

# The Classical Capacity of Additive Quantum Queue-Channels

Prabha Mandayam, *Member, IEEE*, Krishna Jagannathan, *Member, IEEE*, and Avhishek Chatterjee

**Abstract**—We consider a setting where a stream of qubits is processed sequentially. We derive fundamental limits on the rate at which classical information can be transmitted using qubits that decohere as they wait to be processed. Specifically, we model the sequential processing of qubits using a single server queue, and derive expressions for the classical capacity of such a quantum ‘queue-channel.’ Focusing on two important noise models, namely the *erasure channel* and the *depolarizing channel*, we obtain explicit single-letter capacity formulas in terms of the stationary waiting time of qubits in the queue. Our capacity proof also implies that a ‘classical’ coding/decoding strategy is optimal, i.e., an encoder which uses only orthogonal product states, and a decoder which measures in a fixed product basis, are sufficient to achieve the classical capacity of both queue-channels. Our proof technique for the converse theorem generalizes readily — in particular, whenever the underlying quantum noise channel is *additive*, we can obtain a single-letter upper bound on the classical capacity of the corresponding quantum queue-channel. More broadly, our work begins to quantitatively address the impact of decoherence on the performance limits of quantum information processing systems.

## I. INTRODUCTION

Unlike classical bits, quantum bits (or qubits) undergo rapid *decoherence* in time, due to certain unavoidable physical interactions with their environment [3]. Information stored in a qubit may be completely or partially lost as the qubit decoheres. The nature and the rate of the noise process depends on the particular physical implementation of the quantum state, as well as other factors such as the environment and temperature. For example, superconducting Josephson junction based qubits have average coherence times that are typically of the order of a few tens of microseconds [4, Table 2]. Decoherence or noise poses a major challenge to the scalability of quantum information processing systems — therefore, it is imperative to obtain a quantitative understanding of the impact of decoherence on the performance limits of quantum information processing systems.

In this paper, we consider a setting where a stream of qubits is processed *sequentially*. There are several possible scenarios where qubits may wait to be processed. For example,

in the context of quantum computation, subsets of qubits are often ‘idling’ in a quantum circuit, waiting their turn while other qubits are undergoing gate operations. Such idle qubits undergo decoherence as they wait to get processed, and the resulting errors are modelled as ‘storage errors’ in the context of quantum fault tolerance [5].

In the context of quantum communication, such a sequential processing of qubits arises naturally when we study consider the problem of entanglement distribution over quantum networks [6]. In this case, classical information is encoded in pairs of entangled photons which are transmitted over lossy fibre. To sustain the entanglement over long distances, the photons undergo an entanglement swap operation at quantum repeaters that are placed at intermediate distances. Every photon that arrives at a repeater has to wait for its partner to arrive in order to be processed, in which time they may get erased or depolarized, depending on the nature of communication channel as the quantum memory used to realise the quantum repeater [7]. We derive fundamental limits on the rate at which classical information can be transmitted using qubits that *decohere as they wait to be processed*.

To be more precise, we model the sequential processing of a stream of qubits using a *single server queue* [8]. The qubits arrive to be processed at a ‘server’ according to some stationary point process. For example, in the context of quantum communication, qubits are generated by optical sources which have an inherent randomness due to the underlying physical processes. The commonly used heralded single-photon sources, for example, rely on a nonlinear physical interaction which is probabilistic in nature [9]. The server processes the qubits at a fixed average rate. The qubits undergo decoherence (leading to errors) as they wait to be processed, and the probability of error/erasure of each qubit is modeled as a function of the time spent in the queue by *that* qubit. After the processing completes, the qubits are measured and interpreted as classical bits. We call this system a ‘quantum queue-channel’ and characterise the classical capacity of such a quantum queue-channel.

### A. Related Work

An information theoretic notion of reliability of a queuing system with state-dependent errors was introduced and studied in [10], where the authors considered queue-length dependent errors motivated mainly by human computation and crowd-sourcing. The classic paper of Anantharam and Verdú considered timing channels where information is encoded in the times between consecutive information packets, and these

P. Mandayam is with the Department of Physics, IIT Madras, Chennai, India (e-mail: prabhamd@iitm.ac.in).

K. Jagannathan and A. Chatterjee are with the Department of Electrical Engineering, IIT Madras, Chennai, India (e-mail: {krishnaj,avhishek}@ee.iitm.ac.in). AC acknowledges the Department of Science and Technology, Govt. of India for its support through SERB/SRG/2019/001809 and INSPIRE/04/2016/001171.

The material in this paper was presented in part at the 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) [1], and the National Conference on Communications 2019 [2].

packets are subsequently processed according to some queueing discipline [11]. Due to randomness in the sojourn times of packets through servers, the encoded timing information is distorted, which the receiver must decode. In contrast to [11], *we are not concerned with information encoded in the timing between packets — in our work, all the information is in the qubits.*

In a recent paper [2], we considered the queue-channel problem described above, and derived the capacity of an erasure queue-channel under certain technically restrictive conditions. Specifically, [2] restricts the encoder to using only orthogonal product states, and the decoder measures in a *fixed* product basis. In this restricted setting, qubits are essentially made to behave like ‘classical bits that decohere,’ and the underlying quantum channel effectively simulates a classical channel known as the *induced classical* channel. In general, the classical capacity of the underlying quantum channel could be *larger* than the capacity of the induced classical channel, because the former allows for entangled channel uses and more general (joint) measurements at the decoder.

As an aside, we remark that the erasure queue-channel treated in [2] can be used to model a multimedia-streaming scenario, where information packets become useless (erased) after a certain time.

## B. Our Contributions

In this paper, we completely characterise the classical capacity of quantum queue-channel for two important noise models, namely the *erasure channel* and the *depolarizing channel*. Specifically, we allow for possibly entangled channel uses by the encoder, and arbitrary measurements at the decoder.

Obtaining the queue-channel capacity for an arbitrary noise model involves overcoming some technical challenges. First, the quantum queue-channel is non-stationary. Second, the erasure events corresponding to consecutive qubits are correlated through their waiting times, which are in turn governed by the queuing process. This leads to memory across consecutive channel uses.

Interestingly, we note that the capacity result in [2, Theorem 1] for the induced classical channel readily offers an ‘achievable rate’ for the quantum erasure queue-channel — after all, any rate that is achievable with the restrictions in [2] can be achieved without those restrictions. Much of the technical challenge therefore lies in proving a ‘converse theorem,’ i.e., in showing a capacity upper bound that matches the expression in [2, Theorem 1]. The key contributions in this paper can be summarised as follows:

1) *Upper bound on the capacity of additive quantum queue-channels:* We first show that whenever the underlying quantum noise channel is *additive*, we can obtain a single-letter upper bound on the classical capacity of the corresponding quantum queue-channel. Our upper bound proof proceeds via the following key steps. The first step involves showing a certain conditional independence of  $n$  consecutive channel uses, *conditioned* on the sequence of qubit waiting times  $(W_1, W_2, \dots, W_n)$ . Specifically, we show that the  $n$ -qubit queue-channel factors into a *tensor product* of single-use

channels, for any given sequence of the waiting times. Next, we use a general capacity upper bound proved in [12, Lemma 5] for non-stationary channels, which we then simplify using the conditional independence result from the first step.

2) *The Quantum Erasure Queue-Channel:* The general upper bound proved for any additive quantum queue-channel, along with the celebrated *additivity* result of Holevo [13] for the quantum erasure channel immediately gives us a single-letter capacity upper bound. For the erasure queue-channel, we show that the upper bound is indeed the *same* as the capacity expression derived in [2, Theorem 1] for the induced classical channel. That is, the capacity of the quantum erasure queue-channel does not increase by allowing for entangled channel uses by the encoder, and arbitrary measurements at the decoder. In other words, the classical coding/decoding strategy in [2] (proposed for the induced classical channel), is indeed sufficient to realise the classical capacity of the underlying quantum erasure queue-channel. Furthermore, we show that the capacity remains the same, regardless of whether the arrival and departure times of the qubits (and hence their waiting times) is available at the decoder.

In hindsight, this result may not be altogether surprising, considering that the classical capacity of the *memoryless* quantum erasure channel is the same as the capacity of the classical erasure channel. In other words, a classical coding strategy is sufficient to realise the classical capacity of the (memoryless) quantum erasure channel — see [14].

3) *The Depolarising Queue-Channel:* When the underlying noise model is a depolarising channel, we obtain the capacity expression for the queue channel, assuming that the decoder has access to the sequence of arrival and departure times (and hence the waiting times) of the qubits. The upper bound proof is technically similar to the erasure case, except that we appeal to the additivity result for the depolarising channel, due to King [15]. For the achievability part, we consider the corresponding induced classical channel (which turns out to be a binary symmetric queue-channel) and use a classical coding/decoding strategy to obtain an achievable rate which matches the upper bound. When the waiting times are not available at the decoder, the achievable rate that we obtain is strictly smaller than the upper bound enforced by additivity arguments — in such a case, we can only identify the interval in which the capacity lies.

## C. Organisation

The remainder of this paper is organised as follows. Section II details the system model and preliminaries. Section III presents the relevant channel capacity definitions and proceeds to present our main technical results. Section IV deals with the induced classical channels and their capacities — the achievability part of our capacity results also come from this section. Section V concludes the paper and proposes directions for future work.

## II. SYSTEM MODEL AND PRELIMINARIES

The model we study is similar to the one considered in [2]. Specifically, a source generates a classical bit stream, which is

encoded into qubits. These qubits arrive at a continuous-time single-server queue according to a stationary point process of rate  $\lambda$ . To be more explicit, the single-server queue is characterised by (i) A server that processes the qubits in the order in which they arrive, i.e., in a First Come First Served (FCFS) fashion<sup>1</sup>, and (ii) An ‘unlimited buffer’ — that is, there is no limit on the number of qubits that can queue up as they wait to be processed. We assume that the service times for the qubits are independent and identically distributed (i.i.d.) random variables. The service time of the  $j$ th qubit is denoted by  $S_j$  and has a cumulative distribution  $F_S$ . The average service rate of each qubit is  $\mu$ , i.e.,  $\mathbf{E}_{F_S}[S] = 1/\mu$ . For stability of the queue, we assume  $\lambda < \mu$ . For ease of notation let us assume  $\mu = 1$ . (Our results easily extend to general  $\mu$ ). Let  $A_j$  and  $D_j$  be the arrival and the departure epochs of  $j$ th qubit, respectively and  $W_j = D_j - A_j$  be the total time that  $j$ th qubit spends in the queue. Adopting a standard convention, we use upper case letters to denote random variables, and the corresponding lower case letters to denote the realized values of the random variable. For example,  $W_j$  is the random variable that denotes the total time spent in the queue by the  $j$ th qubit, while  $w_j$  denotes a specific realization of  $W_j$ .

The probability that the  $j$ th qubit undergoes erasure/error is modelled as an explicit function of its waiting time  $w_j$ , as we specify in the next subsection. After getting processed at the server, the qubits are measured and interpreted as classical bits (see Fig. 1).

