Deploy-As-You-Go Wireless Relay Placement: An Optimal Sequential Decision Approach using the Multi-Relay Channel Model

Arpan Chattopadhyay, Abhishek Sinha, Marceau Coupechoux, and Anurag Kumar *Fellow, IEEE*

Abstract—We use information theoretic achievable rate formulas for the multi-relay channel to study the problem of as-you-go deployment of relay nodes. The achievable rate formulas are for full-duplex radios at the relays and for decode-and-forward relaying. Deployment is done along the straight line joining a source node and a sink node at an unknown distance from the source. The problem is for a deployment agent to walk from the source to the sink, deploying relays as he walks, given the knowledge of the wireless path-loss model, and given that the distance to the sink node is exponentially distributed with known mean. As a precursor to the formulation of the deploy-as-you-go problem, we apply the multi-relay channel achievable rate formula to obtain the optimal power allocation to relays placed along a line, at fixed locations. This permits us to obtain the optimal placement of a given number of nodes when the distance between the source and sink is given. Numerical work for the fixed source-sink distance case suggests that, at low attenuation, the relays are mostly clustered close to the source in order to be able to cooperate among themselves, whereas at high attenuation they are uniformly placed and work as repeaters. We also prove that the effect of path-loss can be entirely mitigated if a large enough number of relays are placed uniformly between the source and the sink. The structure of the optimal power allocation for a given placement of the nodes, then motivates us to formulate the problem of as-you-go placement of relays along a line of exponentially distributed length, and with the exponential path-loss model, so as to minimize a cost function that is additive over hops. The hop cost trades off a capacity limiting term, motivated from the optimal power allocation solution, against the cost of adding a relay node. We formulate the problem as a total cost Markov decision process, establish results for the value function, and provide insights into the placement policy and the performance of the deployed network via numerical exploration.

Index Terms—Multi-relay channel, optimal relay placement, impromptu wireless networks, as-you-go relay placement.

❖

1 INTRODUCTION

Wireless interconnection of devices such as smart phones, or wireless sensors, to the wireline communication infrastructure is an important requirement. These are battery operated, resource constrained devices. Hence, due to the physical placement of these devices, or due to channel conditions, a direct one-hop link to the infrastructure "base-station" might not be feasible. In such situations, other nodes could serve as *relays* in order to realize a multi-hop path between the source device and the infrastructure. In the wireless sensor network context, the relays could be other wireless sensors or battery operated radio routers deployed specifically as relays. The relays are also resource constrained and a

Arpan Chattopadhyay and Anurag Kumar are with the Electrical Communication Engineering (ECE) Department, Indian Institute of Science (IISc), Bangalore-560012, India (e-mail: arpanc.ju@gmail.com, anurag@ece.iisc.ernet.in). Abhishek Sinha is with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139 (e-mail: sinhaa@mit.edu). Marceau Coupechoux is with Telecom ParisTech and CNRS LTCI, Dept. Informatique et Réseaux, 23, avenue d'Italie, 75013 Paris, France (e-mail: marceau.coupechoux@telecom-paristech.fr). This work was done during the period when he was a Visiting Scientist in the ECE Deparment, IISc.

Figure 1: A source and a sink connected by a multi-hop path comprising N relay nodes along a line.

cost might be involved in placing them. Hence, there arises the problem of *optimal relay placement*. Such a problem involves the joint optimization of node placement and operation of the resulting network, where by "operation" we mean activities such as transmission scheduling, power allocation, and channel coding.

Our work in this paper is motivated by recent interest in problems of *impromptu (as-you-go)* deployment of wireless relay networks in various situations; for example, "first responders" in emergency situations, or quick deployment (and redeployment) of sensor networks in large terrains, such as forests (see [3], [4], [5], [6], [7]). In this paper, we are concerned with the situation in which a deployment agent walks *from the source node to the sink node*, along the line joining these two nodes, and places wireless relays (in an "as-you-go" manner) so as to create a source-to-sink multi-relay channel network with high data rate; see Figure 1. We first consider the scenario where the length L of the line in Figure 1 is known; the results of this case are used to formulate the as-you-go deployment in the case where L is a priori unknown, but has exponential distribution with known mean $\overline{L}$.

In order to capture the fundamental trade-offs in-

volved in such problems, we consider an information theoretic model. For a placement of the relay nodes and allocation of transmission powers to these relays, we model the "quality" of communication between the source and the sink by the information theoretic achievable rate of the multi-relay channel (see [8], [9] and [10] for the single and multi-relay channel models). The relays are equipped with full-duplex radios[1], and carry out decode-and-forward relaying. We consider scalar, memoryless, time-invariant, additive white Gaussian noise (AWGN) channels. We assume synchronous operation across all transmitters and receivers, and consider the exponential path-loss model for radio wave propagation.

1.1 Related Work

A formulation of the relay placement problem requires a model of the wireless network at the physical (PHY) and medium access control (MAC) layers. Most researchers have adopted the link scheduling and interference model, i.e., a scheduling algorithm determines radio resource allocation (channel and power) and interference is treated as noise (see [15]); treating interference as noise leads to the model that simultaneous transmissions "collide" at receiving nodes, and transmission scheduling aims to avoid collisions.

However, node placement for throughput maximization with this model is intractable because the optimal throughput is obtained by first solving for the optimum schedule assuming fixed node locations, followed by an optimization over those locations. Hence, with such a model, there appears to be little work on the problem of jointly optimizing the relay node placement and the transmission schedule. Reference [16] is one such work where the authors considered placing a set of nodes in an existing network such that a certain network utility is optimized subject to a set of linear constraints on link rates, under the link scheduling and interference model. They posed the problem as one of geometric programming assuming exponential path-loss, and proposed a distributed solution. The authors of [17] consider relay placement for utility maximization, assuming there are several source nodes, sink nodes and a few candidate locations for placing relays; they ignore interference because of highly directional antennas used in 60 GHz mmWave networks, which may not always be valid. Relay placement for capacity enhancement has been studied in [18], but there interference is mitigated by scheduling transmissions over multiple channels.

On the other hand, an information theoretic model for a wireless network often provides a closed-form expression for the channel capacity, or at least an achievable rate region. These results are asymptotic, and make idealized assumptions such as full-duplex radios, perfect interference cancellation, etc., but provide algebraic expressions that can be used to formulate tractable optimization problems which can provide useful insights. In the context of optimal relay placement, some researchers

1. Full-duplex radios are becoming practical; see [11], [12], [13], [14].

have already exploited this approach. Thakur et al., in [19], report on the problem of placing a single relay node to maximize the capacity of a broadcast relay channel in a wideband regime. Lin et al., in [20], numerically solve the problem of a single relay node placement, under power-law path loss and individual power constraints at the source and the relay; however, our work is primarily focused on multi-relay placement, under the exponential path-loss model and a sum power constraint among the nodes. The linear deterministic channel model ([21]) is used by Appuswamy et al. in [22] to study the problem of placing two or more relay nodes along a line so as to maximize the end-to-end data rate. Our present paper is in a similar spirit; however, we use the achievable rate formulas for the N-relay channel (with decode and forward relays) to study the problem of placing relays on a line having length L, under a sum power constraint over the nodes.

The most important difference of our paper with the literature reported above is that we address the problem of *sequential placement* of relay nodes along a line of an unknown random length. This paper extends our previous work in [1], which presents the analysis for the case of given L and N; the study under given L and N is a precursor to the formulation of as-you-go deployment problem, since it motivates an additive cost structure that is essential for the formulation of the sequential deployment problem as a Markov decision process (MDP).

The deploy-as-you-go problem has been addressed by previous researchers. For example, Howard et al., in [4], provide heuristic algorithms for incremental deployment of sensors in order to cover a deployment area. Souryal et al., in [3], propose heuristic deployment algorithms for the problem of impromptu wireless network deployment, with an experimental study of indoor RF link quality variation. The authors of [5] propose a collaborative deployment method for multiple deployment agents, so that the contiguous coverage area of relays is maximized subject to a total number of relays constraint. However, until the work in [23] and [6], there appears to have been no effort to rigorously formulate as-you-go deployment problem in order to derive optimal deployment algorithms. The authors of [23] and [6] used MDP based formulations to address the problem of placing relay nodes sequentially along a line and along a random lattice path, respectively. The formulations in [23] and [6] are based on the so-called "lone packet traffic model" under which, at any time instant, there can be no more than one packet traversing the network, thereby eliminating contention between wireless links. This work was later extended in [7] to the scenario where the traffic is still lone packet, but a measurement-based approach is employed to account for the spatial variation of link qualities due to shadowing.

In this paper, we consider as-you-go deployment along a line, but move away from the lone-packet traffic assumption by employing information theoretic achiev-

able rate formulas (for full-duplex radios and decode-and-forward relaying). We assume exponential path-loss model (see [24] and Section 6.1). To the best of our knowledge, there is no prior work that considers as-you-go deployment under this physical layer model.

1.2 Our Contribution

- **Optimal Offline Deployment:** Given the location of N full-duplex relays to connect a source and a sink separated by a given distance L, and under the exponential path-loss model and a sum power constraint among the nodes, the optimal power split among the nodes and the achievable rate are expressed (Theorem 1) in terms of the channel gains. We find expression for optimal relay location in the single relay placement problem (Theorem 2). For the N relay placement problem, numerical study shows that, the relay nodes are clustered near the source at low attenuation and are placed uniformly at high attenuation. Theorem 5 shows that, by placing large number of relays uniformly, we can achieve a rate arbitrarily close to the AWGN capacity. Only this part of our current paper was published in the conference version [1].
- **Optimal As-You-Go Deployment:** In Section 4, we consider the problem of placing relay nodes in a deploy-as-you-go manner, so as to connect a source and a sink separated by an unknown distance, modeled as an exponentially distributed random variable L. Specifically, the problem is to start from a source, and walk along a line, placing relay nodes as we go, until the line ends, at which point the sink is placed. With a sum power constraint, the aim is to maximize a capacity limiting term derived from the deployment problem for known L, while constraining the expected number of relays. We "relax" the expected number of relays constraint via a Lagrange multiplier, and formulate the problem as a total cost MDP with uncountable state space and non-compact action sets. We prove the existence of an optimal policy and convergence of value iteration (Theorem 7); these results for uncountable state space and non-compact action space are not evident from standard literature. We study properties of the value function analytically. This is the first time that the as-you-go deployment problem is formulated to maximize the end-to-end data rate under the full-duplex multi-relay channel model.
- **Numerical Results on As-You-Go Deployment:** In Section 5, we study the policy structure numerically. We also demonstrate numerically that the proposed as-you-go algorithm achieves an end-to-end data rate sufficiently close to the maximum possible achievable data rate for offline placement. This is particularly important since there is no other benchmark in the literature, with which we can make a fair comparison of our policy.
- The material in Section 4 and Section 5 were absent in the conference version [1].

1.3 Organization of the Paper

In Section 2, we describe our system model and notation. In Section 3, we address the problem of relay placement on a line of known length. Section 4 deals with the problem of as-you-go deployment along a line of unknown random length. Numerical work on as-you-go deployment has been presented in Section 5. Some discussions are provided in Section 6. Conclusions are drawn in Section 7.

2 SYSTEM MODEL AND NOTATION

2.1 The Multi-Relay Channel

The multi-relay channel was studied in [10] and [9] and is an extension of the single relay model presented in [8]. We consider a network deployed on a line with a source node and a sink node at the end of the line, and N full-duplex relay nodes as shown in Figure 1. The relay nodes are numbered as $1, 2, \cdots, N$. The source and sink are indexed by 0 and $N+1$, respectively. The distance of the k-th node *from the source* is denoted by $y_k := r_1 + r_2 + \cdots + r_k$. Thus, $y_{N+1} = L$. As in [10] and [9], we consider the *scalar, time-invariant, memoryless, AWGN setting.*

We use the model that a symbol transmitted by node i is received at node j after multiplication by the (positive, real valued) channel gain $h_{i,j}$ (an assumption often made in the literature, see e.g., [10] and [25]). The *power gain* from Node i to Node j is denoted by $g_{i,j} = h_{i,j}^2$. We define $g_{i,i} = 1$ and $h_{i,i} = 1$. The Gaussian additive noise at any receiver is independent and identically distributed from symbol to symbol and has variance σ^2.

