left(\text{proj}(\widehat{\lambda}^+ + \alpha \mathbf{1}, \text{cl}(\Lambda)) \geq \widehat{\lambda}^+ + \alpha \mathbf{1}, \widehat{\lambda}^+ + \alpha \mathbf{1} \geq \lambda + \frac{\varepsilon}{3N} \mathbf{1}\right) \\ &\quad \cdot \left(\lambda + \frac{\varepsilon}{3N} \mathbf{1}\right) \\ &\geq \mathbb{P}(E) \left(\lambda + \frac{\varepsilon}{3N} \mathbf{1}\right) \geq \left(1 - \frac{\varepsilon}{12KN}\right) \left(\lambda + \frac{\varepsilon}{3N} \mathbf{1}\right) \\ &\geq \lambda + \frac{\varepsilon}{3N} \mathbf{1} - \frac{\varepsilon}{12KN} \cdot K \mathbf{1} \\ &= \lambda + \frac{\varepsilon}{4N} \mathbf{1}, \end{aligned} \quad (\text{C.10})$$

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

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

$$\mathbb{E}(A[n]) = B\lambda, \quad \text{and} \quad \mathbb{E}[D[n]] = (B - B_1)\mathbb{E}[D]. \quad (\text{C.11})$$

Using (C.8), (C.10) and (C.11), we have

$$\begin{aligned} \mathbb{E}[D[n]] &\geq \left(1 - \frac{\varepsilon}{8KN}\right) B \left(\lambda + \frac{\varepsilon}{4N} \mathbf{1}\right) \\ &\geq B \left(\lambda + \frac{\varepsilon}{4N} \mathbf{1}\right) - \frac{\varepsilon}{8KN} \cdot BK \mathbf{1} \\ &= B \left(\lambda + \frac{\varepsilon}{8N} \mathbf{1}\right), \end{aligned} \quad (\text{C.12})$$

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

Similar to the proof of Theorem 9.5, we consider the Markov chain $W[\cdot]$ sampled at the beginnings of the episodes. Noting that $\{(A[n], D[n])\}$ are i.i.d. sequences, and using (C.11), (C.12), and Proposition 6.1, the sampled chain $W[\cdot]$ is positive recurrent under the EGL pair (ϕ, ψ) with parameters given in (C.8), whenever $\lambda \in (1 - \varepsilon)\Lambda$. This implies that $\rho^*(\phi, \psi, C) \geq 1 - \varepsilon$. But $\varepsilon \in (0, 1)$ is arbitrary, so we must have $\rho_{0, \infty}^*(C) = 1$. \square

C.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 $D(t)$ 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 C.2. *Under no memory-feedback, $\rho_{\log N, \infty}^*(C) = 1$.*

The proof of the theorem uses a simple, modified version of EGL, and is then essentially that of Theorem C.1 verbatim. Hence, we only provide a proof sketch here. Recall the EGL policy pair described in Appendix C.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 $t \equiv i \pmod{N}$, 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 C.1 can be used to show that under no memory-feedback, $\rho_{\log N, \infty}^*(C) = 1$. 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.

D FINDING THE CAPACITY FACTOR AND OPTIMAL POLICIES

We demonstrate in this section how to calculate the capacity factor $\rho_{\infty, \ell}^*(C)$ as the optimal value of a polynomial optimization problem over finite-dimensional matrices. Note that, by Theorem 3.3, this will cover all $\rho_{k, \ell}^*(C)$ for $k \geq K(\Pi, X)$, and serve as an upper bound on $\rho_{k, \ell}^*$ for $k < K(\Pi, X)$. 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 9.3, Section 9.2) to achieve the capacity factor.

We first consider the case where $\ell \geq 1$ (i.e., the receiver is not memoryless). Fix $\ell \in \mathbb{N}$. The chain $\{(X(t), M_r(t), D(t))\}_{t \in \mathbb{N}}$ is Markov with the following transition dynamics. Let a simple encoding policy (Definition 9.1) be parameterized by the $(|M_r| |\mathcal{X}|) \times |\mathcal{X}|$ row-stochastic matrix, G^E , with

$$G_{(m, x), x'}^E = \mathbb{P}(X(t+1) = x' | M_r(t) = m, X(t) = x). \quad (\text{D.1})$$

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

$$G_{(m,y),m'}^A = \mathbb{P}(M_r(t+1) = m' \mid M_r(t) = m, Y(t) = y), \quad (\text{D.2})$$

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

$$H_{(m,y),\mathbf{d}}^A = \mathbb{P}(D(t) = \mathbf{d} \mid M_r(t) = m, Y(t) = y). \quad (\text{D.3})$$