*Remark 1:* The main technical results in this paper hold as long as the queueing system described above is both *stationary* and *ergodic*. Stationarity essentially means that the queue eventually reaches a ‘steady-state’ behaviour. In particular it implies that random variables of interest (such as waiting time of a qubit, queue length etc.) converge in distribution to a corresponding stationary distribution. Ergodicity means that the long-term *time average* of waiting times, queue lengths, or functions thereof, converge (almost surely) to the corresponding *ensemble average* (or expectation), taken over the stationary distribution. Most stable queueing models of interest (including the M/GI/1 queue and GI/GI/1 queue with non-arithmetic inter-arrival/service distributions) satisfy these two properties.

We further remark that while we consider the queueing systems that are stationary, the queue-channel is non-stationary, because we start with an *empty* queue, and send a *finite* length codeword. Each symbol then undergoes waiting time dependent errors, which corresponds to a non-stationary channel.

### A. Noise Model

As the qubits wait to be served, they undergo decoherence, leading to errors at the receiver. This decoherence is modeled mathematically as a completely positive trace preserving (CPTP) map on the the qubit states [3].

Let the finite set  $\mathcal{X}$  denote the input alphabet. In general, a given input symbol  $X_j \in \mathcal{X}$ , is encoded as a positive trace-class operator (called the *density operator*)  $\rho_j$  on a Hilbert

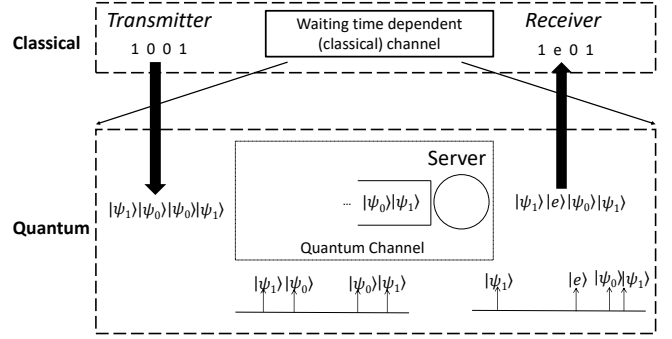


Fig. 1. Schematic of the queue-channel depicting the case of quantum erasure.

space  $\mathcal{H}^A$  of dimension  $|\mathcal{X}|$ . The noise process is denoted as the map  $\mathcal{N} : \mathcal{S}(\mathcal{H}^A) \rightarrow \mathcal{S}(\mathcal{H}^B)$ , where  $\mathcal{S}(\mathcal{H}^A)$  ( $\mathcal{S}(\mathcal{H}^B)$ ) denotes the set of density operators on the *input* Hilbert space  $\mathcal{H}^A$  (*output* Hilbert space  $\mathcal{H}^B$ ). The map  $\mathcal{N}$  is often referred to as the (quantum) *noise channel*.

The probability that a given state  $\rho_j$  undergoes decoherence is modelled as a function of the waiting time  $W_j$ , and the noise channel is accordingly parameterized in terms of the waiting time as  $\mathcal{N}_{W_j}$ . The noisy output state after the action of the map  $\mathcal{N}_{W_j}$  is denoted as  $\sigma_j \equiv \mathcal{N}_{W_j}(\rho_j)$ . This noisy state is measured by the receiver by performing a general quantum measurement and decoded as the output symbol  $Y_j \in \mathcal{Y}$ .

An  $n$ -length transmission over the above channel is denoted as follows. Inputs are drawn from the set  $\mathcal{X}^{(n)}$  of length  $n$  symbols  $\{X^n = (X_1 X_2 \dots X_n)\}$ , and encoded into density operators  $\rho_{X^n} \in \mathcal{S}((\mathcal{H}^A)^{\otimes n})$  on an  $n$ -fold tensor product of the input Hilbert space  $\mathcal{H}^A$ . The  $n$ -qubit channel is denoted  $\mathcal{N}_{W^n}^{(n)}$  and parameterized by the sequence of waiting times  $W^n = (W_1, W_2, \dots, W_n)$ . Note that the encoded state  $\rho_{X^n}$  could be entangled across multiple channel uses. Furthermore, in general, the  $n$ -qubit queue-channel is not a stationary, memoryless channel and does not automatically factor into an  $n$ -fold tensor product of single qubit channels. We refer to the sequence of  $n$ -qubit channels  $\vec{\mathcal{N}}_{\vec{W}} = \{\mathcal{N}_{W^{(n)}}^{(n)}\}_{n=1}^{\infty}$ , which are parameterised by the corresponding waiting time sequences  $\{W^{(n)}\}_{n=1}^{\infty}$ , as a *quantum queue-channel*, and characterize the classical capacity of this system (in bits/sec).

In this paper, we model the erasure/error probability of a qubit as an explicit function of its time spent in the queue. Thus, for the  $j$ th state  $\rho_j$  with waiting time  $w_j$ , we denote by  $p(w_j)$  the probability its error/erasure, where  $p : [0, \infty) \rightarrow [0, 1]$  is typically increasing. We pay special attention to two important noise models:

- (i) The quantum erasure channel [16] which acts on the  $j$ th state  $\rho_j$  with waiting time  $w_j$  as follows:  $\rho_j$  remains unaffected with probability  $1 - p(w_j)$ , and is erased to a (fixed) erasure state  $|e\rangle\langle e|$  with probability  $p(w_j)$ .
- (ii) The quantum depolarizing channel, which acts on a given qubit  $\rho_j$  as,  $\rho_j \rightarrow (1 - p(w_j))\rho_j + p(w_j)\frac{I}{2}$ , where  $I/2$  is the maximally mixed state [3].

In several physical scenarios, the decoherence time of a single qubit may be modelled as an exponential random variable.

<sup>1</sup>The FCFS assumption is not required for our results to hold, but it helps the exposition.

Thus, the probability of a qubit erasure/error after waiting for a time  $w$  is given by  $p(w) = 1 - e^{-\kappa w}$ , where  $1/\kappa$  is a characteristic time constant of the physical system under consideration [3, Section 8.3].

### III. QUANTUM QUEUE-CHANNEL CAPACITY

We are interested in defining and computing the information capacity of the quantum queue-channel. We follow the conventions and the definitions from [11] that were adopted in [10] for defining the capacity of queue-channels.

#### A. Definitions

Let  $M$  be the message transmitted from a set  $\mathcal{M}$  and  $\hat{M} \in \mathcal{M}$  be its estimate at the receiver. We now define an  $(n, R, \epsilon, T)$  code for classical communication over a quantum-queue channel  $\vec{\mathcal{N}}_{\vec{W}}$ . Let  $A^n, D^n$  denote the arrival epoch and departure epoch sequences, respectively.

*Definition 1:* An  $(n, R, \epsilon, T)$  quantum code is characterized by an encoding function  $X^n = f(M)$ , leading to an encoded  $n$ -qubit quantum state  $\rho_{X^n}$  corresponding to message  $M$ , and a decoding function

$$\hat{M} = g(\Lambda, \mathcal{N}^{(n)}(\rho_{X^n}), A^n, D^n)$$

corresponding to a measurement  $\Lambda$  at the receiver's end, where the cardinality of the message set  $|\mathcal{M}| = 2^{nR}$ , the expected total time for all the symbols of any codeword to reach the receiver is less than  $T$ , and the average probability of error of the code is less than  $\epsilon$ .

The measurement  $\Lambda$  at the decoder is a positive operator valued measure (POVM) with operator elements  $\{\Lambda_{\hat{M}}\}$  corresponding to message  $M$ . The probability of error corresponding to message  $M$  is

$$p_e(M) = \text{Tr} \left[ (I - \Lambda_{\hat{M}}) \mathcal{N}_{W^{(n)}}^{(n)}(\rho_{X^n}) \right],$$

and the average probability of error is given by  $\frac{1}{|\mathcal{M}|} \sum_{M \in \mathcal{M}} p_e(M)$ .

As the average probability of error at the decoder for an  $(n, R, \epsilon, T)$  quantum code is less than  $\epsilon$ , the code is called an  $\epsilon$ -achievable code with rate  $\frac{R}{T}$ .

*Definition 2:* A rate  $\bar{R}$  is said to be achievable if for any  $\epsilon \in (0, 1)$ , there exists an  $\epsilon$ -achievable code with rate  $\bar{R}$ , or equivalently, if for any  $\epsilon \in (0, 1)$  there exists an  $(n, R, \epsilon, T)$  code with  $\bar{R} = \frac{R}{T}$ .

*Definition 3:* The information capacity of the queue-channel is the supremum of all achievable rates for a given arrival and service process, and is denoted by  $C$  bits per unit time.

Note that the information capacity of the queue-channel depends on the arrival process, the service process, and the noise model. We assume that the receiver knows the realizations of the arrival and the departure times of each symbol, although we point out some results that hold even without this assumption.

There are several physical scenarios where the arrival and departure times of each qubit is known, the simplest example

being that of heralded single-photon sources which are ubiquitous in quantum communication networks (see [9] for example). The source typically produces a pair of entangled photons, one of which is used as a reference by the encoder/sender and the other photon is used in the actual communication protocol. The reference photon provides information about the arrival time, whereas the departure time can be obtained from the time-stamp on the receiver's detector.

Even in scenarios where this assumption may not be practical or realistic, conditioning on the waiting times offers a useful conceptual handle towards tackling the problem. We also note that two key results in the paper—namely the general upper bound in Theorem 1, and the exact capacity of the erasure queue-channel in Theorem 2, hold irrespective of the receiver's knowledge of the waiting times.

#### B. Holevo Information and Additivity

Before we proceed, we briefly review some well established concepts from quantum Shannon theory, which will be useful in analyzing the queue-channel capacity. An important measure that characterizes the classical capacity of quantum channels is the Holevo information of a quantum channel, also referred to as the Holevo capacity of a quantum channel [17].

*Definition 4 (Holevo Information):* The Holevo Information of a quantum channel  $\mathcal{N}$  is defined as the entropy difference [16], [18],

$$\begin{aligned} \chi(\mathcal{N}) &:= \sup_{\{P_x, \rho_x\}} \chi(\{P_x, \mathcal{N}(\rho_x)\}) \\ &= \sup_{\{P_x, \rho_x\}} H \left( \sum_x P_x \mathcal{N}(\rho_x) \right) - \sum_x p_x H(\mathcal{N}(\rho_x)), \end{aligned} \quad (1)$$

where the supremum is taken over all input ensembles  $\{P_x, \rho_x\}$  and  $H(\rho) = -\text{Tr}[\rho \log \rho]$  is the von Neumann entropy associated with a density operator  $\rho$ .

The celebrated Holevo-Schumacher-Westmoreland theorem states that the classical capacity of a quantum channel, denoted as  $C(\mathcal{N})$ , is the *regularized* Holevo information across independent channel uses [19], [20]. In other words,

$$C(\mathcal{N}) = \lim_{n \rightarrow \infty} \frac{\chi(\mathcal{N}^{\otimes n})}{n},$$

where  $\mathcal{N}^{\otimes n}$  is the product channel corresponding to  $n$  independent channel uses. The Holevo information of the channel  $\mathcal{N}$  is said to be *additive* if it satisfies  $\chi(\mathcal{N} \otimes \mathcal{N}) = 2\chi(\mathcal{N})$ . It is easy to see that for quantum channels whose Holevo information is additive, the classical capacity of the channel is simply equal to the Holevo information. Additivity of the Holevo information further implies that the classical capacity of such channels is achievable via a *classical* encoding and decoding strategy – the optimal coding strategy does not require entangled inputs at the encoder or collective measurements at the decoder. Examples of channels for which the Holevo information is known to be additive include the quantum erasure channel [13] and the quantum depolarizing channel [15].