2.2 An Inner Bound to the Capacity

For the multi-relay channel, we denote the symbol transmitted by the i-th node at time t (t is discrete) by $X_i(t)$ for $i = 0, 1, \cdots, N$. $Z_k(t) \sim \mathcal{N}(0, \sigma^2)$ is the additive white Gaussian noise at node k and time t, and is assumed to be *independent and identically distributed across k and t.* Thus, at symbol time t, node $k, 1 \leq k \leq N+1$ receives:

$$Y_k(t) = \sum_{j \in \{0,1,\cdots,N\}, j \neq k} h_{j,k} X_j(t) + Z_k(t) \qquad (1)$$

An inner bound to the capacity of this network, under any path-loss model, is given by (see [10]):

$$R = \min_{1 \leq k \leq N+1} C\left(\frac{1}{\sigma^2} \sum_{j=1}^{k} (\sum_{i=0}^{j-1} h_{i,k}\sqrt{P_{i,j}})^2\right) \qquad (2)$$

where $C(x) := \frac{1}{2}\log_2(1+x)$, and node i transmits to node j at power $P_{i,j}$ (expressed in mW).

In Appendix A, we provide a descriptive overview of the coding and decoding scheme proposed in [10]. A sequence of messages are sent from the source to the sink; each message is encoded in a block of symbols and transmitted by using the relay nodes. The scheme involves *coherent transmission* by the source and relay nodes (this requires *symbol-level synchronization among the nodes*), and *successive interference cancellation* at the relay nodes and the sink. A node receives information

about a message in two ways (i) by the message being directed to it cooperatively by all the previous nodes, and (ii) by overhearing previous transmissions of the message to the previous nodes. Thus node k receives codes corresponding to a message k times before it attempts to decode the message (a discussion on the practical feasibility of full-duplex decode-and-forward relaying scheme is provided in Section 6.3). Note that, $C\left(\frac{1}{\sigma^2}\sum_{j=1}^{k}(\sum_{i=0}^{j-1}h_{i,k}\sqrt{P_{i,j}})^2\right)$ in (2), for any k, denotes a possible rate that can be achieved by node k from the transmissions from nodes $0, 1, \cdots, k-1$. The smallest of these terms becomes the bottleneck, see (2).

For the single relay channel, $N = 1$. Thus, by (2), an achievable rate is given by (see also [8]):

$$R = \min \left\{ \; C\left(\frac{g_{0,1}P_{0,1}}{\sigma^2}\right), \right.$$
$$\left. C\left(\frac{g_{0,2}P_{0,1} + (h_{0,2}\sqrt{P_{0,2}} + h_{1,2}\sqrt{P_{1,2}})^2}{\sigma^2}\right) \right\} \quad (3)$$

Here, the first term in the $\min\{\cdot, \cdot\}$ of (3) is the achievable rate at node 1 (i.e., the relay node) due to the transmission from the source. The second term in the $\min\{\cdot, \cdot\}$ corresponds to the possible achievable rate at the sink node due to direct coherent transmission from the source and the relay and due to the overheard transmission from the source to the relay. The higher the channel attenuation, the less will be the contribution of farther nodes, "overheard" transmissions become less relevant, and coherent transmission reduces to a simple transmission from the previous relay. The system is then closer to simple store-and-forward relaying.

The authors of [10] have shown that any rate strictly less than R is achievable through the coding and decoding scheme. This achievable rate formula can also be obtained from the capacity formula of a physically degraded multi-relay channel (see [9]), since the capacity of the degraded relay channel is a lower bound to the actual channel capacity. *In this paper, we will seek to optimize R in (2) over power allocations to the nodes and the node locations, keeping in mind that R is a lower bound to the actual capacity. We denote the value of R optimized over power allocation and relay locations by R^*.*

2.3 Path-Loss Model

We model the power gain via the exponential path-loss model: the power gain at a distance r is $e^{-\rho r}$ where $\rho > 0$. This is a simple model used for tractability (see [16], [22]) and [26, Section 2.3] for prior work assuming exponential path-loss). However, for propagation scenarios involving randomly placed scatterers (as would be the case in a dense urban environment, or a forest, for example) analytical and experimental support has been provided for the exponential path-loss model (a discussion has been provided in Section 6.1). We also discuss in Section 6.4 how the insights obtained from the results for exponential path-loss can be used for power-law path-loss (power gain at a distance r is $r^{-\eta}$, $\eta > 0$).

Deployment with other path-loss models is left in this paper as a possible future work.

Under exponential path-loss, the channel gains and power gains in the line network become multiplicative, e.g., $h_{i,i+2} = h_{i,i+1}h_{i+1,i+2}$ and $g_{i,i+2} = g_{i,i+1}g_{i+1,i+2}$ for $i \in \{0, 1, \cdots, N-1\}$.

We discuss in Section 6.2 how shadowing and fading can be taken care of in our model, by providing a fade-margin in the power at each transmitter.

2.4 Motivation for the Sum Power Constraint

In this paper we consider the sum power constraint $\sum_{i=0}^{N} P_i = P_T$ (in mW) over the source and the relays. This constraint has the following motivation. Let the fixed power expended in a relay (for reception and driving the electronic circuits) be denoted by P_{rcv} (expressed in mW), and the initial battery energy in each node be denoted by E (in mJ unit). The information theoretic approach utilized in this paper requires that the nodes in the network are always on. Hence, the lifetime of node $i, 1 \leq i \leq N$, is $\tau_i = \frac{E}{P_i + P_{\text{rcv}}}$, the lifetime of the source is $\tau_0 = \frac{E}{P_0}$, and that of the sink is $\tau_{N+1} = \frac{E}{P_{\text{rcv}}}$. The rate of battery replacement at node i is $\frac{1}{\tau_i}$. Hence, the rate at which we have to replace the batteries in the network is $\sum_{i=0}^{N+1}\frac{1}{\tau_i} = \frac{1}{E}(\sum_{i=0}^{N} P_i + (N+1)P_{\text{rcv}})$. The depletion rate $\frac{P_{\text{rcv}}}{E}$ is inevitable at any node, and it does not affect the achievable data rate. Hence, in order to reduce the battery replacement rate, we must reduce the sum transmit power in the entire network.

3 PLACEMENT ON A LINE OF KNOWN LENGTH

As a precursor to addressing the deploy-as-you-go problem over a line of unknown length, in this section we solve the problem of power constrained deployment of *a given number of relays on a line of known length*. We will often refer to this problem as *offline deployment problem*. The results of this section provide (i) first insights into the relay placements we obtain using the multi-relay channel model, (ii) a starting point for the formulation of as-you-go deployment problem, and (iii) a benchmark with which we can compare the performance of our as-you-go deployment algorithm.

3.1 Optimal Power Allocation

In this section, we consider the optimal placement of relay nodes on a line of given length, L, so as to to maximize R (see (2)), subject to a total power constraint on the source and relay nodes given by $\sum_{i=0}^{N} P_i = P_T$. We will first maximize R in (2) over $P_{i,j}, 0 \leq i < j \leq (N+1)$ for any given placement of nodes (i.e., given $y_1, y_2, \cdots, y_N$). This will provide an expression of achievable rate in terms of channel gains, which has to be maximized over $y_1, y_2, \cdots, y_N$. Let $\gamma_k := \sum_{i=0}^{k-1} P_{i,k}$ for $k \in \{1, 2, \cdots, N+1\}$ (expressed in mW). Hence, the sum power constraint becomes $\sum_{k=1}^{N+1} \gamma_k = P_T$.

Theorem 1: (i) Under the exponential path-loss model, for fixed location of relay nodes, the optimal

power allocation that maximizes the achievable rate for the sum power constraint is given by:

$$P_{i,j} = \begin{cases} \frac{g_{i,j}}{\sum_{l=0}^{j-1} g_{l,j}} \gamma_j & \forall 0 \le i < j \le (N+1) \\ 0, & \text{if } j \le i \end{cases} \quad (4)$$

where

$$\gamma_1 = \frac{P_T}{1 + g_{0,1} \sum_{k=2}^{N+1} \frac{(g_{0,k-1}-g_{0,k})}{g_{0,k}g_{0,k-1} \sum_{l=0}^{k-1} \frac{1}{g_{0,l}}}} \quad (5)$$

$$\gamma_j = \frac{g_{0,1} \frac{(g_{0,j-1}-g_{0,j})}{g_{0,j}g_{0,j-1} \sum_{l=0}^{j-1} \frac{1}{g_{0,l}}}}{1 + g_{0,1} \sum_{k=2}^{N+1} \frac{(g_{0,k-1}-g_{0,k})}{g_{0,k}g_{0,k-1} \sum_{l=0}^{k-1} \frac{1}{g_{0,l}}}} P_T \quad \forall j \ge 2$$

(ii) The achievable rate optimized over the power allocation for a given placement of nodes is given by:

$$R_{P_T}^{opt}(y_1, y_2, \cdots, y_N) = C\left(\frac{\frac{P_T}{\sigma^2}}{\frac{1}{g_{0,1}} + \sum_{k=2}^{N+1} \frac{(g_{0,k-1}-g_{0,k})}{g_{0,k}g_{0,k-1} \sum_{l=0}^{k-1} \frac{1}{g_{0,l}}}} \right) \quad (6)$$

Proof: The basic idea is to choose the power levels (i.e., $P_{i,j}$, $0 \le i < j \le N+1$) in (2) so that all the terms in the $\min\{\cdot\}$ in (2) become equal. We provide explicit expressions for $P_{i,j}$, $0 \le i < j \le N+1$ and the achievable rate (optimized over power allocation) in terms of the power gains. See Appendix B for the detailed proof. A result on the equality of certain terms under optimal power allocation has also been proved in [9] for the coding scheme used in [9]. But it was proved in the context of a degraded Gaussian multi-relay channel, and the proof depends on an inductive argument, whereas our proof utilizes LP (linear programming) duality. □

Recalling the exponential path-loss parameter ρ, and the source-sink distance L, let us define $\lambda := \rho L$, which can be treated as a measure of attenuation in the line.

Let us now comment on the results of Theorem 1:

- In order to maximize $R_{P_T}^{opt}(y_1, y_2, \cdots, y_N)$, we need to place the relay nodes such that $\frac{1}{g_{0,1}} + \sum_{k=2}^{N+1} \frac{(g_{0,k-1}-g_{0,k})}{g_{0,k}g_{0,k-1} \sum_{l=0}^{k-1} \frac{1}{g_{0,l}}}$ is minimized. This quantity is viewed as the net attenuation the power P_T faces.
- When no relay is placed, the attenuation is e^λ. The ratio of attenuation with no relay and attenuation with relays is called the "relaying gain" $G(N, \lambda)$.

$$G(N, \lambda) := \frac{e^\lambda}{\frac{1}{g_{0,1}} + \sum_{k=2}^{N+1} \frac{(g_{0,k-1}-g_{0,k})}{g_{0,k}g_{0,k-1} \sum_{l=0}^{k-1} \frac{1}{g_{0,l}}}} \quad (7)$$

Rate is increasing with the number of relays, and is bounded above by $C(\frac{P_T}{\sigma^2})$. Hence, $G(N, \lambda) \in [1, e^\lambda]$. Also, note that $G(N, \lambda)$ does not depend on P_T.

- By Theorem 1, we have $P_{k,j} \ge P_{i,j}$ for $i < k < j$.
- Note that we have derived Theorem 1 using the fact that $g_{0,k}$ is nonincreasing in k. If there exists some $k \ge 1$ such that $g_{0,k} = g_{0,k+1}$, i.e, if k-th and $(k+1)$-st nodes are placed at the same position, then $\gamma_{k+1} = 0$, i.e., the nodes $i < k$ do not direct any power specifically to relay $k+1$. However, relay $k+$
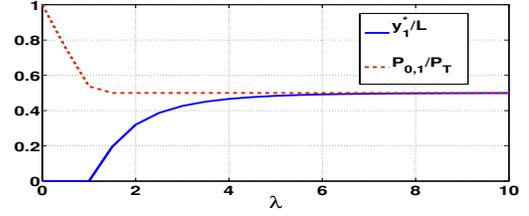
Figure 2: Single relay placement, total power constraint, exponential path-loss: $\frac{y_1^*}{L}$ and optimum $\frac{P_{0,1}}{P_T}$ versus λ.

1 can decode the symbols received at relay k, and those transmitted by relay k. Then relay $(k+1)$ can transmit coherently with the nodes $l \le k$ to improve effective received power in the nodes $j > k+1$.

3.2 Optimal Placement of a Single Relay Node

In the following theorem, we derive the optimal power allocation, node location and data rate when a single relay is placed.