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

$$\begin{aligned} G_{(m,x,\mathbf{d}), (m',x',\mathbf{d}') }^S \\ = \mathbb{P}(X(t+1) = x', M_r(t+1) = m', D(t+1) = \mathbf{d}' \\ \mid M_r(t) = m, X(t) = x, D(t) = \mathbf{d}). \end{aligned} \quad (\text{D.4})$$

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

$$G_{(m,x,\mathbf{d}), (m',x',\mathbf{d}') }^S = G_{(m,x),x'}^E \left(\sum_{y \in \mathcal{Y}} C_{x',y} G_{(m,y),m'}^A H_{(m,y),\mathbf{d}}^A \right), \quad (\text{D.5})$$

where C is the channel matrix, with $C_{x,y} = \mathbb{P}(Y(t) = y \mid X(t) = x)$. Note that G^E , G^A 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

$$\mathbf{p} = (I - G^S)^{-1} \mathbf{1}. \quad (\text{D.6})$$

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

$$\widehat{\mathbf{p}}_{\mathbf{d}} = \sum_{m \in \mathcal{M}_r, x \in \mathcal{X}} \mathbf{p}_{(m,x,\mathbf{d})}^S, \quad \mathbf{d} \in \Pi. \quad (\text{D.7})$$

The resulting stationary service rate is given by

$$\boldsymbol{\mu}(G^E, (G^A, H^A)) = \sum_{\mathbf{d} \in \Pi} \widehat{\mathbf{p}}_{\mathbf{d}} \cdot \mathbf{d}. \quad (\text{D.8})$$

Using the above construction, we now formulate the optimization problem that will lead to the capacity factor. Recall from Lemma 9.3 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 $\rho_{\infty,\ell}^*(C)$ as follows. Let \mathcal{E}' be the set of maximal schedules in \mathcal{E} , where $\mathcal{E}' = \{\mathbf{d}^{(i)}\}_{i=1,\dots,|\mathcal{E}'|}$. Consider the following polynomial optimization problem:

$$\begin{aligned} & \text{maximize } \rho \\ & \text{subject to } \boldsymbol{\mu}(G^E(i), (G^A, H^A)) \geq \rho \mathbf{d}^{(i)}, \\ & \quad i = 1, \dots, |\mathcal{E}'|, \end{aligned} \quad (\text{D.9})$$

where the variables to be optimized are the row-stochastic matrices G^A , H^A , and $\{G^E(i)\}_{i=1,\dots,|\mathcal{E}'|}$. Denote by $\bar{\rho}$ and $\{\bar{G}^E(i)\}_{i=1,\dots,|\mathcal{E}'|}$,

(\bar{G}^A, \bar{H}^A) the optimal value and an optimal solution of (D.9), respectively. We have that the optimal value corresponds to the capacity factor:

$$\rho_{\infty,\ell}^*(C) = \bar{\rho}. \quad (\text{D.10})$$

Furthermore, to construct the Episodic Max Weight policy that achieves the capacity factor, the allocation policy (ψ^ϵ in Lemma 9.3) and the set of simple encoding policies (Φ^ϵ in Lemma 9.3) are given by those associated with (\bar{G}^A, \bar{H}^A) and $\{\bar{G}^E(i)\}_{i=1,\dots,|\mathcal{E}'|}$, respectively.

Special case of memoryless receiver. When $\ell = 0$, the optimization problem (D.9) can be further simplified. Recall the notation of rate allocation matrix M in (6.6), and schedule matrix S in (6.7). The capacity factor $\rho_{\infty,0}^*(C)$ is the optimal value of the following optimization problem:

$$\begin{aligned} & \text{maximize } \rho \\ & \text{subject to } \mathbf{r}^{(i)} C M S \geq \rho \mathbf{d}^{(i)}, \quad i = 1, \dots, |\mathcal{E}'|, \\ & \quad \mathbf{r}^{(i)} \geq 0, \quad \sum_{x \in \mathcal{X}} \mathbf{r}_x^{(i)} = 1, \quad i = 1, \dots, |\mathcal{E}'|, \end{aligned} \quad (\text{D.11})$$

where the variables to be optimized are the probability vectors $\{\mathbf{r}^{(i)}\}_{i=1,\dots,|\mathcal{E}'|}$ and row-stochastic matrix M . 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 M^* in an optimal solution of (D.11).

Notably, enabled by our theoretical results, the above derivations show that $\rho_{\infty,\ell}^*(C)$ can be calculated by solving polynomial optimization problems over *finite-dimensional* matrices. In contrast, without Theorem 3.3, it was not at all clear *a priori* how to compute $\rho_{\infty,\ell}^*(C)$, 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 (D.9) and (D.11) still have their drawbacks: they may be non-convex and could scale poorly as the size of the receiver memory, $|\ell|$, or that of the maximal schedules, $|\mathcal{E}'|$, becomes large. It would be an interesting future direction to investigate how these optimization problems can be solved efficiently.