*Definition 5 (Additive quantum queue-channel):* A quantum queue-channel  $\mathcal{N}_{\vec{W}}$  is said to be additive if the Holevo information of the underlying single-use quantum channel  $\mathcal{N}$  is additive. Specifically, additivity of the Holevo information of the quantum channel  $\mathcal{N}$  implies

$$\chi(\mathcal{N}_{W_1} \otimes \mathcal{N}_{W_2}) = \chi(\mathcal{N}_{W_1}) + \chi(\mathcal{N}_{W_2}).$$

### C. Upper bound on the Capacity for Additive Quantum Queue-Channels

We first invoke the capacity formula obtained in [12], for the classical capacity of general quantum channels which are neither stationary nor memoryless.

*Proposition 1:* The capacity of the quantum queue-channel (in bits/sec) described in Sec. II is given by

$$C = \lambda \sup_{\{\vec{P}, \vec{\rho}\}} \mathbb{I}(\{\vec{P}, \vec{\rho}\}, \vec{\mathcal{N}}_{\vec{W}}), \quad (2)$$

where,  $\mathbb{I}(\{\vec{P}, \vec{\rho}\}, \vec{\mathcal{N}}_{\vec{W}})$  is the quantum spectral information rate originally defined in [12]. We have stated this definition in Appendix A for completeness. Here,  $\vec{P}$  is the totality of sequences  $\{P^n(X^n)\}_{n=1}^{\infty}$  of probability distributions (with finite support) over input sequences  $X^n$ , and  $\{\vec{\rho}\}$  denotes the sequences of states  $\{\rho_{X^n}\}$  corresponding to the encoding  $X^n \rightarrow \rho_{X^n}$ . Finally,  $\vec{\mathcal{N}}_{\vec{W}}$  denotes the sequence of channels  $\{\mathcal{N}_{W^{(n)}}^{(n)}\}_{n=1}^{\infty}$ , which are parameterised by the corresponding waiting time sequences  $\{W^{(n)}\}_{n=1}^{\infty}$ .

We are now ready to state and prove a general upper bound for the capacity of additive quantum queue-channels.

*Theorem 1:* For an additive quantum queue-channel  $\vec{\mathcal{N}}_{\vec{W}}$ , the capacity is bounded as,

$$C \leq \lambda \mathbf{E}_{\pi} [\chi(\mathcal{N}_W)] \text{ bits/sec.},$$

irrespective of the receiver's knowledge of the arrival and the departure times. Here  $\chi(\mathcal{N}_W)$  denotes the Holevo information of the single-use quantum channel corresponding to waiting time  $W$  and  $\pi$  is the stationary distribution of the waiting time in the queue.

*Proof Outline:* Obtaining the queue-channel capacity of a quantum channel poses certain technical challenges, since the error probabilities are correlated across different channel uses. In other words, the probability that the  $i^{\text{th}}$  qubit gets affected by an error is a function of its waiting time in the queue, which in turn depends on the waiting time of the previous  $(i-1)^{\text{th}}$  qubit and so on. Furthermore, the channel is non-stationary. However, for the queue-channel model considered here, the  $n$ -qubit queue-channel does factor into a *tensor product* of single-use channels, *conditioned* on the sequence of waiting times  $(W_1, W_2, \dots, W_n)$ . This is formally shown in Lemma 1 below.

In order to obtain an *upper bound on capacity*, we proceed via the following key steps. First, we invoke an upper bound proved in [12, Lemma 5], to bound the capacity as the limit inferior of the *Holevo information* of a sequence of quantum channels. Next, we use the tensor product form of the channel obtained in Lemma 1 in conjunction with the fact that the Holevo information of the channel  $\mathcal{N}$  is additive. Finally, we

invoke the ergodicity of the queue to obtain a single-letter expression for an upper bound on the queue-channel capacity.

*Lemma 1 (Conditional independence):* The  $n$ -qubit quantum queue-channel  $\mathcal{N}_{W^n}^{(n)}$  factors into a tensor product of single-use channels, conditioned on the waiting times  $(W_1, W_2, \dots, W_n)$ .

*Proof:* Consider a sequence of  $n$  qubits transmitted via a quantum queue-channel with associated waiting times  $W_1, W_2, \dots, W_n$ . The  $n$ -qubit channel may be described via  $n$ -fold tensor-product operators of the form  $\{B_1 \otimes B_2 \otimes \dots \otimes B_n\}$ , where each single-qubit operator  $B_i$  is one of a finite set of *noise operators*  $\{E_1, E_2, \dots, E_L\}$ , which characterize the quantum channel  $\mathcal{N}$ . Note that the operators  $E_i \in \mathbb{M}_{2 \times 2}$  which belong to the space of  $2 \times 2$  complex matrices are called the *Kraus operators* associated with the quantum channel  $\mathcal{N}$ .

There are  $L^n$  tensor-product operators of the form  $\{B_1 \otimes B_2 \otimes \dots \otimes B_n\}$ , occurring with probabilities  $q_{k_1 k_2 \dots k_n}(W_1, W_2, \dots, W_n)$ , with the indices  $k_i \in \{E_1, E_2, \dots, E_L\}$ , depending on which noise operator each  $B_i$  corresponds to. This  $n$ -fold channel is a non-iid, correlated quantum channel in general, since the joint distribution  $q_{k_1 k_2 \dots k_n}(W_1, W_2, \dots, W_n)$  does not factor into a product of the individual error probabilities for each qubit. However, conditioned on the the waiting time sequence  $W^{(n)} = (W_1, W_2, \dots, W_n)$ , the joint distribution does factor as  $q_{k_1 k_2 \dots k_n}(W_1, W_2, \dots, W_n) = \prod_i q_{k_i}(W_i)$ . Therefore, conditioned on the waiting times  $(W_1, W_2, \dots, W_n)$  we may represent the action of the  $n$ -qubit channel on any  $n$ -qubit state  $\rho_{12 \dots n}$  as shown in (3). In other words, conditioned on the waiting time sequence  $W^n$ , the  $n$ -qubit channel factors into an  $n$ -fold tensor product of the form  $\mathcal{N}_{W_1} \otimes \mathcal{N}_{W_2} \otimes \dots \otimes \mathcal{N}_{W_n}$ , as desired. ■

We are now ready to prove the upper bound on the capacity of an additive quantum queue-channel. We assume that the sequence of waiting times  $\vec{W}$  is available at the receiver.

*Proof:* (Theorem 1) We start with an upperbound on the quantum inf-information rate proved in [12, Lemma 5]:

$$\mathbb{I}(\{\vec{P}, \vec{\rho}\}, \vec{\mathcal{N}}_{\vec{W}}) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \chi(\{P^{(n)}, \rho_{X^n}\}, \mathcal{N}_{W^{(n)}}^{(n)}),$$

where,  $\chi(\{P^{(n)}, \rho_{X^n}\}, \mathcal{N}_{W^{(n)}}^{(n)})$  is the Holevo information of the ensemble  $\{P^{(n)}(X^n), \mathcal{N}_{W^{(n)}}^{(n)}(\rho_{X^n})\}$ . Consider now the Holevo information  $\chi(\mathcal{N}_{W^{(n)}}^{(n)})$  of the  $n$ -qubit quantum queue-channel, for a given sequence of waiting times  $W^{(n)}$ . Lemma 1 implies that,

$$\chi(\mathcal{N}_{W^n}^{(n)}) = \chi(\mathcal{N}_{W_1} \otimes \mathcal{N}_{W_2} \otimes \dots \otimes \mathcal{N}_{W_n}).$$

Furthermore, the fact that the queue-channel is additive implies for any  $n$ ,

$$\chi(\mathcal{N}_{W^n}^{(n)}) = \sum_{i=1}^n \chi(\mathcal{N}_{W_i}).$$

Rewriting this in terms of the Holevo information of the

$$\begin{aligned}
& \mathcal{N}_{W^{(n)}}^{(n)}(\rho_{12\dots n}) \\
&= \sum_{k_1, k_2, \dots, k_n} q_{k_1 k_2 \dots k_n}(W_1, W_2, \dots, W_n) B_{k_1} \otimes B_{k_2} \dots \otimes B_{k_n} (\rho_{12\dots n}) B_{k_1}^\dagger \otimes B_{k_2}^\dagger \dots \otimes B_{k_n}^\dagger \\
&= \sum_{k_1, k_2, \dots, k_n} q_{k_1}(W_1) q_{k_2}(W_2) \dots q_{k_n}(W_n) B_{k_1} \otimes B_{k_2} \dots \otimes B_{k_n} (\rho_{12\dots n}) B_{k_1}^\dagger \otimes B_{k_2}^\dagger \dots \otimes B_{k_n}^\dagger \\
&= (\mathcal{N}_{W_1} \otimes \mathcal{N}_{W_2} \dots \otimes \mathcal{N}_{W_n})(\rho_{12\dots n}). \tag{3}
\end{aligned}$$

encoding ensemble for the  $n$ -qubit channel, we get,

$$\begin{aligned}
& \sup_{\{P^n(X^n), \rho_{X^n}\}} \chi(\{P_{X^n}^n, \mathcal{N}_{W^{(n)}}^{(n)}(\rho_{X^n})\}) \\
&= \sum_{i=1}^n \sup_{\{P_i(X_i), \rho_i\}} \chi(\{P(X_i), \mathcal{N}_{W_i}(\rho_i)\}). \tag{4}
\end{aligned}$$

Combining the above sequence of steps, we thus get the following upper bound on the capacity of an additive quantum queue-channel:

$$\begin{aligned}
C &\stackrel{(a)}{=} \lambda \sup_{\{\vec{P}, \vec{\rho}\}} \mathbf{I}(\{\vec{P}, \vec{\rho}\}, \vec{\mathcal{N}}_{\vec{W}}) \\
&\stackrel{(b)}{\leq} \lambda \sup_{\{\vec{P}, \vec{\rho}\}} \liminf_{n \rightarrow \infty} \frac{1}{n} \chi(\{P^{(n)}, \rho_{X^n}\}, \mathcal{N}_{W^{(n)}}^{(n)}) \\
&\stackrel{(c)}{\leq} \lambda \liminf_{n \rightarrow \infty} \frac{1}{n} \sup_{\{P^{(n)}, \rho_{X^n}\}} \chi(\{P^{(n)}, \rho_{X^n}\}, \mathcal{N}_{W^{(n)}}^{(n)}) \\
&\stackrel{(d)}{=} \lambda \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \sup_{\{P(X_i), \rho_i\}} \chi(\{P(X_i), \mathcal{N}_{W_i}(\rho_i)\}) \\
&= \lambda \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \chi(\mathcal{N}_{W_i}) \\
&\stackrel{(e)}{=} \lambda \mathbf{E}_\pi [\chi(\mathcal{N}_W)] \text{ a.s.}
\end{aligned}$$

Here, (a) is simply the definition of the queue-channel capacity as stated in Eq. (2), and (b) is the upper bound from [12, Lemma 5]. This upper bound uses the quantum Neyman-Pearson Lemma and the monotonicity of the quantum relative entropy, in addition to invoking arguments from quantum hypothesis testing. The inequality (c) follows from elementary real analysis. The equality in (d) follows from the conditional independence of the  $n$ -use channel and the additivity of the quantum queue-channel (see Eq. (4)), and (e) follows from the ergodicity of the queue. ■

We note that our proof of the upper bound on classical queue-channel capacity assumes knowledge of the sequence of waiting times  $(W_1, W_2, \dots, W_n)$  at the receiver. This automatically implies an upper bound for the scenario where the receiver does not have knowledge of the waiting times.