Theorem 2: For the single relay node placement problem with sum power constraint and exponential path-loss model, the normalized optimum relay location $\frac{y_1^*}{L}$, power allocation and optimized achievable rate are given as follows:[2]

(i) For $\lambda \le \log 3$, $\frac{y_1^*}{L} = 0$, $P_{0,1} = \frac{2P_T}{e^\lambda+1}$, $P_{0,2} = P_{1,2} = \frac{e^\lambda-1}{e^\lambda+1}\frac{P_T}{2}$ and $R^* = C\left(\frac{2P_T}{(e^\lambda+1)\sigma^2}\right)$.

(ii) For $\lambda \ge \log 3$, $\frac{y_1^*}{L} = \frac{1}{\lambda}\log\left(\sqrt{e^\lambda+1}-1\right)$, $P_{0,1} = \frac{P_T}{2}$, $P_{0,2} = \frac{1}{\sqrt{e^\lambda+1}}\frac{P_T}{2}$, $P_{1,2} = \frac{\sqrt{e^\lambda+1}-1}{\sqrt{e^\lambda+1}}\frac{P_T}{2}$ and $R^* = C\left(\frac{1}{\sqrt{e^\lambda+1}-1}\frac{P_T}{2\sigma^2}\right)$

Proof: See Appendix C. □

Discussion: It is easy to check that R^* obtained in Theorem 2 is strictly greater than the AWGN capacity $C\left(\frac{P_T}{\sigma^2}e^{-\lambda}\right)$ for all $\lambda > 0$. This happens because the source and relay transmit coherently to the sink. R^* becomes equal to the AWGN capacity only at $\lambda = 0$. At $\lambda = 0$, we do not use the relay since the sink can decode any message that the relay is able to decode.

The variation of $\frac{y_1^*}{L}$ and $\frac{P_{0,1}}{P_T}$ with λ has been shown in Figure 2. We observe that (from Figure 2 and Theorem 2) $\lim_{\lambda\to\infty} \frac{y_1^*}{L} = \frac{1}{2}$, $\lim_{\lambda\to\infty} P_{0,2} = 0$ and $\lim_{\lambda\to 0} P_{0,1} = P_T$. For large values of λ, source and relay cooperation provides negligible benefit since source to sink attenuation is very high. So it is optimal to place the relay at a distance $\frac{L}{2}$. The relay works as a repeater which forwards data received from the source to the sink. For small λ, the gain obtained from coherent transmission is dominant, and, in order to receive sufficient information (required for coherent transmission) from the source, the relay is placed near the source.

3.3 Optimal Placement for a Multi-Relay Channel

As discussed earlier, we need to place N relay nodes such that $\frac{1}{g_{0,1}} + \sum_{k=2}^{N+1} \frac{(g_{0,k-1}-g_{0,k})}{g_{0,k}g_{0,k-1} \sum_{l=0}^{k-1} \frac{1}{g_{0,l}}}$ is minimized

---

2. $\log(\cdot)$ in this paper will mean the natural logarithm unless the base is specified.

(a) $N = 2$ (b) $N = 3$ (c) $N = 5$
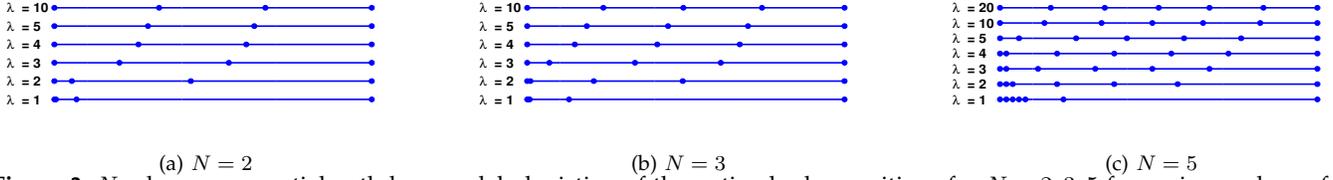
Figure 3: N relays, exponential path-loss model: depiction of the optimal relay positions for $N = 2, 3, 5$ for various values of λ. On any line, the leftmost dot is the source node, the rightmost dot is the sink, and the intermediate dots denote the relays.

Here $g_{0,k} = e^{-\rho y_k}$. We have the constraint $0 \leq y_1 \leq y_2 \leq \cdots \leq y_N \leq y_{N+1} = L$. Now, writing $z_k = e^{\rho y_k}$, and defining $z_0 := 1$, we arrive at the following problem:

$$\min \left\{ z_1 + \sum_{k=2}^{N+1} \frac{z_k - z_{k-1}}{\sum_{l=0}^{k-1} z_l} \right\}$$
$$s.t \quad 1 \leq z_1 \leq \cdots \leq z_N \leq z_{N+1} = e^{\lambda} \qquad (8)$$

The objective function is convex in each of the variables $z_1, z_2, \cdots, z_N$. The objective function is sum of linear fractionals, and the constraints are linear.

Remark: From optimization problem (8) we observe that optimum $z_1, z_2, \cdots, z_N$ depend only on $\lambda := \rho L$. Since $z_k = e^{\lambda \frac{y_k}{L}}$, the normalized optimal distance of relays from the source depend only on λ and N.

Theorem 3: For fixed ρ, L and σ^2, the optimized achievable rate R^* for a sum power constraint *strictly* increases with the number of relay nodes.

Proof: See Appendix D.   □

Theorem 4: For any fixed number of relays $N \geq 1$, $G(N, \lambda)$ is increasing in λ.

Proof: See Appendix E.   □

A numerical study of multi-relay placement: We discretize the interval $[0, L]$ and run a search program to find normalized optimal relay locations for different values of λ and N. The results are summarized in Figure 3.

We observe that at low attenuation (small λ), relay nodes are clustered near the source node and are often at the source node, whereas at high attenuation (large λ) they are almost uniformly placed along the line. For large λ, the effect of long distance between any two adjacent nodes dominates the gain obtained by coherent relaying. Hence, it is beneficial to minimize the maximum distance between any two adjacent nodes and thus multihopping is a better strategy in this case. For small λ, the gain obtained by coherent transmission is dominant. In order to allow this, relays should be able to receive sufficient information from their previous nodes. Thus, they tend to be clustered near the source.

In Figure 4 we plot the relaying gain $G(N, \lambda)$ in dB vs. the number of relays N, for various values of λ. As proved in Theorem 3, we see that $G(N, \lambda)$ increases with N for fixed λ. On the other hand, $G(N, \lambda)$ increases with λ for fixed N, as proved in Theorem 4.

3.4 Uniformly Placed Relays, Large N

When the relays are uniformly placed, the behaviour of $R_{P_T}^{opt}(y_1, \cdots, y_N)$ (called R_N in the next theorem) for
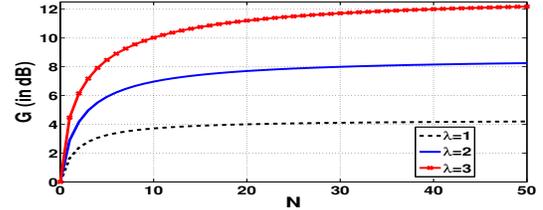
Figure 4: $G(N, \lambda)$ vs N for total power constraint.

large number of relays is captured by the following:

Theorem 5: For exponential path-loss and sum power constraint, if N relay nodes are placed uniformly between the source and the sink, resulting in R_N achievable rate, then $\lim_{N \to \infty} R_N = C\left(\frac{P_T}{\sigma^2}\right)$.

Proof: See Appendix F.   □

Remark: From Theorem 5, it is clear that we can achieve a rate arbitrarily close to $C(\frac{P_T}{\sigma^2})$ (i.e., the effect of path-loss can completely be mitigated) by placing a large enough number of relay nodes. In this context, we would like to mention that the variation of broadcast capacity as a function of the number of nodes N (located randomly inside a unit square) was studied in [27]; but the broadcast capacity in their paper increases with N since they assume per-node power constraint.

4 AS-YOU-GO DEPLOYMENT OF RELAYS ON A LINE OF UNKNOWN LENGTH

Having developed the problem of placing a given number of relays over a line of fixed, given length, we now turn to the deploy-as-you-go problem. An agent walks along a line, starting from the source and heading towards the sink which is at an unknown distance from the source location, deploying relays as he goes, so as to achieve a multi-relay network when he encounters the sink location (and places the sink there). We model the distance from the source to sink as an exponentially distributed random variable with mean $\overline{L} = \frac{1}{\beta}$.[3] The deployment objective is to achieve a high data rate from the source to the sink, subject to a total power constraint and a constraint on the expected number of relays placed (note that, the number of relay nodes, N, is a random variable here, due to the randomness in L). Using the rate expression from Theorem 1, we formulate the problem as a total cost MDP.

---

3. A motivation for the use of the exponential distribution, given the prior knowledge of the mean length $\overline{L}$, is that it is the maximum entropy continuous probability density function with the given mean. By using the exponential distribution, we are leaving the length of the line as uncertain as we can, given the prior knowledge of its mean.

Such a deployment problem could be motivated by a situation where it is required to place a sensor (say, a video camera) to monitor an event or an object from a safe distance (e.g., the battlefront in urban combat, or a suspicious object that needs to be detonated, or a group of animals in a forest). In such a situation, the deployment agent, after placing the sensor, walk away from the scene of the event, along a forest trail, or a road, or a building corridor, placing relays as he walks, until a suitable safe sink location is found, in such a way that the number of relays is kept small while the end-to-end data rate is maximized.

4.1 Formulation as an MDP

We now formulate the as-you-go deployment problem as an MDP.

4.1.1 Deployment Policies

In the as-you-go placement problem, the person carries a number of nodes and places them as he walks, under the control of a placement policy. A deployment policy π is a sequence of mappings $\{\mu_1, \mu_2, \mu_3, \cdots\}$ from the state space to the action space; at the k-th decision instant (i.e., after placing the $(k-1)$-st relay), μ_k provides the distance at which the next relay should be placed (provided that the line does not end before that point), given the system state which is a function of the locations of previously placed nodes. Thus, the decisions are made based on the locations of the relays placed earlier. The first decision instant is the start of the line, and the subsequent decision instants are the placement points of the relays. Let Π denote the set (possibly uncountable) of all deployment policies. Let $\mathbb{E}_\pi$ denote the expectation under policy π.

4.1.2 The Unconstrained Problem

We recall from (6) that for a fixed length L of the line and a fixed N, $e^{\rho y_1} + \sum_{k=2}^{N+1} \frac{e^{\rho y_k} - e^{\rho y_{k-1}}}{1 + e^{\rho y_1} + \cdots + e^{\rho y_{k-1}}}$ has to be minimized in order to maximize $R_{P_T}^{opt}(y_1, y_2, \cdots, y_N)$. $e^{\rho y_1} + \sum_{k=2}^{N+1} \frac{e^{\rho y_k} - e^{\rho y_{k-1}}}{1 + e^{\rho y_1} + \cdots + e^{\rho y_{k-1}}}$ is basically a scaling factor which captures the effect of attenuation and relaying on the maximum possible SNR $\frac{P_T}{\sigma^2}$.

Let $\xi > 0$ be the cost of placing a relay. We are interested in solving the following problem:

$$\inf_{\pi \in \Pi} \mathbb{E}_\pi \left(\left(e^{\rho y_1} + \sum_{k=2}^{N+1} \frac{e^{\rho y_k} - e^{\rho y_{k-1}}}{1 + e^{\rho y_1} + \cdots + e^{\rho y_{k-1}}} \right) + \xi N \right) \quad (9)$$

The "cost" function inside the outer parentheses in (9) has two terms, one is the denominator of $G(N, \lambda)$ in (7), and the other is a linear multiple of the number of relays. Thus, the cost function captures the tradeoff between the cost of placing relays (quantified as ξ per relay), and the need to achieve high end-to-end data rate by making the denominator of $G(N, \lambda)$ small. Note that, due to the

randomness in the length of the line, the $y_k, k \geq 1$, and N are all random variables.[4]

We will see in Theorem 7 that an optimal policy always exists for this problem.

4.1.3 The Constrained Problem

Solving the problem in (9) also helps in solving the following constrained problem:

$$\inf_{\pi \in \Pi} \mathbb{E}_\pi \left(e^{\rho y_1} + \sum_{k=2}^{N+1} \frac{e^{\rho y_k} - e^{\rho y_{k-1}}}{1 + e^{\rho y_1} + \cdots + e^{\rho y_{k-1}}} \right)$$
$$s.t., \quad \mathbb{E}_\pi(N) \leq M \quad (10)$$

where $M > 0$ is a constraint on the expected number of relays. [5] The following standard result (see [28, Theorem 4.3]) gives the optimal ξ^*:

Lemma 1: If there exists $\xi^* > 0$ and a policy $\pi^*_{\xi^*} \in \Pi$ such that $\pi^*_{\xi^*}$ is an optimal policy for the unconstrained problem (9) under ξ^* and $\mathbb{E}_{\pi^*_{\xi^*}} N = M$, then $\pi^*_{\xi^*}$ is also optimal for the constrained problem (10).

The motivation behind formulation (10) is as follows. Suppose that one seeks to solve the following problem:

$$sup_\pi \mathbb{E}_\pi \log_2 \left(1 + \frac{\frac{P_T}{\sigma^2}}{e^{\rho y_1} + \sum_{k=2}^{N+1} \frac{e^{\rho y_k} - e^{\rho y_{k-1}}}{1 + e^{\rho y_1} + \cdots + e^{\rho y_{k-1}}}} \right)$$
$$s.t., \mathbb{E}_\pi(N) \leq M \quad (11)$$

Since $\log_2(1 + \frac{1}{x})$ is convex in x, we can argue by Jensen's inequality that by solving (10) we actually find a relay placement policy that maximizes a lower bound to the expected achievable data rate obtained from (11). But formulation (10) (and hence formulation (9), by Lemma 1) allows us to write the objective function as a summation of hop-costs; this motivates us to formulate the as-you-go deployment problem as an MDP, resulting in a substantial reduction in policy computation. However, in Section 5, we will show numerically that solving (10) is a reasonable approach to deal with the computational complexity of (11); we will see that formulation (10) allows us to achieve a reasonable performance.

We now formulate the above "as-you-go" relay placement problem (9) as a total cost Markov decision process.

4.1.4 State Space, Action Space and Cost Structure

Let us define $s_0 := 1$, $s_k := \frac{e^{\rho y_k}}{1 + e^{\rho y_1} + \cdots + e^{\rho y_k}} \, \forall \, k \geq 1$. Also, recall that $r_{k+1} = y_{k+1} - y_k$. Thus, we can rewrite (9) as:

$$\inf_{\pi \in \Pi} \mathbb{E}_\pi \left(1 + \sum_{k=0}^{N} s_k(e^{\rho r_{k+1}} - 1) + \xi N \right) \quad (12)$$

When the person starts walking from the source along the line, the state of the system is set to $s_0 := 1$. At

---

4. Recall Section 2.4. The battery depletion rate $\frac{P_{rcv}}{E}$ of a node due to the receive power alone can be absorbed into the relay cost ξ.

5. The constraint on the mean number of relays can be justified if we consider the relay deployment problem for multiple source-sink pairs over several lines of mean length $\overline{L}$, given a large pool of relays, and we are only interested in keeping small the total number of relays over all these deployments.

this instant the placement policy provides the location at which the first relay should be placed. The person walks towards the prescribed placement point. If the sink placement location is encountered before reaching this point, the sink is placed; if not, then the first relay is placed at the placement point. In general, the state after placing the k-th relay is denoted by s_k (a function of the location of the nodes up to the k-th instant), for $k = 1, 2, \cdots$. At state s_k, the action is the distance r_{k+1} where the next relay has to be placed (action ∞ means that no further relay will be placed). If the line ends before this distance, the sink node has to be placed at the end. *The randomness is coming from the random residual length of the line.* Let l_k denote the residual length at the k-th instant.

With this notation, the state of the system evolves as:

$$s_{k+1} = \begin{cases} \frac{s_k e^{\rho r_{k+1}}}{1 + s_k e^{\rho r_{k+1}}}, & \text{if } l_k > r_{k+1}, \\ \textbf{EOL}, & \text{else.} \end{cases} \quad (13)$$

Here **EOL** denotes the end of the line, i.e., the termination state.

The single stage cost (for problem (12)) for state s, action a and residual length l, is:

$$c(s, a, l) = \begin{cases} \xi + s(e^{\rho a} - 1), & \text{if } l > a, \\ s(e^{\rho l} - 1), & \text{else.} \end{cases} \quad (14)$$

Also, $c(\textbf{EOL}, a, \cdot) = 0$ for all a.

From (13), it is clear that the next state s_{k+1} depends on the current state s_k, the current action r_{k+1} and the residual length of the line. Since the length of the line is exponentially distributed, from any placement point, the residual line length is exponentially distributed, and independent of the history of the process. The cost incurred at the k-th decision instant is given by (14), which depends on s_k, r_{k+1} and l_k.

Hence, our formulation in (12) is an MDP with state space $\mathcal{S} := (0, 1] \cup \{\textbf{EOL}\}$ and action space $\mathcal{A} \cup \{\infty\}$ where $\mathcal{A} := [0, \infty)$.

Remark: An optimal policy (if it exists) for the problem (9) will be used to place relay nodes along a line whose length is a sample from an exponential distribution with mean $\frac{1}{\beta}$. After the deployment is over, the power P_T will be shared optimally among the source and the deployed relay nodes (according to Theorem 1).

4.2 Optimal Value Function

Suppose $s_k = s$ for some $k \geq 0$. Then, the optimal value function (cost-to-go) at state s is defined by:

$$J_\xi(s) = \inf_{\pi \in \Pi} \mathbb{E}\left(\sum_{n=k}^{\infty} c(s_n, a_n, l_n) | s_k = s\right)$$

If we decide to place the next relay at a distance $a < \infty$ and follow the optimal policy thereafter, the expected cost-to-go at a state $s \in (0, 1]$ becomes:

$$\int_0^a \beta e^{-\beta z} s(e^{\rho z} - 1)dz + e^{-\beta a}\left(s(e^{\rho a} - 1) + \xi + J_\xi\left(\frac{s e^{\rho a}}{1 + s e^{\rho a}}\right)\right) \quad (15)$$

The first term in (15) corresponds to the case in which the line ends at a distance less than a and we are forced to place the sink node. The second term corresponds to the case where the residual length of the line is greater than a and a relay is placed at a distance a.

Note that our MDP has an uncountable state space $\mathcal{S} = (0, 1] \cup \{\textbf{EOL}\}$ and a non-compact action space $\mathcal{A} = [0, \infty) \cup \{\infty\}$. Several technical issues arise in this kind of problems, such as the existence of optimal or ϵ-optimal policies, measurability of the policies, etc. We, therefore, invoke the results provided by Schäl [29], which deal with such issues. Our problem is one of minimizing total, undiscounted, non-negative costs over an infinite horizon. Equivalently, in the context of [29], we have a problem of total reward maximization where the rewards are the negative of the costs. Thus, our problem specifically fits into the negative dynamic programming setting of [29] (i.e., the N case where single-stage rewards are non-positive).

Now, the state **EOL** is absorbing. Also, no action is taken at this state and the cost at this state is 0. Hence, we can think of this state as state 0 in order to make our state space a Borel subset of the real line.

Theorem 6: [[29], Equation (3.6)] The optimal value function $J_\xi(\cdot)$ satisfies the Bellman equation. $\square$

Thus, $J_\xi(\cdot)$ satisfies the following Bellman equation for each $s \in (0, 1]$:

$$\begin{aligned} J_\xi(s) = \min\Bigg\{ & \inf_{a \geq 0}\Bigg[\int_0^a \beta e^{-\beta z} s(e^{\rho z} - 1)dz \\ & + e^{-\beta a}\left(s(e^{\rho a} - 1) + \xi + J_\xi\left(\frac{s e^{\rho a}}{1 + s e^{\rho a}}\right)\right)\Bigg], \\ & \int_0^\infty \beta e^{-\beta z} s(e^{\rho z} - 1)dz\Bigg\} \end{aligned} \quad (16)$$

where the second term inside $\min\{\cdot, \cdot\}$ is the cost of not placing any relay (i.e., $a = \infty$).

We analyze the MDP for $\beta > \rho$ and $\beta \leq \rho$.

4.2.1 Case I ($\beta > \rho$)

We observe that the cost of not placing any relay (i.e., $a = \infty$) at state $s \in (0, 1]$ is given by:

$$\int_0^\infty \beta e^{-\beta z} s(e^{\rho z} - 1)dz = \theta s$$

where $\theta := \frac{\rho}{\beta - \rho}$ (using the fact that $\beta > \rho$). Since not placing a relay (i.e., $a = \infty$) is a possible action for every s, it follows that $J_\xi(s) \leq \theta s$.

The cost in (15), upon simplification, can be written as:

$$\theta s + e^{-\beta a}\left(-\theta s e^{\rho a} + \xi + J_\xi\left(\frac{s e^{\rho a}}{1 + s e^{\rho a}}\right)\right) \quad (17)$$

Since $J_\xi(s) \leq \theta$ for all $s \in (0, 1]$, the expression in (17) is strictly less that θs for large enough $a < \infty$. Hence, according to (16), it is not optimal to not place any relay and the Bellman equation (16) can be rewritten as:

$$J_\xi(s) = \theta s + \inf_{a \geq 0} e^{-\beta a}\left(-\theta s e^{\rho a} + \xi + J_\xi\left(\frac{s e^{\rho a}}{1 + s e^{\rho a}}\right)\right) \quad (18)$$

4.2.2 Case II ($\beta \leq \rho$)

Here the cost in (15) is ∞ if we do not place a relay (i.e., if $a = \infty$). Let us consider a policy π_1 where we place the next relay at a fixed distance $0 < a < \infty$ from the current relay, irrespective of the current state. If the residual length of the line is z at any state s, we will place less than $\frac{z}{a}$ additional relays, and for each relay a cost less than $(\xi + (e^{\rho a} - 1))$ is incurred (since $s \leq 1$). At the last step when we place the sink, a cost less than $(e^{\rho a} - 1)$ is incurred. Thus, the value function of this policy is upper bounded by:

$$\int_0^\infty \beta e^{-\beta z} \frac{z}{a} (\xi + (e^{\rho a} - 1)) dz + (e^{\rho a} - 1)$$
$$= \frac{1}{\beta a} (\xi + (e^{\rho a} - 1)) + (e^{\rho a} - 1) \qquad (19)$$

Hence, $J_\xi(s) \leq \frac{1}{\beta a} (\xi + (e^{\rho a} - 1)) + (e^{\rho a} - 1) < \infty$. Thus, by the same argument as in the case $\beta > \rho$, the minimizer in the Bellman equation lies in $[0, \infty)$, i.e., the optimal placement distance lies in $[0, \infty)$. Hence, (16) can be rewritten as:

$$J_\xi(s) = \inf_{a \geq 0} \left\{ \int_0^a \beta e^{-\beta z} s(e^{\rho z} - 1) dz + e^{-\beta a} \left(s(e^{\rho a} - 1) + \xi + J_\xi \left(\frac{s e^{\rho a}}{1 + s e^{\rho a}} \right) \right) \right\} \qquad (20)$$

4.3 Upper Bound on the Optimal Value Function

Proposition 1: If $\beta > \rho$, then $J_\xi(s) < \theta s$ for all $s \in (0, 1]$.

Proof: We know that $J_\xi(s) \leq \theta s \leq \theta$. Now, let us consider the Bellman equation (18). It is easy to see that $(-\theta s e^{\rho a} + \xi + J_\xi(\frac{s e^{\rho a}}{1 + s e^{\rho a}}))$ is strictly negative for sufficiently large a. Hence, the R.H.S of (18) is strictly less than θs. $\square$

Corollary 1: For $\beta > \rho$, $\lim_{s \to 0} J_\xi(s) \to 0$ for any $\xi > 0$.

Proof: Follows from Proposition 1. $\square$

Proposition 2: If $\beta > 0$ and $\rho > 0$ and $0 < a < \infty$, then $J_\xi(s) < \frac{1}{\beta a} (\xi + (e^{\rho a} - 1)) + (e^{\rho a} - 1)$ for all $s \in (0, 1]$.

Proof: Follows from (19), since the analysis is valid even for $\beta > \rho$. $\square$

4.4 Convergence of the Value Iteration

The value iteration for our MDP is given by:

$$J_\xi^{(k+1)}(s) = \inf_{a \geq 0} \left\{ \int_0^a \beta e^{-\beta z} s(e^{\rho z} - 1) dz + e^{-\beta a} \left(s(e^{\rho a} - 1) + \xi + J_\xi^{(k)} \left(\frac{s e^{\rho a}}{1 + s e^{\rho a}} \right) \right) \right\} \qquad (21)$$

Here $J_\xi^{(k)}(s)$ is the k-th iterate of the value iteration. Let us start with $J_\xi^{(0)}(s) := 0$ for all $s \in (0, 1]$. We set $J_\xi^{(k)}(\mathbf{EOL}) = 0$ for all $k \geq 0$. $J_\xi^{(k)}(s)$ is the optimal value function for a problem with the same single-stage cost and the same transition structure, but with the horizon length being k (instead of infinite horizon as in our original problem) and 0 terminal cost. Here, by horizon length k, we mean that there are k number of relays available for deployment.