#### D. Erasure Queue-Channels

Erasure channels are ubiquitous in classical as well as quantum information theory. Our model of an erasure queue-channel captures a quantum information system where qubits decohere with time into erased (or non-informative) quantum states.

A single-use quantum erasure channel, for a qubit with waiting time  $W_1 = w_1$  is characterized by a pair of noise operators (or *Kraus* operators), namely the operator  $E : \rho_1 \rightarrow |e\rangle\langle e|$  which maps any input density operator  $\rho_1$  to a fixed erasure state  $|e\rangle\langle e|$ , and the identity operator  $I : \rho_1 \rightarrow \rho_1$ . If the input state has a waiting time  $W_1 = w_1$ , the erasure and identity operations occur with probabilities  $q_E(w_1) = p(w_1)$  and  $q_I(w_1) = 1 - p(w_1)$  respectively. Thus, the final state after the action of the erasure queue-channel for a given input state  $\rho_1$ , with waiting time  $W_1 = w_1$  is

$$\mathcal{N}_{W_1}(\rho_1) = q_E(w_1) E \rho_1 E^\dagger + q_I(w_1) \rho_1.$$

We now evaluate the upper bound proved in Theorem 1 for the case of the quantum erasure queue-channel and show that is achievable using a classical coding strategy. This leads us to the following result on the classical capacity of the quantum erasure queue-channel.

*Theorem 2:* For the erasure queue-channel defined above, the capacity is given by  $C = \lambda \mathbf{E}_\pi [1 - p(W)]$  bits/sec, irrespective of the receiver's knowledge of the arrival and the departure times, where  $\pi$  is the stationary distribution of the waiting time in the queue.

*Proof Outline:* The upper bound simply follows from the upper bound proved in Theorem 1 above, since the Holevo capacity of the quantum erasure channel is additive [13]. The *achievability* proof follows by fixing a classical encoding and decoding strategy and showing that the capacity of the induced classical channel does indeed coincide with the upper bound.

*Proposition 2 (Upper bound on Capacity):* The capacity of the quantum erasure queue-channel satisfies  $C \leq \lambda \mathbf{E}_\pi [1 - p(W)]$  bits/sec.

*Proof:* In order to obtain an *upper bound on capacity*, we invoke the celebrated additivity result of Holevo [13] for the quantum erasure channel. Next, we use the fact that the Holevo information of a single use erasure channel corresponding to waiting time  $W_i$  is given by [14],

$$\chi(\mathcal{N}_{W_i}) \equiv \sup_{\{P(x), \rho_x\}} \chi(\{P_x, \mathcal{N}_{W_i}(\rho_x)\}) = 1 - p(W_i). \tag{5}$$

This, along with the upper bound proved in Theorem 1 above, results in the desired upper bound. ■

*Proposition 3 (Lower bound on Capacity (Achievability)):* The capacity of the quantum erasure queue-channel satisfies  $C \geq \lambda \mathbf{E}_\pi [1 - p(W)]$  bits/sec.

*Proof:* We prove the lower bound by producing a particular encoding/decoding strategy that achieves the said capacity expression. In particular, we employ a classical strategy in

which classical bits 0 and 1 are encoded into two fixed orthogonal states (say  $|\psi_0\rangle$  and  $|\psi_1\rangle$ ), and the decoder also measures in a fixed basis. The input codewords are unentangled across multiple channel uses and the decoder simply performs product measurements. In this setting, the qubits essentially behave as classical bits, and the quantum erasure channel simulates the *induced classical* channel. The capacity of the corresponding induced classical channel is equal to  $\lambda \mathbf{E}_\pi [1 - p(W)]$ , as shown in Theorem 7 in Sec. IV-B below. ■

We remark that the above capacity result does not depend on the specific functional form of  $p(\cdot)$ . Further, the capacity result holds for any stationary and ergodic queue — i.e., it does not assume any specific queueing model. If we assume the functional form  $p(W) = 1 - \exp(-\kappa W)$  for the erasure probability, the following corollary is immediate.

*Corollary 1:* When the decoherence time of each qubit is exponentially distributed, i.e.,  $p(W) = 1 - \exp(-\kappa W)$ , the erasure queue-channel capacity is given by  $\lambda \mathbf{E}_\pi [e^{-\kappa W}]$  bits/sec., where  $\pi$  is the stationary distribution of the waiting time in the queue.

#### E. M/M/1 and M/GI/1 examples

We remark that the capacity expression  $\lambda \mathbf{E}_\pi [e^{-\kappa W}]$  is simply  $\lambda$  times the Laplace transform of the stationary waiting time  $W$ , evaluated at  $\kappa$ , which is the rate of decoherence. We now derive closed form expressions for the capacity for the M/M/1 and M/GI/1 queues. An M/GI/1 queue is a well-studied model of a single server queue in which the arrival process is a Poisson process, the service times are independent of the arrival process, and generally distributed (but i.i.d across qubits). See [8, Chapter 5] for a detailed treatment of the M/GI/1 queue. The M/M/1 queue is a special case of an M/GI/1 queue, where the service times are exponentially distributed.

Using Pollaczek-Khinchin formula (see [8, Eq. (5.105)]) for an FCFS M/GI/1 queue, we can obtain a closed-form expression for the capacity and the optimal arrival rate.

*Theorem 3:* For an FCFS M/GI/1 erasure queue-channel (with  $p(W) = 1 - \exp(-\kappa W)$ ),

- (i) the capacity is given by  $\frac{\lambda(1-\lambda)}{1-\alpha\lambda}$  bits/sec, and
- (ii) the capacity is maximised at

$$\lambda_{M/GI/1} = \frac{1}{\alpha} (1 - \sqrt{1-\alpha}) = \frac{1}{1 + \sqrt{1-\alpha}},$$

where  $\alpha = \frac{1-\tilde{F}_S(\kappa)}{\kappa}$ , and  $\tilde{F}_S(u) = \int \exp(-ux) dF_S(x)$  is the Laplace transform of the service time distribution.

*Proof:* First, note that for the particular form of  $p(\cdot)$  considered here, the capacity expression in Theorem 2 simplifies to  $\lambda \mathbf{E}[\exp(-\kappa W)]$ .

Next, using the Pollaczek-Khinchin formula for the M/GI/1 queue (with  $\mu = 1$ ), we write

$$\mathbf{E}[\exp(-\kappa W)] = \frac{(1-\lambda)\kappa}{\kappa - \lambda(1 - \tilde{F}_S(\kappa))} = \frac{1-\lambda}{1-\alpha\lambda},$$

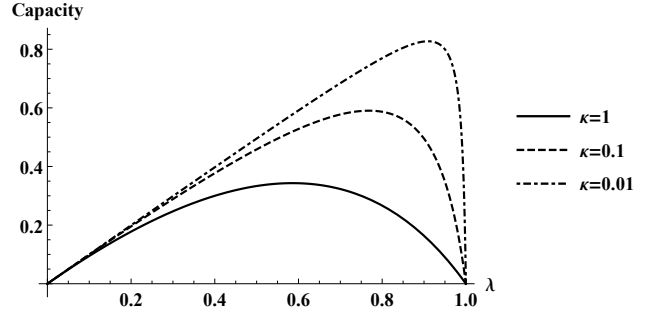


Fig. 2. The capacity of the M/M/1 erasure queue-channel (in bits/sec) plotted as a function of the arrival rate  $\lambda$  for different values of the decoherence parameter  $\kappa$ .

where  $\alpha = \frac{(1-\tilde{F}_S(\kappa))}{\kappa}$ . Thus, the capacity is  $\frac{\lambda(1-\lambda)}{1-\alpha\lambda}$  bits/sec, and the capacity maximising arrival rate is

$$\arg \max_{\lambda \in [0,1]} \frac{\lambda(1-\lambda)}{1-\alpha\lambda}.$$

The objective function in the above optimization problem is concave in  $\lambda$ . This implies that the value of  $\lambda$  that maximises the capacity is the one at which the derivative of the capacity with respect to  $\lambda$  is zero. Taking the derivative, we obtain a quadratic function in  $\lambda$  which when equated to zero yields two solutions for  $\lambda: \frac{1}{\alpha} \pm \frac{\sqrt{1-\alpha}}{\alpha}$ . The only valid solution for which  $\lambda \in [0, 1)$  is given by  $\frac{1}{\alpha} - \frac{\sqrt{1-\alpha}}{\alpha}$ . ■

This result offers interesting insights into the relation between the information capacity and the characteristic time-constant of the quantum states. Fig. 2 plots capacity versus arrival rate for an M/M/1 queue of unit service rate. We note that  $\kappa = 0.01$  corresponds to an average coherence time which is two orders of magnitude longer than the service time — a setting reminiscent of superconducting qubits [4]. We also notice from the shape of the capacity curve for  $\kappa = 0.01$  that there is a drastic drop in the capacity, if the system is operated beyond the optimal arrival rate  $\lambda_{M/M/1}$ . This is due to the drastic increase in delay induced decoherence as the arrival rate of qubits approaches the server capacity.

Theorem 3 characterizes an optimal  $\lambda$  for given arrival and service distributions. One can also ask after the best service distribution for a given arrival process and a fixed server rate. This question is of interest in designing the server characteristics like gate operations [21], or photon detectors for quantum communications. The following theorem is useful in such scenarios.

*Theorem 4:* For an erasure queue-channel with  $p(W) = 1 - \exp(-\kappa W)$  and Poisson arrivals at a given rate  $\lambda$ , the capacity is maximised by the service time distribution  $F_S(x) = \mathbf{1}(x \geq 1)$ , i.e., a deterministic service time maximises capacity, among all service distributions with unit mean and  $F_S(0) = 0$ .

*Proof:* As derived in the proof of Theorem 3, the capacity is

$$\frac{\lambda(1-\lambda)\kappa}{\kappa - \lambda(1 - \tilde{F}_S(\kappa))} = \frac{(1-\lambda)\kappa}{\frac{\kappa-\lambda}{\lambda} + \tilde{F}_S(\kappa)}.$$

Thus, for any given  $\lambda$ , among all service distribution with unit mean, the capacity is maximised by that service distribution for which  $\tilde{F}_S(\kappa)$  is minimised. For any service random variable  $S$ , by Jensen's inequality, we have  $\tilde{F}_S(\kappa) = \mathbf{E}[\exp(-\kappa S)] \geq \exp(-\kappa \mathbf{E}[S])$ . Therefore,  $\tilde{F}_S(\kappa)$  is minimised by  $S = \mathbf{E}[S]$ , i.e., a deterministic service time. ■

Theorems 3 and 4 hold for an M/GI/1 queue when  $p(W) = 1 - e^{-\kappa W}$ . For the M/M/1 case, we can derive a capacity expression for any functional form of the erasure probability  $p(\cdot)$ .