Let $\Gamma_k(s)$ be the set of minimizers of (21) at the k-th iteration at state s, if the infimum is achieved at some $a < \infty$. Let $\Gamma_\infty(s) := \{a \in \mathcal{A} : a$ be an accumulation point of some sequence $\{a_k\}$ where each $a_k \in \Gamma_k(s)\}$. Let $\Gamma^*(s)$ be the set of minimizers in (20). In Appendix G, we show that $\Gamma_k(s)$ for each $k \geq 1$, $\Gamma_\infty(s)$ and $\Gamma^*(s)$ are nonempty.

Theorem 7: The value iteration given by (21) has the following properties:

(i) $J_\xi^{(k)}(s) \to J_\xi(s)$ for all $s \in (0, 1]$, i.e., the value iteration converges to the optimal value function.

(ii) $\Gamma_\infty(s) \subset \Gamma^*(s)$.

(iii) There is a stationary optimal policy $f^\infty = \{f, f, f, \cdots\}$ where $f : (0, 1] \to \mathcal{A}$ and $f(s) \in \Gamma_\infty(s)$ for all $s \in (0, 1]$.

Proof: The proof is given in Appendix G. It uses some results from [29], which have been discussed first. Next, we provide a general theorem (Theorem 11) on the convergence of value iteration, which has been used to prove Theorem 7. $\square$

Remark: Since the action space is noncompact, it is not obvious from standard results whether the optimal policy exists. However, we are able to show that in our problem, for each state $s \in (0, 1]$, the optimal action will lie in a compact set of the from $[0, a(s)]$, where $a(s)$ is continuous in s, and $a(s)$ could possibly go to ∞ as $s \to 0$. The results of [29] allow us to work with the scenario where for each state s, it is sufficient to focus only on a compact action space $[0, a(s)]$.

4.5 Properties of the Value Function $J_\xi(s)$

Proposition 3: $J_\xi(s)$ is increasing and concave over $s \in (0, 1]$.

Proposition 4: $J_\xi(s)$ is increasing and concave in ξ for all $s \in (0, 1]$.

Proposition 5: $J_\xi(s)$ is continuous in s over $(0, 1]$ and continuous in ξ over $(0, \infty)$.

See Appendix H for the proofs of these propositions.

4.6 A Useful Normalization

Note that, βL is exponentially distributed with mean 1. Defining $\Lambda := \frac{\rho}{\beta}$ and $\tilde{z}_k := \beta y_k$, $k = 1, 2, \cdots, (N+1)$, we can rewrite (9) as follows:

$$\inf_{\pi \in \Pi} \mathbb{E}_\pi \left(e^{\Lambda \tilde{z}_1} + \sum_{k=2}^{N+1} \frac{e^{\Lambda \tilde{z}_k} - e^{\Lambda \tilde{z}_{k-1}}}{1 + e^{\Lambda \tilde{z}_1} + \cdots + e^{\Lambda \tilde{z}_{k-1}}} + \xi N \right) \qquad (22)$$

Note that, Λ plays the same role as λ played in the known L case (see Section 3.2). Since $\frac{1}{\beta}$ is the mean length of the line, Λ can be considered as a measure of attenuation in the network. We can think of the new problem (22) in the same way as (9), but with the length of the line being exponentially distributed with mean 1 ($\beta' = 1$) and the path-loss exponent being changed to $\rho' = \Lambda = \frac{\rho}{\beta}$. The relay locations are also normalized ($\tilde{z}_k = \beta y_k$). One can solve the new problem (22) and obtain the optimal policy. Then the solution to (9) can be obtained by multiplying each control distance (from the optimal policy of (22)) with the constant $\frac{1}{\beta}$. Hence, it suffices to work with $\beta = 1$.
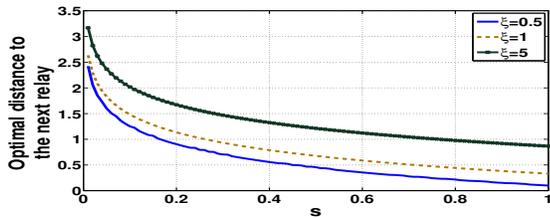
Figure 5: $\beta = 1$, $\Lambda := \frac{\rho}{\beta} = 2$; a^* vs. s.
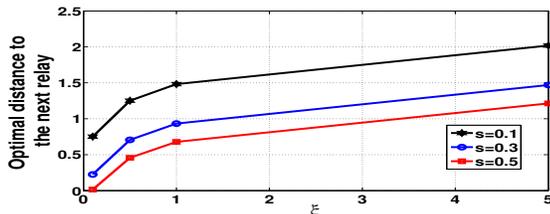
Figure 6: $\beta = 1$, $\Lambda := \frac{\rho}{\beta} = 2$; a^* vs. ξ.
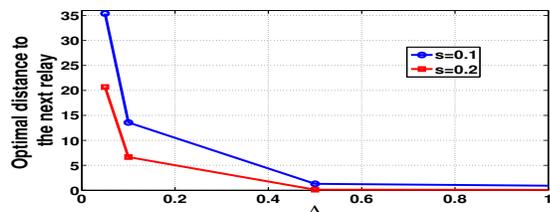
Figure 7: $\beta = 1$, $\xi = 0.01$; a^* vs. Λ.

5 A NUMERICAL STUDY OF AS-YOU-GO DEPLOYMENT

Let us recall that the state of the system after placing the k-th relay is given by $s_k = \frac{e^{\Lambda \tilde{z}_k}}{\sum_{i=0}^{k} e^{\Lambda \tilde{z}_i}}$. The action is the normalized distance of the next relay to be placed from the current location. The single stage cost function for our total cost minimization problem is given by (14).

In our numerical work, we discretized the state space $(0, 1]$ into 100 steps as $\{0.01, 0.02, \cdots, 0.99, 1\}$, and discretized the action space into steps of size 0.001, i.e., the action space becomes $\{0, 0.001, 0.002, \cdots\}$.

5.1 Structure of the Optimal Policy

We performed numerical experiments to study the structure of the optimal policy obtained through value iteration for $\beta = 1$ and some values of $\Lambda := \frac{\rho}{\beta}$. The value iteration in these experiments converged and we obtained a stationary optimal policy, though Theorem 7 does not guarantee the uniqueness of the stationary optimal policy.

Figure 5 shows that the *normalized* optimal placement distance a^* is decreasing with the state $s \in (0, 1]$. This can be understood as follows. The state s (at a placement point) is small only if a sufficiently large number of relays have already been placed.[6] Hence, if several relays have already been placed and $\sum_{i=0}^{k} e^{\Lambda \tilde{z}_i}$ is sufficiently large compared to $e^{\Lambda \tilde{z}_k}$ (i.e., s_k is small), the $(k + 1)$-st relay will be able to receive sufficient amount of power

---

6. $\frac{e^{\Lambda \tilde{z}_k}}{\sum_{i=0}^{k} e^{\Lambda \tilde{z}_i}} \geq \frac{e^{\Lambda \tilde{z}_k}}{(k+1)e^{\Lambda \tilde{z}_k}} = \frac{1}{k+1}$; hence, if s_k is small, k must be large enough.

Table 1: Sequential placement on a line of length 10 for various Λ, using the corresponding optimal policies for $\xi = 0.001$.

Λ	Normalised Optimal distances of the nodes from the source	No. of relays
0.01	0, 0, 8.4180, 10.0000	3
0.1	0, 0, 0, 0, 0, 0, 0, 0.2950, 0.5950, 0.9810, 1.3670, 1.7530, 2.1390, $\cdots$, 9.0870, 9.4730, 9.8590, 10.0000	33
5	0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.0020, 0.0080, 0.0140, 0.0200, $\cdots$, 9.9860, 9.9920, 9.9980, 10.0000	1677

Table 2: Evolution of state in the process of sequential placement on a line of length 10 for various values of Λ, using the corresponding optimal policies for $\xi = 0.001$.

Λ	Evolution of state in the process of sequential placement
0.01	1, 0.5, 0.34, 0.27
0.1	1, 0.5, 0.34, 0.26, 0.21, 0.18, 0.16, 0.14, 0.13, 0.12, 0.12, $\cdots$
5	1, 0.5, 0.34, 0.26, 0.21, 0.18, 0.16, 0.14, 0.13, 0.12, 0.11, 0.1, 0.1, 0.1, $\cdots$

Table 3: Sequential placement on a line of length 10 for various Λ, using the corresponding optimal policies for $\xi = 0.1$.

Λ	Normalised Optimal distances of the nodes from the source	No. of relays
0.01	10	0
0.1	5.3060, 10.0000	1
5	0, 0.0050, 0.0510, 0.1220, 0.1930, 0.2640, $\cdots$, 9.9910, 10.0000	143
8	0, 0.003, 0.019, 0.06, 0.101, $\cdots$, 9.982, 10	246
20	0, 0.001, 0.003, 0.016, 0.031, 0.046, $\cdots$, 9.991, 10	669

Table 4: Evolution of state in the process of sequential placement on a line of length 10 for various values of Λ, using the corresponding optimal policies for $\xi = 0.1$.

Λ	Evolution of state in the process of sequential placement
0.01	1
0.1	1, 0.63
5	1, 0.5, 0.34, 0.3, 0.3, $\cdots$
8	1, 0.5, 0.34, 0.28, 0.28, $\cdots$
20	1, 0.5, 0.34, 0.27, 0.26, 0.26, $\cdots$

from the previous nodes, and hence does not need to be placed close to the k-th relay. A value of s_k close to 1 indicates that there is a large gap between relay k and relay $k-1$, the power received at the next relay from the previous relays is small and hence the next relay must be placed closer to the previous one.

On the other hand, a^* is increasing with ξ (see Figure 6). Recall that ξ is the price of placing a relay. This figure confirms the intuition that if the relay price is high, then the relays should be placed less frequently.

Figure 7 shows that a^* is decreasing with Λ, for fixed values of ξ and s. This happens because increased attenuation will require frequent placement of the relays.

5.2 Relay Placement Patterns

The policy that we use corresponds to a line having exponentially distributed length with mean 1, but it is applied to the scenario where the actual realization of the (normalised) length (see Section 4.6) of the line is 10.

Tables 1, 3, and 5 illustrate some examples of as-you-go placement of relay nodes along a line of normalised length 10, using various values of Λ and ξ. Tables 2, 4, and 6 illustrate the corresponding evolution of state

Table 5: Sequential placement on a line of length 10 for of ξ, using the corresponding optimal policies for $\Lambda = 20$.

ξ	Normalised Optimal distances of the nodes from the source	No. of relays
0.2	0 , 0.008, 0.03, 0.052, $\cdots$, 9.996, 10	456
1	0.022, 0.069, 0.116, $\cdots$, 9.986, 10	213
2	0.042, 0.103, 0.163, 0.223, $\cdots$, 9.943, 10	166
10	0.099, 0.205, 0.311, $\cdots$, 9.957, 10	94

Table 6: Evolution of state in the process of sequential placement on a line of length 10 for various values of ξ, using the corresponding optimal policies for $\Lambda = 20$.

ξ	Evolution of state in the process of sequential placement
0.2	1, 0.5, 0.37, 0.37, $\cdots$
1	1, 0.61, 0.61, $\cdots$
2	1, 0.7, 0.71, 0.71, $\cdots$
10	1, 0.88, 0.88, $\cdots$

as the relays are placed in the examples in Tables 1, 3 and 5. If the line actually ends at some point before (normalised) distance 10, the process would end there with the corresponding placement of relays (as can be obtained from Tables 1, 3, and 5) before the sink being placed at the end-point. Thus, for example, reading from Table 3 for $\xi = 0.1$ and $\Lambda = 5$, if the actual normalised length of the line is 0.99, then one relay will be placed at 0 (the source itself), followed by 15 relays at normalised distances $0.005, 0.051, 0.122, 0.193, 0.264, \cdots, 0.974$ from the source, and finally the sink is placed at a normalised distance 0.99, the end of the line.

We observe that as Λ increases, more relays need to be placed since the optimal control decreases with Λ for each s (see Figure 7). On the other hand, the number of relays decreases with increasing ξ (the relay cost); this is in confirmation of the observations from Figure 6.

Note that, initially one or more relays are placed at or near the source if $a^*(s = 1)$ is 0 or small. But, after some relays have been placed, the relays are placed equally spaced apart. We see that this happens because, after a few relays have been placed, the state, s, does not change, hence, resulting in the relays being subsequently placed equally spaced apart. This phenomenon is evident in Table 2, Table 4, Table 6, and Figure 8. The state s will remain unchanged after a relay placement if $s = \lceil \frac{se^{\Lambda a^*(s)}}{0.01(1+se^{\Lambda a^*(s)})} \rceil \times 0.01$, since we have discretized the state space. After some relays are placed, the state becomes equal to a fixed point s' of the function $\lceil \frac{se^{\Lambda a^*(s)}}{0.01(1+se^{\Lambda a^*(s)})} \rceil \times 0.01$. Note that the deployment starts from $s_0 := 1$, but for any value of s_0 (even with s_0 smaller than s'), we numerically observe the same phenomenon. Hence, s' is an absorbing state.

5.3 Numerical Examples for Practical Deployment

In order to provide a more concrete illustration we adopt a path loss parameter from [24]. Figure 4 of [24] shows that the attenuation in the received signal power in a dense urban environment is roughly 50 dB when we move from 50 m distance to 300 m distance away from the transmitter. This yields a value of ρ to be 0.04 per
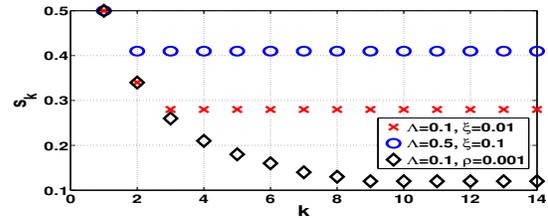
Figure 8: Evolution of the state s_k with k: initial state $s_0 = 1$.

Table 7: Comparison of the performance (in terms of H; see text) of optimal sequential placement over a line of random length, with the optimal placement if the length was known. Results from 10000 samples of exponentially distributed line lengths.

ξ	Λ	Average percentage difference	Mean number of relays used	Number of cases where no relay was used	Maximum percentage difference
0.001	0.01	0.0068	2.0002	0	0.7698
0.001	0.1	0.3996	9.4849	0	6.8947
0.01	0.01	0	0	10000	0
0.01	0.1	0.3517	2.2723	0	4.6618
0.01	0.5	1.5661	7.7572	0	4.7789
0.1	0.01	0	0	10000	0
0.1	0.1	0.1259	0.0056	9944	25.9098
0.1	0.5	2.9869	1.8252	0	12.5907
0.1	2	4.7023	7.1530	0	9.0211
0.1	20	3.5472	27.9217	0	6.6223
0.1	8	4.0097	21.0671	0	7.8264
1	8	8.0286	7.8886	495	27.7362
1	20	5.2158	11.2342	402	26.0026
5	20	10.3341	7.1950	597	61.7460

meter for the exponential path-loss (see the discussion in Section 6.1 on the motivation for choosing the exponential path-loss model in the light of the results from [24]). Then, $\frac{1}{\beta} = 200$ m corresponds to $\Lambda = 8$, and $\frac{1}{\beta} = 500$ m corresponds to $\Lambda = 20$. For $\Lambda = 20$, normalised relay locations and state evolution $\{s_k\}_{k \geq 1}$ are available in Tables 3-6, and, for $\Lambda = 8$, normalised relay locations and state evolution $\{s_k\}_{k \geq 1}$ are available in Tables 3-4. Note that, under $\rho = 0.04$ per meter and $\Lambda = 20$, one unit normalised distance in the tables correspond to 500 m distance in the dense urban environment (due to the normalization as in Section 4.6).

For the sake of illustration, let us consider the sample deployment for $\xi = 10$, $\Lambda = 20$ (Table 5). In this case, the first relay will be placed at a distance $0.099 \times 500 = 49.5$ m from the source, the second relay will be placed at a distance $0.205 \times 500 = 102.5$ m from the source, etc. Also, if we choose ξ such that few relays will be placed on a typical line whose length is several hundreds of meters, then the relays will be placed almost uniformly on the line. But, for small ξ, more relays will be placed and some of them will be clustered near the source (see the deployment for $\Lambda = 8$ and $\xi = 0.1$ in Table 3).

5.4 Comparison with Optimal Offline Deployment

Since there is no prior work in the literature with which we can make a fair comparison of our as-you-go deployment policy for the full-duplex wireless multi-relay network, we compare the performance of our policy with optimal offline deployment. Thus, the numerical

experiments reported in Table 7 are a result of asking the following question: how does the cost of as-you-go deployment over a line of exponentially distributed length compare with the cost of placing the same number of relays optimally over the line, once the length of the line has been revealed?

For several combinations of Λ and ξ, we generated 10000 random numbers independently from an exponential distribution with parameter $\beta = 1$. Each of these numbers was considered as a possible realization of the length of the line. Then we computed the placement of the relay nodes for each realization by optimal sequential placement policy, which gave us $H = \frac{1}{g_{0,1}} + \sum_{k=2}^{N+1} \frac{(g_{0,k-1}-g_{0,k})}{g_{0,k}g_{0,k-1}\sum_{l=0}^{k-1}\frac{1}{g_{0,l}}}$, a quantity that we use to evaluate the quality of the relay placement. The significance of H can be recalled from (6) where we found that the rate $C(1 + \frac{P_T/\sigma^2}{H})$ can be achieved if total power P_T is available to distribute among the source and the relays; i.e., H can be interpreted as the net effective attenuation after power has been allocated optimally over the nodes. Also, for each realization, we computed H for optimal relay placement, assuming that the length of the line is known before deployment and that the number of relays available is the *same* as the number of relays used by the corresponding sequential placement policy. For a given combination of Λ and ξ, for the k-th realization of the length of the line, let us denote the two H values by $H_{\text{sequential}}^{(k)}$ and $H_{\text{optimal}}^{(k)}$. Then the percentage *difference* for the k-th realization is:

$$e_k := \frac{|H_{\text{optimal}}^{(k)} - H_{\text{sequential}}^{(k)}|}{H_{\text{optimal}}^{(k)}} \times 100 \qquad (23)$$

The average percentage difference in Table 7 is the quantity $\frac{\sum_{k=1}^{10000} e_k}{10000}$. The maximum percentage difference is the quantity $\max_{k\in\{1,2,\cdots,10000\}} e_k$.

Discussion of Table 7:

(i) For small enough ξ, some relays will be placed at the source itself. For example, for $\Lambda = 0.01$ and $\xi = 0.001$, we will place two relays at the source (Table 1). After placing the first relay, the next state will become $s = 0.5$, and $a^*(s = 0.5) = 0$. The state after placing the second relay becomes $s = 0.34$, for which $a^*(s = 0.34) = 8.41$ (see the placement in Table 1). Now, the line having an exponentially distributed length with mean 1 will end before $a^*(s = 0.34) = 8.41$ distance with high probability, and the probability of placing the third relay will be very small. As a result, the mean number of relays will be 2.0002. In case only 2 relays are placed by the sequential deployment policy and we seek to place 2 relays optimally for the same length of the line (with the length known), the optimal locations for both relays are close to the source location if the length of the line is small (i.e., if the attenuation λ is small, recall the definition of λ from Section 3.2). If the line is long (which has a very small proba-

bility), the optimal placement will be significantly different from the sequential placement. Altogether, the difference (from (23)) will be small.

(ii) For $\Lambda = 0.01$ and $\xi = 0.1$, $a^*(1)$ is so large that with high probability the line will end in a distance less than $a^*(1)$ and no relay will be placed.

(iii) From (6) we know that for a given placement of relays on a line of given length L, the optimal power allocation yields an achievable rate $\log_2(1 + \frac{P_T/\sigma^2}{H})$. At the end of as-you-go deployment the power is allocated optimally among the nodes deployed, and a rate $\log_2(1 + \frac{P_T/\sigma^2}{H_{sequential}})$ can be achieved. If the same number of relays are optimally placed over the same line, with the same total power, then the inner bound is given by $\log_2(1 + \frac{P_T/\sigma^2}{H_{optimal}})$. We seek to compare these two rates numerically.

The maximum fractional difference in Table 7 is less than $\frac{2}{3}$, and substantially smaller than $\frac{2}{3}$ in most cases. Since, in (23), $H_{sequential}^{(k)}$ is always greater than $H_{optimal}^{(k)}$, we have $H_{sequential}^{(k)} \leq \frac{5}{3}H_{optimal}^{(k)}$ for all $k \geq 1$ (i.e., for all realizations of L in the simulation). Now, by the monotonicity of $\log_2(\cdot)$:

$$\frac{1}{2}\log_2\left(1 + \frac{P_T/\sigma^2}{H_{optimal}^{(k)}}\right) - \frac{1}{2}\log_2\left(1 + \frac{P_T/\sigma^2}{H_{sequential}^{(k)}}\right)$$
$$\leq \quad \frac{1}{2}\log_2\left(1 + \frac{P_T/\sigma^2}{H_{optimal}^{(k)}}\right) - \frac{1}{2}\log_2\left(1 + \frac{P_T/\sigma^2}{\frac{5}{3}H_{optimal}^{(k)}}\right)$$

Since $\log_2(\cdot)$ is a concave function, for any $x > y > 0$, we have $\log_2(1 + x) - \log_2(1 + y) \leq \log_2(x) - \log_2(y)$. Using this inequality, we can upper bound the difference in achievable rate from the previous equation by:

$$\frac{1}{2}\log_2\left(\frac{P_T/\sigma^2}{H_{optimal}^{(k)}}\right) - \frac{1}{2}\log_2\left(\frac{P_T/\sigma^2}{\frac{5}{3}H_{optimal}^{(k)}}\right) = 0.3685$$

This calculation implies that, for the large number of cases reported in Table 7, by using the approximation in (10) and by using the corresponding optimal policy for as-you-go deployment, we lose at most 0.3685 bits per channel use, compared to the case when the realization of the exponentially distributed source to sink distance is known apriori and when we use the same number of relays as used in the as-you-go deployment case. Note that, the statement of this claim holds *with high probability* since the maximum difference is taken over 10000 sample deployments. Hence, it is reasonable to solve (10) instead of (11) which is intractable.

6 DISCUSSION

6.1 Exponential Path-Loss Model

Exponential path-loss model has been used before in the context of relay placement (see [16], [22]) and in the context of cellular networks (see [26, Section 2.3]). Analytical

and experimental support for the exponential path-loss model have been provided by Franceschetti et al. ([24]). Franceschetti et al. used a random scattering model (applicable to an urban environment, or a forest environment) to show that the path-loss in such an environment is the product of an exponential function and a power function of the distance (see [24, Equation (14)]). Figure 4 of their paper, which is obtained from measurements made in an urban environment, shows that path-loss (in dB) varies linearly with distance beyond a distance of $40 - 50$ meters, which implies exponential path-loss for longer distance. These distances are practical for urban scenarios where the network is deployed over several hundreds of meters or several kilometers.

Exponential path-loss was also proposed by Marano and Franceschetti for urban environment, and validated by theory and experiment (see [30, Figure 10]).[7]

6.2 Incorporating Shadowing and Fading

Shadowing (which is typically viewed as being static once a link is deployed) and time varying fading, can be incorporated in our setting by providing a fade-margin in the power at each transmitter. Thus, when expressed in dBm, the actual transmit power for any transmitter-receiver pair is the fade-margin plus the power used in the information theoretic capacity formulas; this fade-margin does not depend on the distance between the transmitter-receiver pair. Note that, this approach, though very conservative in nature, can remove the complexity in analysis arising out of fading in the network. Also note that, if the actual power gain between two nodes r distance apart is $c_0 e^{-\rho r}$ with $c_0 > 0$, then c_0 can be absorbed in the fade margin.

6.3 Full-Duplex Decode-and-Forward Relaying

Full-duplex radios might become a reality soon; see [11], [12], [13], [14] for recent efforts to realize them. Decode-and-forward relaying requires symbol-level synchronous operation across all nodes in the network. The requirement of globally coherent transmission and reception seems to be restrictive at the moment, but this problem will be solved with the advent of better clocks (with less drift) and efficient clock synchronization algorithms. Any research on impromptu deployment assuming imperfect synchronization, or half-duplex communication, or no interference cancellation, can use this paper as a benchmark for performance analysis.