*Theorem 5:* For the M/M/1 erasure queue-channel, the capacity expression for any functional form of the erasure probability  $p(\cdot)$  is given by

$$\lambda \left( 1 - \frac{1-\lambda}{\lambda} \tilde{p} \left( \frac{1-\lambda}{\lambda} \right) \right) \text{ bits/sec.},$$

where for any  $u > 0$ ,  $\tilde{p}(u) := \int \exp(-ux) p(x) dx$  is the Laplace transform of  $p(\cdot)$ . Furthermore, the capacity maximizing arrival rate is

$$1 - \arg \min_{u \in (0,1)} u \left( 1 + \tilde{p} \left( \frac{u}{1-u} \right) \right). \quad (6)$$

*Proof:* This proof uses the exponential waiting time distribution of M/M/1 queue to relate the capacity to Laplace transform of  $p(\cdot)$ . It is known that the waiting time in M/M/1 is distributed as  $\exp(\frac{1-\lambda}{\lambda})$  for  $\mu = 1$ . Thus,

$$\begin{aligned} \mathbf{E}[p(W)] &= \int_0^\infty p(w) \frac{1-\lambda}{\lambda} \exp\left(\frac{1-\lambda}{\lambda} w\right) dw \\ &= \frac{1-\lambda}{\lambda} \tilde{p} \left( \frac{1-\lambda}{\lambda} \right). \end{aligned}$$

Thus, the capacity is given by  $\lambda \left( 1 - \frac{1-\lambda}{\lambda} \tilde{p} \left( \frac{1-\lambda}{\lambda} \right) \right)$  bits/sec.

Therefore, the capacity maximising arrival rate is the one that maximises this expression:

$$\begin{aligned} &\arg \max_{\lambda \in (0,1)} \lambda \left( 1 - \frac{1-\lambda}{\lambda} \tilde{p} \left( \frac{1-\lambda}{\lambda} \right) \right) \\ \iff &\arg \max_{\lambda \in (0,1)} \left( \lambda - (1-\lambda) \tilde{p} \left( \frac{1-\lambda}{\lambda} \right) \right) \\ \iff &1 - \arg \max_{u \in (0,1)} \left( 1 - u - u \tilde{p} \left( \frac{u}{1-u} \right) \right) \\ \iff &1 - \arg \min_{u \in (0,1)} u \left( 1 + \tilde{p} \left( \frac{u}{1-u} \right) \right). \end{aligned}$$

## F. Quantum Depolarizing Queue-channel

Our techniques can be easily extended to study another important additive quantum channel, namely the depolarizing channel. Suppose that the  $j$ th qubit spends  $W_j = w_j$  seconds in the queue. The quantum depolarizing channel acts on a given qubit  $\rho_j$  as,  $\rho_j \rightarrow (1 - p(w_j))\rho_j + p(w_j)\frac{I}{2}$ , where  $I/2$  is the maximally mixed state [3]. Additivity of the Holevo information of the depolarizing channel was shown by King [15], leading to a single-letter formula for its classical capacity, originally noted in [14]. Here, we evaluate the classical capacity of the quantum depolarizing queue-channel as stated below.

*Theorem 6:* The classical capacity of a quantum depolarizing queue-channel with arrival rate  $\lambda$  is,

$$C = \lambda \mathbf{E}_\pi \left[ 1 - h \left( \frac{p(W)}{2} \right) \right] \text{ bits/sec},$$

provided the receiver knows the arrival and the departure times of the qubits. Here,  $h(\cdot)$  denotes the binary entropy function, and the expectation is taken over the stationary distribution of the waiting time  $W$  in the queue.

*Proof:* Since the depolarising channel satisfies additivity [15], the upper bound in Theorem 1 holds. Furthermore, the Holevo information for a single-qubit depolarising channel is evaluated to be [14]

$$\chi(\mathcal{N}_W) = 1 - h \left( \frac{p(W)}{2} \right).$$

To prove achievability, we first note that the induced classical channel corresponding to the quantum depolarizing channel is a binary symmetric channel with crossover probability  $\phi(w) = p(w)/2$  (see [14]). Then, we invoke the capacity result proved in Theorem 8 in Sec. IV-C for the induced classical channel. ■

*Remark 2:* We note that for the depolarising queue channel, when the arrival and departure times are known to the receiver, we get matching upper and lower bounds from Theorems 1 and 8, thus characterising the capacity exactly. However, when the arrival and departure times are not known to the receiver, there is a 'Jensen gap' between the upper and lower bounds in Theorems 1 and 8 respectively. Thus, when the waiting times are not known to the receiver, we are unable to precisely characterise the capacity — we know only that it lies in between  $\lambda \left( 1 - h \left( \mathbf{E}_\pi \left[ \frac{p(W)}{2} \right] \right) \right)$  and  $\lambda \mathbf{E}_\pi \left[ 1 - h \left( \frac{p(W)}{2} \right) \right]$  bits/sec.

*Remark 3:* We remark that the upper bound and the capacity results which have been stated and proved for qubit queue-channels, do extend readily to the case of qudit ( $d$ -dimensional) queue-channels as well. The upper bound on the queue-channel capacity stated in Theorem 1 is of course independent of the dimension of the states and channels involved; rather it depends only on the additivity of the Holevo information for the channel. However, the capacity expressions do change with the dimensionality of the system. Specifically, the classical capacity of a qudit quantum erasure queue-channel is  $C_d = \lambda \log_2 d \mathbf{E}_\pi [1 - p(W)]$  bits/sec. and the classical capacity of a qudit depolarizing queue-channel (when the waiting times are known to the receiver) is

$$\begin{aligned} C_d &= \lambda \mathbf{E}_\pi \left[ \log_2 d + \left( 1 - p + \frac{p}{d} \right) \log_2 \left( 1 - p + \frac{p}{d} \right) + \right. \\ &\quad \left. (d-1) \frac{p}{d} \log_2 \frac{p}{d} \right] \text{ bits/sec.} \end{aligned}$$

## IV. CAPACITY OF THE INDUCED CLASSICAL QUEUE-CHANNEL

In this section we define and characterize the capacity of the induced classical channel corresponding to a general quantum queue-channel. Formally, the  $j$ th symbol  $X_j \in \mathcal{X} = \{0, 1\}^2$ ,

<sup>2</sup>Results extend to any finite alphabet.



is now encoded as one of a set of orthogonal states  $\{|\psi_{X_j}\rangle\}$  belonging to a Hilbert space  $\mathcal{H}$  of dimension 2. The noisy output state  $|\tilde{\psi}_j\rangle$  is measured by the receiver in some fixed basis, and decoded as the output symbol  $Y_j \in \mathcal{Y}$ . This measurement induces a conditional probability distribution  $\mathbf{P}(Y_j|X_j, W_j)$ . The *induced classical channel* is a sequence of conditional distributions from  $\mathcal{X}^n$  to  $\mathcal{Y}^n$ :

$$\mathbf{P}(Y^n|X^n, W^n) = \prod_{j=1}^n \mathbf{P}(Y_j|X_j, W_j) \text{ for } n \in \{1, 2, \dots\}. \quad (7)$$

As argued in Lemma 1,  $Y_j$  is conditionally independent of other random variables, given  $X_j$  and  $W_j$ , which leads to the product form above. Throughout,  $Z^n = (Z_1, Z_2, \dots, Z_n)$  denotes an  $n$ -dimensional vector and  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n, \dots)$  denotes an infinite sequence of random variables.

We are interested in defining and computing the information capacity of this induced classical queue-channel. As mentioned earlier, we restrict ourselves to using a fixed set of orthogonal states to encode the classical symbols at the sender's side, and measuring in some fixed basis at the receiver's end. For this reason, the capacity of the induced classical channel is not the same as the *classical capacity of the quantum channel* resulting from the underlying decoherence model.

The motivation behind the study of the induced classical queue-channel is twofold. Firstly, the achievability result (i.e., the capacity lower bound) for this channel is applicable to the classical capacity of the quantum queue-channel studied in the previous section. Indeed, as we show below, when the arrival and departure times of the symbols are known at the receiver, the capacities of the induced classical erasure and binary symmetric queue-channels are the same as that of the classical capacities of the corresponding quantum queue-channels. Secondly, the induced queue-channel has other interesting applications as well. It closely models delay-sensitive communication systems where data packets become useless after a deadline [2], [22].

#### A. Definitions

Let  $M$  be the message transmitted from a set  $\mathcal{M}$  and  $\hat{M} \in \mathcal{M}$  be its estimate at the receiver.

*Definition 6:* An  $(n, \bar{R}, \epsilon, T)$  code consists of the encoding function  $X^n = f(M)$  and the decoding function  $\hat{M} = g(X^n, A^n, D^n)$ , where the cardinality of the message set  $|\mathcal{M}| = 2^{n\bar{R}}$ , the expected total time for all the symbols of any codeword to reach to the receiver is less than  $T$ , and the average probability of error at the decoder is less than  $\epsilon$ .

If the average probability of error of a code is less than  $\epsilon$  the code is called an  $\epsilon$ -achievable code.

*Definition 7:* For any  $0 < \epsilon < 1$ , if there exists an  $\epsilon$ -achievable code  $(n, \bar{R}, \epsilon, T)$ , the rate  $R = \frac{\bar{R}}{T}$  is said to be achievable.

*Definition 8:* The information capacity of the induced classical queue-channel is the supremum of all achievable rates for a given arrival and service process and is denoted by  $C_{Ind}$  bits per unit time.

We assume that the transmitter knows the arrival process statistics, but not the realizations before it does the encoding.

However, depending on the scenario, the receiver may or may not know the realization of the arrival and the departure time of each symbol.

The following result is a consequence of the general channel capacity expression in [23] and is useful in deriving the main results on single letter capacity expression.

*Proposition 4:* The capacity of the queue-channel (in bits/sec) described above is given by

$$C_{Ind} = \lambda \sup_{\mathbf{P}(\mathbf{X})} \mathbf{I}(\mathbf{X}; \mathbf{Y}|\mathbf{W}), \quad (8)$$

when the receiver knows the arrival and the departure time of each symbol. On the other hand, when the receiver does not have that information, the capacity is,

$$C_{Ind} = \lambda \sup_{\mathbf{P}(\mathbf{X})} \mathbf{I}(\mathbf{X}; \mathbf{Y}). \quad (9)$$

Here,  $\mathbf{I}$  is the usual notation for classical *inf-information rate* [23]. Specifically,  $\mathbf{I}(\mathbf{X}; \mathbf{Y})$  is defined as the limit inferior in probability of the *information density* sequence  $\left\{ \frac{1}{n} \log \frac{\mathbf{P}(Y^n|X^n)}{\mathbf{P}(Y^n)} \right\}_{n=1}^{\infty}$ .

*Proof:* The case where arrival and departure times of the qubits are not known follows directly from [23] using (limiting) stationarity and ergodicity of the arrival and the departure processes of the queue. Note that a string of  $n$  qubits see the joint channel

$$\mathbf{P}(Y^n|X^n) = \int_{W^n} \prod_{j=1}^n \mathbf{P}(Y_j|X_j, W_j) d\mathbf{P}(W^n),$$

and the number of qubits coming out of the queue per unit time (asymptotically) is  $\lambda$ . Combining these two facts with standard information spectrum results we get the desired capacity result.