6.4 Insights on Power-Law Path-Loss

In [1], we studied the problem of single-relay placement under a per-node power constraint at the source and the relay, for both exponential and power-law path-loss models. The variation of optimal relay location,

as the amount of attenuation in the network varies, follow slightly different (but mostly similar) trends (see Figures 2 and 3 of [1]) because of the fact that power-law model allows unbounded power gain (unlike the exponential model) when the distance r tends to 0 ($\lim_{r \to 0} r^{-\eta} = \infty$). The findings are even more similar when we bound the power gain from above by some constant value in case of the power-law model (power gain is $\min\{r^{-\eta}, b^{-\eta}\}$ for some $b > 0$); see the similarity between Figures 2 and 4 in [1]. The results on the fixed node power case provide the insight that when the power gain is $r^{-\eta}$ or $\min\{r^{-\eta}, b^{-\eta}\}$; under the sum power constraint, the variation of the relay locations as a function of attenuation will follow a pattern similar to that in case of exponential path-loss.

7 CONCLUSION

Motivated by the problem of as-you-go deployment of wireless relay networks, we first studied the problem of placing relay nodes along a line, in order to connect a sink at the end of the line to a source at the start of the line, so as to maximize the end-to-end achievable data rate. For the multi-relay channel with exponential path-loss and sum power constraint, we derived an expression for the achievable rate in terms of the power gains among all possible node pairs, and formulated an optimization problem in order to maximize the end-to-end data rate. Numerical work for the fixed source-sink distance suggests that at low attenuation the relays are mostly clustered close to the source in order to be able to cooperate among themselves, whereas at high attenuation they are uniformly placed and work as repeaters. Next, the deploy-as-you-go sequential placement problem was addressed; a sequential relay placement problem along a line having unknown random length was formulated as an MDP, the value function was characterized analytically, and the policy structure was investigated numerically. We found numerically that at the initial stage of the deployment process the inter-relay distances are smaller, and, as deployment progresses, the inter-relay distances increase gradually, and finally the relays start being placed at regular intervals.

Our results are based on information theoretic achievable rate results. In order to utilize currently commercially available wireless devices, we have also been exploring non-information theoretic, packet forwarding models for optimal relay placement, with the aim of obtaining placement algorithms that can be easily reduced to practice (see [7] for reference). The study of as-you-go deployment under the information theoretic model and under the packet forwarding model provides two complementary approaches for two different conditions in the physical layer and the MAC layer, and provides a more comprehensive development of the problem.

REFERENCES

[1] A. Chattopadhyay, A. Sinha, M. Coupechoux, and A. Kumar, "Optimal capacity relay node placement in a multi-hop network

---

7. Marano and Franceschetti ([30]) modeled a city as a random lattice, and the distance from the transmitter to the receiver is measured along the edges of the lattice instead of the Euclidean distance. Hence, this result renders the analysis in our paper valid even for deployment along the streets of a city with turns; deployment algorithm in that case will only consider the distances along the streets and not on the actual Euclidean distances.

on a line," in *Proc. of the 8th International workshop on Resource Allocation and Cooperation in Wireless Networks (RAWNET), 2012 (In conjunction with WiOpt 2012)*. IEEE, 2012, pp. 452—459.

[2] ——, "Optimal capacity relay node placement in a multi-hop network on a line," *http://arxiv.org/abs/1204.4323v3*.

[3] M. Souryal, J. Geissbuehler, L. Miller, and N. Moayeri, "Real-time deployment of multihop relays for range extension," in *Proc. of the International Conference on Mobile Systems, Applications, and Services (MobiSys)*. ACM, 2007, pp. 85–98.

[4] M. Howard, M. Matarić, and G. Sukhatme, "An incremental self-deployment algorithm for mobile sensor networks," *Kluwer Autonomous Robots*, vol. 13, no. 2, pp. 113–126, 2002.

[5] J. Bao and W. Lee, "Rapid deployment of wireless ad hoc backbone networks for public safety incident management," in *Proc. of the Global Telecommunications Conference (GLOBECOM)*. IEEE, 2007, pp. 1217—1221.

[6] A. Sinha, A. Chattopadhyay, K. Naveen, M. Coupechoux, and A. Kumar, "Optimal sequential wireless relay placement on a random lattice path," *Ad Hoc Networks Journal (Elsevier).*, vol. 21, pp. 1–17, 2014.

[7] A. Chattopadhyay, M. Coupechoux, and A. Kumar, "Measurement based impromptu deployment of a multi-hop wireless relay network," in *Proc. of the 11th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*. IEEE, 2013.

[8] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, 1979.

[9] A. Reznik, S. Kulkarni, and S. Verdú, "Degraded gaussian multirelay channel: capacity and optimal power allocation," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3037–3046, 2004.

[10] L. Xie and P. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 748–767, 2004.

[11] A. Khandani, "Two-way (true full-duplex) wireless," in *Proc. of the 13th Canadian Workshop on Information Theory (CWIT)*. IEEE, 2013, pp. 33—38.

[12] ——, "Methods for spatial multiplexing of wireless two-way channels," in *http://www.google.ca/patents/US7817641*, 2010.

[13] J. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking (MobiCom), 2010*. ACM, 2010, pp. 1–12.

[14] M. Jain, J. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking (MobiCom)*. ACM, 2011, pp. 301–312.

[15] L. Georgiadis, M. Neely, and L. Tassiulas, *Resource allocation and cross-layer control in wireless networks*. Now Pub, 2006.

[16] S. Firouzabadi and N. Martins, "Optimal node placement in wireless networks," in *3rd International Symposium on Communications, Control and Signal Processing (ISCCSP), 2008*. IEEE, 2008, pp. 960–965.

[17] G. Zheng, C. Hua, R. Zheng, and Q. Wang, "A robust relay placement framework for 60GHz mm Wave wireless personal area networks," *http://arxiv.org/abs/1207.6509*.

[18] H. Lu, W. Liao, and Y. Lin, "Relay station placement strategy in IEEE 802.16j WiMAX networks," *IEEE Transactions on Communications*, vol. 59, no. 1, January 2011.

[19] M. Thakur, N. Fawaz, and M. Médard, "Optimal relay location and power allocation for low SNR broadcast relay channels," *Arxiv preprint arXiv:1008.3295*, 2010.

[20] B. Lin, P. Ho, L. Xie, and X. Shen, "Optimal relay station placement in IEEE 802.16j networks," in *Proc. of the international conference on Wireless communications and mobile computing*. ACM, 2007, pp. 25—30.

[21] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: a deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, 2011.

[22] R. Appuswamy, E. Atsan, C. Fragouli, and M. Franceschetti, "On relay placement for deterministic line networks," in *Wireless Network Coding Conference (WiNC), 2010*. IEEE, 2010, pp. 1–9.

[23] P. Mondal, K. Naveen, and A. Kumar, "Optimal Deployment of Impromptu Wireless Sensor Networks," in *Proc. of the IEEE National Conference on Communications (NCC),2012*. IEEE, 2012.

[24] M. Franceschetti, J. Bruck, and L. Schulman, "A random walk model of wave propagation," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 5, pp. 1304–1317, 2004.

[25] P. Gupta and P. Kumar, "Towards an information theory of large networks: An achievable rate region," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1877–1894, August, 2003.

[26] E. Altman, M. Hanawal, R. El-Azouzi, and S. Shamai, "Trade-offs in green cellular networks," *ACM SIGMETRICS Performance Evaluation Review*, vol. 39, no. 3, pp. 67—71, December, 2011.

[27] R. Zheng, "Information dissemination in power-constrained wireless networks," in *Proc. of the International Conference on Computer Communications (INFOCOM)*. IEEE, 2006, pp. 1—10.

[28] F. J. Beutler and K. W. Ross, "Optimal policies for controlled markov chains with a constraint," *Journal of Mathematical Analysis and Applications*, vol. 112, pp. 236–252, 1985.

[29] M. Schäl, "Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal," *Probability theory and related fields*, vol. 32, no. 3, pp. 179–196, 1975.

[30] S. Marano and M. Franceschetti, "Ray propagation in a random lattice: A maximum entropy, anomalous diffusion process," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 6, June, 2005.

Arpan Chattopadhyay obtained his B.E. in Electronics and Telecommunication Engineering from Jadavpur University, Kolkata, India in the year 2008, and M.E. and Ph.D in Telecommunication Engineering from Indian Institute of Science, Bangalore, India in the year 2010 and 2015, respectively. He is currently working in INRIA, Paris as a postdoctoral researcher. His research interests include networks and machine learning.

Abhishek Sinha is currently a graduate student in the Laboratory for Information and Decision Systems (LIDS), at Massachusetts Institute of Technology, Cambridge, MA. Prior to joining MIT, he completed his Master's studies in Telecommunication Engineering at the Indian Institute of Science, Bangalore, in the year 2012. His areas of interests include stochastic processes, information theory and network control.

Marceau Coupechoux is an Associate Professor at Telecom ParisTech since 2005. He obtained his master from Telecom ParisTech in 1999 and from University of Stuttgart, Germany in 2000, and his Ph.D. from Institut Eurecom, Sophia-Antipolis, France, in 2004. From 2000 to 2005, he was with Alcatel-Lucent (Bell Labs former Research & Innovation and then in the Network Design department). In the Computer and Network Science department of Telecom ParisTech, he is working on cellular networks, wireless networks, ad hoc networks, cognitive networks, focusing mainly on layer 2 protocols, scheduling and resource management. From August 2011 to August 2012 he was a visiting scientist at IISc Bangalore.

Anurag Kumar obtained his B.Tech. degree from the Indian Institute of Technology at Kanpur, and the PhD degree from Cornell University, both in Electrical Engineering. He was then with Bell Laboratories, Holmdel, N.J., for over 6 years. Since 1988 he has been on the faculty of the Indian Institute of Science (IISc), Bangalore, in the Department of Electrical Communication Engineering. He is currently also the Director of the Institute. From 1988 to 2003 he was the Coordinator at IISc of the Education and Research Network Project (ERNET), India's first wide-area packet switching network. His area of research is communication networking, specifically, modeling, analysis, control and optimisation problems arising in communication networks and distributed systems. Recently his research has focused primarily on wireless networking. He is a Fellow of the IEEE, of the Indian National Science Academy (INSA), of the Indian Academy of Science (IASc), of the Indian National Academy of Engineering (INAE), and of The World Academy of Sciences (TWAS). He is a recepient of the Indian Institute of Science Alumni Award for Engineering Research for 2008.

Supplementary Material for "Deploy-As-You-Go Wireless Relay Placement: An Optimal Sequential Decision Approach using the Multi-Relay Channel Model"

APPENDIX A
A BRIEF DESCRIPTION OF THE CODING SCHEME OF [10]

Transmissions take place via block codes of T symbols each. The transmission blocks at the source and the N relays are synchronized. The coding and decoding scheme is such that a message generated at the source at the beginning of block $b, b \geq 1$, is decoded by the sink at the end of block $b + N$, i.e., $N + 1$ block durations after the message was generated (with probability tending to 1, as $T \to \infty$). Thus, at the end of B blocks, $B \geq N + 1$, the sink is able to decode $B - N$ messages. It follows, by taking $B \to \infty$, that, if the code rate is R bits per symbol, then an information rate of R bits per symbol can be achieved from the source to the sink.

As mentioned earlier, we index the source by 0, the relays by $k, 1 \leq k \leq N$, and the sink by $N + 1$. There are $(N+1)^2$ independent Gaussian random codebooks, each containing 2^{TR} codes, each code being of length T; these codebooks are available to all nodes. At the beginning of block b, the source generates a new message w_b, and, at this stage, we assume that each node $k, 1 \leq k \leq N+1$, has a reliable estimate of all the messages $w_{b-j}, j \geq k$. In block b, the source uses a new codebook to encode w_b. In addition, relay $k, 1 \leq k \leq N$, and *all* of its previous transmitters (indexed $0 \leq j \leq k - 1$), use *another* codebook to encode w_{b-k} (or their estimate of it). Thus, if the relays $1, 2, \cdots, k$ have a perfect estimate of w_{b-k} at the beginning of block b, they will transmit the same codeword for w_{b-k}. Therefore, in block b, the source and relays $1, 2, \cdots, k$ *coherently transmit* the codeword for w_{b-k}. In this manner, in block b, transmitter $k, 0 \leq k \leq N$, generates $N + 1 - k$ codewords, corresponding to $w_{b-k}, w_{b-k-1}, \cdots, w_{b-N}$, which are transmitted with powers $P_{k,k+1}, P_{k,k+2}, \cdots, P_{k,N+1}$. In block b, node $k, 1 \leq k \leq N+1$, receives a superposition of transmissions from all other nodes. Assuming that node k knows all the powers, and all the channel gains, and recalling that it has a reliable estimate of all the messages $w_{b-j}, j \geq k$, it can subtract the interference from transmitters $k+1, k+2, \cdots, N$. At the end of block b, after subtracting the signals it knows, node k is left with the k received signals from nodes $0, 1, \cdots, (k-1)$ (received in blocks $b, b-1, \cdots, b-k+1$), which all carry an encoding of the message w_{b-k+1}. These k signals are then jointly used to decode w_{b-k+1}, using joint typicality decoding. The codebooks are cycled through in a manner so that in any block all nodes encoding a message (or their estimate of it) use the same codebook, but different (thus, independent) codebooks are used for different messages. Under this encoding and decoding scheme, any rate strictly less than R displayed in (2) is achievable.