When the arrival and the departures times of the qubits are known at the receiver, the channel behaves like  $X^n \rightarrow (Y^n, A^n, D^n)$ . In that case it follows from [23] that the capacity of this channel is

$$\lambda \sup_{\mathbf{P}(\mathbf{X})} \mathbf{I}(\mathbf{X}; (\mathbf{Y}, \mathbf{A}, \mathbf{D})).$$

Next, note that  $W_j = D_j - A_j$  for all  $j$ , and since no information is encoded in timings,  $\mathbf{X}$  is independent of  $(\mathbf{A}, \mathbf{D})$ . By definition,  $\mathbf{I}(\mathbf{X}; (\mathbf{Y}, \mathbf{A}, \mathbf{D}))$  is the limit inferior in probability of the following sequence:

$$\begin{aligned} & \frac{1}{n} \log \frac{\mathbf{P}(Y^n, A^n, D^n|X^n)}{\mathbf{P}(Y^n, A^n, D^n)} \\ &= \frac{1}{n} \log \frac{\mathbf{P}(A^n, D^n|X^n) \mathbf{P}(Y^n|X^n, A^n, D^n)}{\mathbf{P}(A^n, D^n) \mathbf{P}(Y^n|A^n, D^n)} \\ &= \frac{1}{n} \log \frac{\mathbf{P}(Y^n|X^n, A^n, D^n)}{\mathbf{P}(Y^n|A^n, D^n)} \\ &= \frac{1}{n} \log \frac{\mathbf{P}(Y^n|X^n, W^n)}{\mathbf{P}(Y^n|W^n)} \end{aligned} \quad (10)$$

The step before the last is due to independence of  $X^n$  and  $(A^n, D^n)$ . Also, as per the channel model we have the Markov property  $(A^n, D^n) \rightarrow W^n \rightarrow Y^n$ , which leads to the final expression. The liminf in probability of the quantity in (10), is by definition, equal to  $\mathbf{I}(\mathbf{X}; \mathbf{Y}|\mathbf{W})$ . ■

## B. Erasure Queue-Channels

Erasure channels are ubiquitous in classical as well as quantum information theory. We consider a classical erasure queue-channel [14] induced by the quantum erasure queue-channel where the  $j$ th state  $|\psi_{X_j}\rangle$  remains unaffected with probability  $1 - p(W_j)$ , and is erased to a state  $|e\rangle$  with probability  $p(W_j)$ , where  $p : [0, \infty) \rightarrow [0, 1]$  is typically increasing. Such a model also captures the communication scenarios where information packets become useless (erased) after a deadline. For such an erasure channel, a single letter expression for capacity can be obtained.

*Theorem 7:* For the classical erasure queue-channel defined above, the capacity (in bits/sec) is given by  $\lambda \mathbf{E}_\pi [1 - p(W)]$ , irrespective of the receiver's knowledge of the arrival and the departure times of symbols.

*Remark 4:* For a general queue-channel, the capacity could depend on the receiver's knowledge (or lack thereof) of the arrival and the departure times of symbols. However, for the erasure case, the received symbol is either correct or is erased, but is never wrong. This makes the knowledge of the arrival and departure times irrelevant, and the capacity remains the same in both cases.

*Proof of Theorem 7:* We first prove the result for the case when the receiver knows the arrival and the departure times.

The proof uses an upper-bound on  $\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W})$  in terms of two conditional sup-entropy rates rate and shows that  $\mathbf{E}_\pi [1 - p(W)]$  is an upper-bound on  $\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W})$ . On the other hand, using a similar lower-bound on  $\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W})$  we show that for a choice of distribution of  $\{X_n\}$  (namely, i.i.d. uniform),  $\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W})$  is no smaller than  $\mathbf{E}_\pi [1 - p(W)]$ . More precisely, we use the following two bounds, which follow from the properties of limit superior and limit inferior [24, Chap. 3]:

$$\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W}) \leq \overline{\mathbf{H}}(\mathbf{Y}|\mathbf{W}) - \overline{\mathbf{H}}(\mathbf{Y}|\mathbf{X}, \mathbf{W}), \quad (11)$$

$$\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W}) \geq \underline{\mathbf{H}}(\mathbf{Y}|\mathbf{W}) - \overline{\mathbf{H}}(\mathbf{Y}|\mathbf{X}, \mathbf{W}), \quad (12)$$

where  $\overline{\mathbf{H}}(\mathbf{Y}|\mathbf{W})$  and  $\overline{\mathbf{H}}(\mathbf{Y}|\mathbf{X}, \mathbf{W})$  are respectively the lim-sup in probability of the sequences  $\left\{ \frac{1}{n} \log \frac{1}{\mathbf{P}(Y^n|W^n)} \right\}_{n=1}^\infty$  and  $\left\{ \frac{1}{n} \log \frac{1}{\mathbf{P}(Y^n|X^n, W^n)} \right\}_{n=1}^\infty$ . Similarly,  $\underline{\mathbf{H}}(\mathbf{Y}|\mathbf{W})$  is the lim-inf in probability of  $\left\{ \frac{1}{n} \log \frac{1}{\mathbf{P}(Y^n|W^n)} \right\}_{n=1}^\infty$ .

We now state three lemmas, which lead to the capacity result.

*Lemma 2:* For the induced classical erasure queue-channel,

$$\overline{\mathbf{H}}(\mathbf{Y}|\mathbf{X}, \mathbf{W}) = \mathbf{E}_\pi [h(p(W))],$$

irrespective of the choice of  $\mathbf{P}_X$ .

*Lemma 3:* For the induced classical erasure queue-channel,

$$\overline{\mathbf{H}}(\mathbf{Y}|\mathbf{W}) \leq 1 + \mathbf{E}_\pi [h(p(W))] - \mathbf{E}_\pi [p(W)],$$

irrespective of the choice of  $\mathbf{P}_X$ .

*Lemma 4:* For uniform and i.i.d  $\{X_i\}$

$$\underline{\mathbf{H}}(\mathbf{Y}|\mathbf{W}) = 1 + \mathbf{E}_\pi [h(p(W))] - \mathbf{E}_\pi [p(W)].$$

The proofs of Lemmas 2, 3, and 4 appear in Appendix B. Combining Lemmas 2 and 3 we get a  $\mathbf{P}_X$ -independent upper-bound on  $\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W})$ :

$$\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W}) \leq 1 - \mathbf{E}_\pi [p(W)].$$

A matching lower bound on  $\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W})$  is readily obtained using Lemmas 2 and 4. Thus, for the erasure queue-channel,

$$\sup_{\mathbf{P}_X} \underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W}) = 1 - \mathbf{E}_\pi [p(W)].$$

We obtain the capacity of this channel by multiplying the above expression by  $\lambda$  (see Proposition 4).

When the arrival and departure times are *not* available at the receiver, one way to prove the capacity result is to analyze  $\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y})$ . However, it is easier to use a dominance argument for the probability of errors in the two scenarios: arrival and departure times known, versus unknown, at the receiver. Consider any encoding and decoding scheme which attains capacity when the arrival and departure times are known at the receiver. Note that under this encoding scheme the rate is close to the capacity and the probability of error is tending to 0. By the optimality of MAP (maximum a posteriori probability) decoding, if we replace the decoding scheme by a MAP decoder the probability error cannot increase. For the case where the receiver does not know arrival and departure times, we can use the *same* encoding scheme, followed by a MAP decoding. For an erasure channel, the MAP decoding depends only on the prior probabilities of the codewords that match the received codeword at the non-erased symbols. Therefore, the knowledge of  $\{p(W_i)\}$  (and hence that of arrival and departure times) has no effect on the MAP decoding. Thus, it follows that the same rate and the probability of error are also achievable in the absence of arrival and departure times. ■

Note that the above theorem, particularly the achievability result, completes the proof of Proposition 3. Furthermore, since the capacity of the induced classical erasure queue-channel is same as the classical capacity of the quantum erasure queue-channel, the results in Theorem 3, 4, and 5 are applicable here as well.

## C. Binary Symmetric Queue-channels

In this section, we state our results for a binary symmetric queue-channel. This classical channel is induced by the depolarizing quantum queue-channel. The probability of a bit flip is captured by a function  $\phi : [0, \infty) \rightarrow [0, 0.5]$ ,  $\phi(W_i)$ , where  $\mathbf{P}(Y_i \neq X_i | X_i, W_i) = \phi(W_i)$ . Note that the  $p(W)$  for the depolarizing channel and  $\phi(W)$  are related as  $\phi(W) = \frac{p(W)}{2}$ .

*Theorem 8:* For a binary symmetric queue-channel, the capacity is  $\lambda(1 - \mathbf{E}_\pi [h(\phi(W))])$  bits/sec when the receiver knows the arrival and departure times of the symbols. The capacity is lower bounded by  $\lambda(1 - h(\mathbf{E}_\pi [\phi(W)]))$  when the receiver does not know the arrival and departure times of the symbols.

*Proof:*

*Case 1: arrival and departures times are known at the receiver* These steps are similar to the proof of the capacity result for the erasure queue-channel. Once again, we use the inequalities in (11) and (12) respectively, for the upper and the lower bounds. For proving the upper-bound, we argue along the same lines as Lemmas 2 and 3 to obtain

$$\begin{aligned} \overline{\mathbf{H}}(\mathbf{Y}|\mathbf{W}) &\leq 1, \text{ and} \\ \overline{\mathbf{H}}(\mathbf{Y}|\mathbf{X}, \mathbf{W}) &= \mathbf{E}_\pi [h(\phi(W))]. \end{aligned} \quad (13)$$

The last equality uses the following simple insight: the channel probability distribution for binary symmetric queue-channel can be written as

$$\mathbf{P}(Y^n|X^n, W^n) = \mathbf{P}(N^n|W^n),$$

where  $Y_i = X_i + N_i$  (modulo 2) and  $N_i \sim \text{Bernoulli}(\phi(W_i))$ . The rest follows using conditional independence of  $\{N_i\}$  given  $\{W_i\}$ , and the ergodicity of the process  $\{W_i\}$ . This completes the proof of the upper-bound.

For deriving the lower bound, we fix the input distribution to be i.i.d. Bernoulli(0.5) independent of  $\mathbf{W}$ . An elementary calculation then shows that  $\mathbf{Y}$  is also i.i.d. Bernoulli(0.5) and is independent of  $\mathbf{W}$ . Thus,

$$\underline{\mathbf{H}}(\mathbf{Y}|\mathbf{W}) = 1.$$

Using (12) and (13) we obtain the following lower bound:

$$\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}|\mathbf{W}) \geq 1 - h(\mathbf{E}_\pi[\phi(W)]).$$

*Case II: arrival and departures times are not known at the receiver*

For obtaining the lower-bound in this case we use the following inequality from [24].

$$\underline{\mathbf{I}}(\mathbf{X}; \mathbf{Y}) \geq \underline{\mathbf{H}}(\mathbf{Y}) - \overline{\mathbf{H}}(\mathbf{Y}|\mathbf{X}).$$

Again, for uniform i.i.d.  $\mathbf{X}$ , we have  $\underline{\mathbf{H}}(\mathbf{Y}) = 1$ .

For obtaining a lower bound, we obtain an upper-bound on  $\overline{\mathbf{H}}(\mathbf{Y}|\mathbf{X})$  by upper-bounding the limit

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left( \frac{1}{\mathbf{P}(N^n)} \right) \quad (14)$$

in an almost sure sense.

Since  $\{W_i\}$  is stationary and ergodic, so is  $\{N_i\}$ . To bound the limit in (14), we use successive Markov approximations of this ergodic process, where the Markov approximation approaches the ‘true’ process, as the order or the memory of the Markov process tends to infinity. (A celebrated application of this successive Markov approximation approach appears in proving the asymptotic optimality of the LZ78 universal source coding algorithm — see [25, Section 13.5.2]).

To be precise, we first consider the case where  $N_i$  is a  $k$ -th order stationary and ergodic Markov process and obtain a bound on the above limit, which does not depend on  $k$ . Let  $\{\tilde{N}_t\}$  be a  $k$ -th order Markov process with a transition kernel

$$\begin{aligned} & \prod_{j=1}^n \mathbf{P}(\tilde{N}_j = u_j | \tilde{N}_{j-1} = u_{j-1}, \dots, \tilde{N}_{(j-k)_+} = u_{(j-k)_+}) \\ &= \prod_{j=1}^n \mathbf{P}(N_j = u_j | N_{j-1} = u_{j-1}, \dots, N_{(j-k)_+} = u_{(j-k)_+}), \end{aligned}$$

and  $\mathbf{P}(\tilde{N}_1 = u_1) = \mathbf{P}(N_1 = u_1)$ , where  $(j-k)_+ = \max(j-k, 1)$ . As claimed in the proof of [25, Lemma 13.5.4], (14) is equal to  $\lim_{k \rightarrow \infty} \beta_k$ , where  $\beta_k$  is the almost sure limit of  $\frac{1}{n} \log \left( \frac{1}{\mathbf{P}(\tilde{N}^n)} \right)$  for the  $k$ th order Markov process.

Note that  $\frac{1}{n} \log \left( \frac{1}{\mathbf{P}(\tilde{N}^n)} \right)$  is given by

$$\frac{1}{n} \log \left( \frac{1}{\mathbf{P}(\tilde{N}^k)} \right) - \frac{1}{n} \sum_{t=k+1}^n \log \mathbf{P}(\tilde{N}_t | \tilde{N}_{t-1}, \dots, \tilde{N}_{t-k+1}).$$

By ergodicity, the above expression has the following almost sure limit:

$$\mathbf{E}_{\tilde{N}_{k+1}, \tilde{N}_k, \dots, \tilde{N}_1} \left[ -\log \mathbf{P}(\tilde{N}_{k+1} | \tilde{N}_1, \tilde{N}_2, \dots, \tilde{N}_k) \right].$$

The expectation above is over the joint distribution of  $(\tilde{N}_1, \tilde{N}_2, \dots, \tilde{N}_k, \tilde{N}_{k+1})$ , and can be recognized readily as the conditional entropy of  $\tilde{N}_{k+1}$ , given  $\tilde{N}_1, \tilde{N}_2, \dots, \tilde{N}_k$ . This in turn implies that

$$\begin{aligned} & \mathbf{E}_{\tilde{N}_{k+1}, \tilde{N}_k, \dots, \tilde{N}_1} \left[ -\log \mathbf{P}(\tilde{N}_{k+1} | \tilde{N}_1, \tilde{N}_2, \dots, \tilde{N}_k) \right] \\ &= \mathbf{E}_{\tilde{N}_k, \dots, \tilde{N}_1} \left[ \mathbf{E}_{\tilde{N}_{k+1}} \left[ -\log \mathbf{P}(\tilde{N}_{k+1} | \tilde{N}_1, \tilde{N}_2, \dots, \tilde{N}_k) \right] \right] \quad (15) \end{aligned}$$

$$\begin{aligned} &= \mathbf{E}_{\tilde{N}_k, \dots, \tilde{N}_1} \left[ h(\mathbf{P}(\tilde{N}_{k+1} | \tilde{N}_1, \tilde{N}_2, \dots, \tilde{N}_k)) \right] \\ &\leq h(\mathbf{E}_{\tilde{N}_k, \dots, \tilde{N}_1} [\mathbf{P}(\tilde{N}_{k+1} | \tilde{N}_1, \tilde{N}_2, \dots, \tilde{N}_k)]) \quad (16) \end{aligned}$$

$$= h(\mathbf{P}(\tilde{N}_k)) = h(\mathbf{E}_\pi[\phi(W)]). \quad (17)$$

Here, (15), (16) and (17) follow from iterated expectation, Jensen’s inequality due to concavity of  $h(\cdot)$ , and stationarity of  $\tilde{N}_k$  which is Bernoulli( $\mathbf{E}_\pi[\phi(W)]$ ), respectively. Thus,  $h(\mathbf{E}_\pi[\phi(W)])$  is an upper-bound on  $\mathbf{E}_{\tilde{N}_{k+1}, \tilde{N}_k, \dots, \tilde{N}_1} \left[ -\log \mathbf{P}(\tilde{N}_{k+1} | \tilde{N}_1, \tilde{N}_2, \dots, \tilde{N}_k) \right]$  for all  $k$ . Hence, (14) is upper-bounded by  $h(\mathbf{E}_\pi[\phi(W)])$ . ■

When the receiver does not know the arrival and departure times,  $\lambda(1 - \mathbf{E}_\pi[h(\phi(W))])$  is still an upper-bound on the capacity. Thus, there remains a ‘Jensen’s gap’ between the upper and the lower bound in that case.

We remark that the above capacity result for the binary symmetric queue-channel is a special case of the capacity result for a class of classical queue-channels called random bijective queue-channels — see [22].

## V. CONCLUDING REMARKS AND FUTURE WORK

In this paper, we considered quantum queue-channels as a framework to study the sequential processing of qubits, and derived fundamental capacity limits for sequential quantum information processing. For the class of additive queue-channels, we derived a single-letter upper bound for the capacity in terms of the stationary waiting time in the queue. We then showed that this capacity upper bound is achievable for two important noise models, namely the erasure channel and the depolarising channel. We also showed that a classical coding/decoding strategy is capacity achieving for these channel.

There is ample scope for further work along several directions. First, we can study the queue-channel capacity for other noise channels whose Holevo information is known to be additive. This includes the class of entanglement breaking channels [26] as well as the class of Hadamard channels [16]. The upper bound on the classical capacity of the queue-channel proved here will indeed hold in all of these cases. However, whether the upper bound is achievable or not will depend on the structure of the induced classical channel. As seen in the case of the depolarizing queue-channel, the upper bound may or may not be tight depending on whether the receiver’s knowledge of the waiting times of the qubits. Whether our capacity results can be extended to the larger

$$\mathbf{I}(\{\vec{P}, \vec{\rho}\}, \vec{\mathcal{N}}_{\vec{W}}) = \sup \left\{ a \in \mathbb{R}^+ \left| \lim_{n \rightarrow \infty} \sum_{X^n \in \mathcal{X}^{(n)}} P^n(X^n) \text{Tr} \left[ \mathcal{N}_{W^{(n)}}^{(n)}(\rho_{X^n}) \left\{ \Gamma_{\{P^n(X^n), \rho_{X^n}\}}(a) > 0 \right\} \right] = 1 \right. \right\}. \quad (18)$$

$$\begin{aligned} & -\log \mathbf{P}(Y^n | W^n) = -\log \mathbf{P}(\{Y_i : i \in J^\mathcal{E}\}, \{Y_i : i \notin J^\mathcal{E}\} | W^n) \\ & = -\log \mathbf{P}(\{Y_i : i \in J^\mathcal{E}\} | \{W_i : i \in J^\mathcal{E}\}) - \log \mathbf{P}(\{Y_i : i \notin J^\mathcal{E}\} | \{W_i : i \notin J^\mathcal{E}\}) \end{aligned} \quad (19)$$

$$= -\sum_{i \in J^\mathcal{E}} \log \mathbf{P}(\mathcal{E} | W_i) - \log \mathbf{P}(\{Y_i : i \notin J^\mathcal{E}\} | \{W_i : i \notin J^\mathcal{E}\}) \quad (20)$$

$$= -\sum_{i \in J^\mathcal{E}} \log \mathbf{P}(\mathcal{E} | W_i) - \log \mathbf{P}(\{Y_i \neq \mathcal{E} : i \notin J^\mathcal{E}\}, \{Y_i : i \notin J^\mathcal{E}\} | \{W_i : i \notin J^\mathcal{E}\}) \quad (21)$$

$$= -\sum_{i \in J^\mathcal{E}} \log p(W_i) - \log \mathbf{P}(\{Y_i \neq \mathcal{E} : i \notin J^\mathcal{E}\} | \{W_i : i \notin J^\mathcal{E}\}) - \log \mathbf{P}(\{Y_i : i \notin J^\mathcal{E}\} | \{Y_i \neq \mathcal{E} : i \notin J^\mathcal{E}\}, \{W_i : i \notin J^\mathcal{E}\})$$

$$\begin{aligned} & = -\sum_{i \in J^\mathcal{E}} \log p(W_i) - \sum_{i \notin J^\mathcal{E}} \log(1 - p(W_i)) - \log \mathbf{P}(\{Y_i : i \notin J^\mathcal{E}\} | \{Y_i \neq \mathcal{E} : i \notin J^\mathcal{E}\}, \{W_i : i \notin J^\mathcal{E}\}) \\ & = -\sum_{i=1}^n [\mathbf{1}(Y_i = \mathcal{E}) \log(p(W_i)) + \mathbf{1}(Y_i \neq \mathcal{E}) \log(1 - p(W_i))] - \log \mathbf{P}(\{Y_i : i \notin J^\mathcal{E}\} | \{Y_i \neq \mathcal{E}, W_i : i \notin J^\mathcal{E}\}). \end{aligned} \quad (22)$$

class of additive queue-channels is an interesting avenue for further investigations. In this context, the upper bounds proved in [27] and the corresponding achievability results for finite-length codes might be useful. An immediate application of our work would be in identifying optimal rates for reliable communication across quantum networks. We can also quantitatively evaluate the impact of using quantum codes with finite block lengths to protect qubits from errors. Employing a code would enhance robustness to errors, but would also increase the waiting time due to the increased number of qubits to be processed. It would be interesting to characterise this tradeoff, and identify the regimes where using coded qubits would be beneficial or otherwise.

More generally, the qubits-through-queues paradigm studied here may be used to analyse various performance aspects of quantum circuits, such as, quantifying the best achievable processing rate for asymptotically perfect accuracy. As we enter an era of quantum networks and noisy intermediate-scale quantum technologies [28], our work begins to quantitatively address the impact of decoherence on the performance limits of quantum information processing systems.

## APPENDIX

### A. Quantum Inf-Information Rate

Consider the quantum queue-channel  $\vec{\mathcal{N}}_{\vec{W}}$  comprising a sequence of channels  $\{\mathcal{N}_{W^{(n)}}^{(n)}\}_{n=1}^\infty$ , which are parameterised by the corresponding waiting time sequences  $\{W^{(n)}\}_{n=1}^\infty$ . Let  $\vec{P}$  denote the totality of sequences  $\{P^n(X^n)\}_{n=1}^\infty$  of probability distributions (with finite support) over input sequences  $X^n$ , and  $\{\vec{\rho}\}$  denote the sequences of states  $\{\rho_{X^n}\}$  corresponding to the encoding  $X^n \rightarrow \rho_{X^n}$ . For any  $a \in \mathbb{R}^+$  and  $n$ , we first define the operator

$$\Gamma_{\{P^n(X^n), \rho_{X^n}\}}(a) = \mathcal{N}_{W^{(n)}}^{(n)}(\rho_{X^n}) - e^{an} \sum_{X^n \in \mathcal{X}^{(n)}} P^n(X^n) \mathcal{N}_{W^{(n)}}^{(n)}(\rho_{X^n}).$$

Let  $\{\Gamma_{\{P^n(X^n), \rho_{X^n}\}}(a) > 0\}$  denote the projector onto the positive eigenspace of the operator  $\Gamma_{\{P^n(X^n), \rho_{X^n}\}}(a)$ .

*Definition 9:* The quantum inf-information rate [12]  $\mathbf{I}(\{\vec{P}, \vec{\rho}\}, \vec{\mathcal{N}}_{\vec{W}})$  is defined as in (18).

This is the quantum analogue of the classical inf-information rate originally defined in [23], [24]. The central result of [12] is to show that the classical capacity of the channel sequence  $\vec{\mathcal{N}}_{\vec{W}}$  is given by

$$C = \sup_{\{\vec{P}, \vec{\rho}\}} \mathbf{I}(\{\vec{P}, \vec{\rho}\}, \vec{\mathcal{N}}_{\vec{W}}).$$

### B. Proofs of Lemmas 2, 3 and 4

*Proof of Lemma 2:* We are interested in computing  $\overline{\mathbf{H}}(\mathbf{Y} | \mathbf{X}, \mathbf{W})$ , which is the limsup in probability of the sequence  $\left\{ \frac{1}{n} \log \frac{1}{\mathbf{P}(Y^n | X^n, W^n)} \right\}_{n=1}^\infty$ . We start with the product form (7) for the channel, and also note that  $\mathbf{P}(Y_i | X_i, W_i) = p(W_i)$  if  $Y_i$  is an erasure, else, it is  $1 - p(W_i)$ . Combining these observations we obtain

$$\begin{aligned} \log \frac{1}{\mathbf{P}(Y^n | X^n, W^n)} &= \sum_{i=1}^n \log \frac{1}{\mathbf{P}(Y_i | X_i, W_i)} \\ &= \left[ \sum_{i \in J^\mathcal{E}} \log \frac{1}{\mathbf{P}(Y_i | X_i, W_i)} + \sum_{i \in [n] \setminus J^\mathcal{E}} \log \frac{1}{\mathbf{P}(Y_i | X_i, W_i)} \right] \\ &= -\sum_{i=1}^n [\mathbf{1}(Y_i = \mathcal{E}) \log(p(W_i)) + \mathbf{1}(Y_i \neq \mathcal{E}) \log(1 - p(W_i))], \end{aligned}$$

where  $\mathcal{E}$  represents erasure, and  $J^\mathcal{E}$  denotes the set of indices for which  $Y_i = \mathcal{E}$ . By ergodicity of the queue, the limit of the final expression divided by  $n$  exists almost surely as a finite constant, and is equal to  $\mathbf{E}_\pi[h(p(W))]$ .

Finally, we note that the lim-sup in probability and the lim-inf in probability of a sequence are both equal to the almost

sure limit of that sequence, whenever the latter exists. Hence, the above almost sure limit is the value of  $\bar{\mathbf{H}}(\mathbf{Y}|\mathbf{X}, \mathbf{W})$ . Note that for an erasure queue-channel this limit does not depend on the distribution of  $X^n$ . ■

*Proof of Lemma 3:* We are interested in computing  $\bar{\mathbf{H}}(\mathbf{Y}|\mathbf{W})$ , which is the limsup in probability of the sequence  $\left\{ \frac{1}{n} \log \frac{1}{\mathbf{P}(Y^n|W^n)} \right\}_{n=1}^{\infty}$ . Following standard conditional probability arguments, we get the series of equalities (19)–(22).

The equalities in (19) and (20) follow from the fact that given  $W_i$ , the probability of a symbol getting erased is independent of anything else (even the input symbols). The equality in (21) follows because the event  $\{Y_i \neq \mathcal{E} : i \notin J^{\mathcal{E}}\}$  contains the event  $\{Y_i : i \notin J^{\mathcal{E}}\}$ .

Note that we need to analyze only the second term in (22). Since the first term divided by  $n$ , as argued in the proof of Lemma 2, has an almost sure limit  $\mathbf{E}_{\pi}[h(p(W))]$ .

Note that for an erasure channel, if  $Y_i$  is not an erasure,  $Y_i$  has the same value as that of  $X_i$  and  $\{X_i\}$  are chosen independently of  $\{W_i\}$ . So, for any joint distribution  $\mathbf{P}_{\mathbf{X}}$  of input symbols:

$$\begin{aligned} & -\frac{1}{n} \log \mathbf{P}(\{Y_i : i \notin J^{\mathcal{E}}\} | \{Y_i \neq \mathcal{E}, W_i : i \notin J^{\mathcal{E}}\}) \\ &= -\frac{1}{n} \log \mathbf{P}_{\mathbf{X}}(\{Y_i : i \notin J^{\mathcal{E}}\}) \\ &= -\frac{n - |J^{\mathcal{E}}|}{n} \frac{1}{n - |J^{\mathcal{E}}|} \log \mathbf{P}_{\mathbf{X}}(\{Y_i : i \notin J^{\mathcal{E}}\}). \end{aligned}$$

By ergodicity  $\frac{|J^{\mathcal{E}}|}{n}$  converges a.s. to  $\mathbf{E}_{\pi}[p(W)] < 1$ . So, a.s.  $n - |J^{\mathcal{E}}| \rightarrow \infty$ . Thus, a.s.  $\lim_{n \rightarrow \infty} -\frac{1}{n - |J^{\mathcal{E}}|} \log \mathbf{P}_{\mathbf{X}}(\{Y_i : i \notin J^{\mathcal{E}}\})$  is upper-bounded by the entropy rate of an i.i.d. Bernoulli(0.5) process  $\mathbf{X}$ , i.e., 1. Thus, a.s. the second term in (22) is at most  $1 - \mathbf{E}_{\pi}[p(W)]$ . ■

*Proof of Lemma 4:* This proof is similar to the proof of Lemma 3, except the fact that

$$\lim_{n \rightarrow \infty} -\frac{1}{n - |J^{\mathcal{E}}|} \log \mathbf{P}_{\mathbf{X}}(\{Y_i : i \notin J^{\mathcal{E}}\}) = 1$$

when  $\mathbf{P}_{\mathbf{X}}$  is i.i.d. uniform. ■

## REFERENCES

- [1] P. Mandayam, K. Jagannathan, and A. Chatterjee, “The classical capacity of a quantum erasure queue-channel,” in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019.
- [2] K. Jagannathan, A. Chatterjee, and P. Mandayam, “Qubits through queues: The capacity of channels with waiting time dependent errors,” in *Proc. 25th National Conf. on Communication (NCC)*, Feb. 2019.
- [3] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [4] G. Wendin, “Quantum information processing with superconducting circuits: a review,” *Reports on Progress in Physics*, vol. 80, no. 10, p. 106001, 2017.
- [5] P. Aliferis, D. Gottesman, and J. Preskill, “Quantum accuracy threshold for concatenated distance-3 codes,” *Quantum Inf. Comput.*, vol. 6, no. quant-ph/0504218, pp. 97–165, 2005.
- [6] W. Kozłowski and S. Wehner, “Towards large-scale quantum networks,” in *Proceedings of the Sixth Annual ACM International Conference on Nanoscale Computing and Communication*, 2019, pp. 1–7.
- [7] N. Sangouard, C. Simon, H. De Riedmatten, and N. Gisin, “Quantum repeaters based on atomic ensembles and linear optics,” *Reviews of Modern Physics*, vol. 83, no. 1, p. 33, 2011.
- [8] L. Kleinrock, *Queueing Systems, Volume I: Theory*. John Wiley & Sons, Inc., 1975.

- [9] D. Felinto, C.-W. Chou, J. Laurat, E. Schomburg, H. De Riedmatten, and H. J. Kimble, “Conditional control of the quantum states of remote atomic memories for quantum networking,” *Nature Physics*, vol. 2, no. 12, pp. 844–848, 2006.
- [10] A. Chatterjee, D. Seo, and L. R. Varshney, “Capacity of systems with queue-length dependent service quality,” *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3950–3963, Jun. 2017.
- [11] V. Anantharam and S. Verdú, “Bits through queues,” *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 4–18, Jan. 1996.
- [12] M. Hayashi and H. Nagaoka, “General formulas for capacity of classical-quantum channels,” *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1753–1768, 2003.
- [13] A. Holevo and M. Shirokov, “On Shor’s channel extension and constrained channels,” *Communications in mathematical physics*, vol. 249, no. 2, pp. 417–430, 2004.
- [14] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, “Capacities of quantum erasure channels,” *Physical Review Letters*, vol. 78, no. 16, p. 3217, 1997.
- [15] C. King, “The capacity of the quantum depolarizing channel,” *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 221–229, 2003.
- [16] M. M. Wilde, *Quantum information theory*. Cambridge University Press, 2013.
- [17] A. S. Holevo, “Bounds for the quantity of information transmitted by a quantum communication channel,” *Problemy Peredachi Informatsii*, vol. 9, no. 3, pp. 3–11, 1973.
- [18] J. Watrous, *The theory of quantum information*. Cambridge University Press, 2018.
- [19] A. S. Holevo, “The capacity of the quantum channel with general signal states,” *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 269–273, 1998.
- [20] B. Schumacher and M. D. Westmoreland, “Sending classical information via noisy quantum channels,” *Physical Review A*, vol. 56, no. 1, p. 131, 1997.
- [21] C. Ballance, T. Harty, N. Linke, M. Sepiol, and D. Lucas, “High-fidelity quantum logic gates using trapped-ion hyperfine qubits,” *Physical review letters*, vol. 117, no. 6, p. 060504, 2016.
- [22] A. Chatterjee, K. Jagannathan, and P. Mandayam, “Qubits through queues: The capacity of channels with waiting time dependent errors,” arXiv:1804.00906, 2018.
- [23] S. Verdú and T. S. Han, “A general formula for channel capacity,” *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.
- [24] T. S. Han, *Information-Spectrum Methods in Information Theory*. Berlin: Springer, 2003.
- [25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: John Wiley & Sons, 2006.
- [26] P. W. Shor, “Additivity of the classical capacity of entanglement-breaking quantum channels,” *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4334–4340, 2002.
- [27] X. Wang, K. Fang, and M. Tomamichel, “On converse bounds for classical communication over quantum channels,” *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4609–4619, 2019.
- [28] J. Preskill, “Quantum computing in the NISQ era and beyond,” *Quantum*, vol. 2, p. 79, 2018.

**Prabha Mandayam** is an Assistant Professor at the Department of Physics, IIT Madras. Her research interests lie in the areas of quantum error correction, quantum information and quantum cryptography. She obtained her Masters in Physics from IIT Madras and Ph.D. in Physics from the California Institute of Technology.

**Krishna Jagannathan** is an Associate Professor at the Department of Electrical Engineering, IIT Madras. He obtained his B. Tech. in Electrical Engineering from IIT Madras in 2004, and the S.M. and Ph.D. degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology in 2006 and 2010 respectively. His research interests lie in the areas of stochastic modeling and analysis of communication networks, online learning, network control, and queuing theory.

**Avhishek Chatterjee** is an Assistant Professor at the Department of Electrical Engineering, IIT Madras. He obtained his Ph.D. in Electrical and Computer Engineering from The University of Texas at Austin. His research interest is in stochastic networks and information dynamics.