APPENDIX B
PROOF OF THEOREM 1

We want to maximize R given in (2) subject to the total power constraint, assuming fixed relay locations. Let us consider $C(\frac{1}{\sigma^2} \sum_{j=1}^{k} (\sum_{i=0}^{j-1} h_{i,k} \sqrt{P_{i,j}})^2)$, i.e., the k-th term in the argument of $\min\{\cdots\}$ in (2). By the monotonicity of $C(\cdot)$, it is sufficient to consider $\sum_{j=1}^{k} (\sum_{i=0}^{j-1} h_{i,k} \sqrt{P_{i,j}})^2$. Now since the channel gains are multiplicative, we have:

$$\sum_{j=1}^{k} (\sum_{i=0}^{j-1} h_{i,k} \sqrt{P_{i,j}})^2 = g_{0,k} \sum_{j=1}^{k} \left(\sum_{i=0}^{j-1} \frac{\sqrt{P_{i,j}}}{h_{0,i}} \right)^2$$

Thus our optimization problem becomes:

$$\max_{k \in \{1,2,\cdots,N+1\}} \min \; g_{0,k} \sum_{j=1}^{k} \left(\sum_{i=0}^{j-1} \frac{\sqrt{P_{i,j}}}{h_{0,i}} \right)^2$$

$$s.t \quad \sum_{j=1}^{N+1} \gamma_j \leq P_T \; and \; \sum_{i=0}^{j-1} P_{i,j} = \gamma_j \; \forall \, 1 \leq j \leq (N+1) \quad (24)$$

Let us fix $\gamma_1, \gamma_2, \cdots, \gamma_{N+1}$ such that their sum is equal to P_T. We observe that $P_{i,N+1}$ for $i \in \{0,1,\cdots,N\}$ appear in the objective function only once: for $k = N+1$ through the term $(\sum_{i=0}^{N} \frac{\sqrt{P_{i,N+1}}}{h_{0,i}})^2$. Since we have fixed γ_{N+1}, we need to maximize this term over $P_{i,N+1}, i \in \{0,1,\cdots,N\}$. So we have the following optimization problem:

$$\max \sum_{i=0}^{N} \frac{\sqrt{P_{i,N+1}}}{h_{0,i}} \quad s.t \quad \sum_{i=0}^{N} P_{i,N+1} = \gamma_{N+1} \quad (25)$$

By Cauchy-Schwartz inequality, the objective function in this optimization problem is upper bounded by (using the fact that $g_{0,i} = h_{0,i}^2 \; \forall i \in \{0,1,\cdots,N\}$):

$$\sqrt{(\sum_{i=0}^{N} P_{i,N+1})(\sum_{i=0}^{N} \frac{1}{g_{0,i}})} = \sqrt{\gamma_{N+1} \sum_{i=0}^{N} \frac{1}{g_{0,i}}}$$

The upper bound is achieved if there exists some $c > 0$ such that $\frac{\sqrt{P_{i,N+1}}}{\frac{1}{h_{0,i}}} = c \; \forall i \in \{0,1,\cdots,N\}$. So we have:

$$P_{i,N+1} = \frac{c^2}{g_{0,i}} \; \forall i \in \{0,1,\cdots,N\}$$

Since $\sum_{i=0}^{N} P_{i,N+1} = \gamma_{N+1}$, we obtain $c^2 = \frac{\gamma_{N+1}}{\sum_{l=0}^{N} \frac{1}{g_{0,l}}}$.

Thus, $P_{i,N+1} = \frac{\frac{1}{g_{0,i}}}{\sum_{l=0}^{N} \frac{1}{g_{0,l}}} \gamma_{N+1}$.

Here we have used the fact that $h_{0,0} = 1$. Now $\{P_{i,N} : i = 0, 1, \cdots, (N-1)\}$ appear only through the sum $\sum_{i=0}^{N-1} \frac{\sqrt{P_{i,N}}}{h_{0,i}}$, and it appears twice: for $k = N$ and $k = N+1$. We need to maximize this sum subject to the constraint $\sum_{i=0}^{N-1} P_{i,N} = \gamma_N$. This optimization can be solved in a similar way as before. Thus by repeatedly using this argument and solving optimization problems similar in nature to (25), we obtain:

$$P_{i,j} = \frac{\frac{1}{g_{0,i}}}{\sum_{l=0}^{j-1} \frac{1}{g_{0,l}}} \gamma_j \ \forall 0 \le i < j \le (N+1) \tag{26}$$

Substituting for $P_{i,j}, 0 \le i < j \le (N+1)$ in (24), we obtain the following optimization problem:

$$\max_{k \in \{1,2,\cdots,N+1\}} \min g_{0,k} \sum_{j=1}^{k} \left(\gamma_j \sum_{i=0}^{j-1} \frac{1}{g_{0,i}} \right)$$

$$s.t. \quad \sum_{j=1}^{N+1} \gamma_j \le P_T \tag{27}$$

Let us define $b_k := g_{0,k}$ and $a_j := \sum_{i=0}^{j-1} \frac{1}{g_{0,i}}$. Observe that b_k is decreasing and a_k is increasing with k. Let us define:

$$\tilde{s}_k(\gamma_1, \gamma_2, \cdots, \gamma_{N+1}) := b_k \sum_{j=1}^{k} a_j \gamma_j \tag{28}$$

With this notation, our optimization problem becomes:

$$\max \min_{1 \le k \le N+1} \tilde{s}_k(\gamma_1, \gamma_2, \cdots, \gamma_{N+1}) \ \ s.t. \ \sum_{j=1}^{N+1} \gamma_j \le P_T \tag{29}$$

Claim 1: Under optimal allocation of $\gamma_1, \gamma_2, \cdots, \gamma_{N+1}$ for the optimization problem (29), $\tilde{s}_1 = \tilde{s}_2 = \cdots = \tilde{s}_{N+1}$. □

Proof: (29) can be rewritten as:

$$\max \zeta$$

$$s.t \quad \zeta \le b_k \sum_{j=1}^{k} a_j \gamma_j \ \forall k \in \{1, 2, \cdots, N+1\},$$

$$\sum_{j=1}^{N+1} \gamma_j \le P_T, \ \gamma_j \ge 0 \ \forall 1 \le j \le N+1 \tag{30}$$

The dual of this linear program is given by:

$$\min P_T \theta$$

$$s.t \quad \sum_{k=1}^{N+1} \mu_k = 1, \ \theta \ge 0,$$

$$a_l \sum_{k=l}^{N+1} b_k \mu_k + \nu_l = \theta \ \forall l \in \{1, 2, \cdots, N+1\},$$

$$\mu_l \ge 0, \nu_l \ge 0 \ \forall l \in \{1, 2, \cdots, N+1\} \tag{31}$$

Now, let us consider a primal feasible solution $(\{\gamma_j^*\}_{1 \le j \le N+1}, \zeta^*)$ which satisfies:

$$b_k \sum_{j=1}^{k} a_j \gamma_j^* = \zeta^* \ \forall k \in \{1, 2, \cdots, N+1\},$$

$$\sum_{j=1}^{N+1} \gamma_j^* = P_T \tag{32}$$

Thus we have, $b_1 a_1 \gamma_1^* = \zeta^*$, i.e., $\gamma_1^* = \frac{\zeta^*}{b_1 a_1}$. Again, $b_2(a_1 \gamma_1^* + a_2 \gamma_2^*) = \zeta^*$, which implies $\frac{b_2}{b_1}\zeta^* + b_2 a_2 \gamma_2^* = \zeta^*$.

Thus we obtain $\gamma_2^* = \frac{\zeta^*}{a_2}(\frac{1}{b_2} - \frac{1}{b_1})$. In general, we can write:

$$\gamma_k^* = \frac{\zeta^*}{a_k} \left(\frac{1}{b_k} - \frac{1}{b_{k-1}} \right) \forall k \in \{1, 2, \cdots, N+1\}$$

with $\frac{1}{b_0} := 0$. Now, since $\sum_{k=1}^{N+1} \gamma_k^* = P_T$, we obtain:

$$\zeta^* = \frac{P_T}{\sum_{k=1}^{N+1} \frac{1}{a_k}(\frac{1}{b_k} - \frac{1}{b_{k-1}})}$$

$$\gamma_j^* = \frac{\frac{1}{a_j}\left(\frac{1}{b_j} - \frac{1}{b_{j-1}}\right)}{\sum_{k=1}^{N+1} \frac{1}{a_k}(\frac{1}{b_k} - \frac{1}{b_{k-1}})} P_T, \ j \in \{1, 2, \cdots, N+1\} \tag{33}$$

It should be noted that since b_k is nonincreasing in k, the primal variables above are nonnegative and satisfies feasibility conditions. Again, let us consider a dual feasible solution $(\{\mu_j^*, \nu_j^*\}_{1 \le j \le N+1}, \theta^*)$ which satisfies:

$$\sum_{k=1}^{N+1} \mu_k^* = 1, \ \nu_l^* = 0 \ \forall l \in \{1, 2, \cdots, N+1\}$$

$$a_l \sum_{k=l}^{N+1} b_k \mu_k^* + \nu_l^* = \theta^* \ \forall l \in \{1, 2, \cdots, N+1\} \tag{34}$$

Solving these equations, we obtain:

$$\theta^* = \frac{1}{\sum_{k=1}^{N+1} \frac{1}{b_k} \left(\frac{1}{a_k} - \frac{1}{a_{k+1}} \right)}$$

$$\mu_j^* = \frac{\frac{1}{b_j}(\frac{1}{a_j} - \frac{1}{a_{j+1}})}{\sum_{k=1}^{N+1} \frac{1}{b_k} \left(\frac{1}{a_k} - \frac{1}{a_{k+1}} \right)}, \ j \in \{1, 2, \cdots, N+1\} \tag{35}$$

where $\frac{1}{a_{N+2}} := 0$. Since a_k is increasing in k, all dual variables are feasible. It is easy to check that $\zeta^* = P_T \theta^*$, which means that there is no duality gap. Since the primal is a linear program, the solution $(\gamma_1^*, \gamma_2^*, \cdots, \gamma_{N+1}^*, \zeta^*)$ is primal optimal. Thus we have established the claim, since the primal optimal solution satisfies it. □

So let us obtain $\gamma_1, \gamma_2, \cdots, \gamma_{N+1}$ for which $\tilde{s}_1 = \tilde{s}_2 = \cdots = \tilde{s}_{N+1}$. Putting $\tilde{s}_k = \tilde{s}_{k-1}$, we obtain $b_k \sum_{j=1}^{k} a_j \gamma_j = b_{k-1} \sum_{j=1}^{k-1} a_j \gamma_j$. Thus, we obtain, $\gamma_k = \frac{(b_{k-1}-b_k)}{b_k} \frac{1}{a_k} \sum_{j=1}^{k-1} a_j \gamma_j$

Let $d_k := \frac{(b_{k-1}-b_k)}{b_k} \frac{1}{a_k}$. Hence, $\gamma_k = d_k \sum_{j=1}^{k-1} a_j \gamma_j$. From this recursive equation, we have $\gamma_2 = d_2 a_1 \gamma_1$, $\gamma_3 = d_3(a_1 \gamma_1 + a_2 \gamma_2) = d_3 a_1(1 + a_2 d_2)\gamma_1$, and, in general for $k \ge 3$,

$$\gamma_k = d_k a_1 \Pi_{j=2}^{k-1}(1 + a_j d_j)\gamma_1 \tag{36}$$

Using the fact that $\gamma_1 + \gamma_2 + \cdots + \gamma_{N+1} = P_T$, we obtain:

$$\gamma_1 = \frac{P_T}{1 + d_2 a_1 + \sum_{k=3}^{N+1} d_k a_1 \Pi_{j=2}^{k-1}(1 + a_j d_j)} \tag{37}$$

Thus if $\tilde{s}_1 = \tilde{s}_2 = \cdots = \tilde{s}_{N+1}$, there is a unique allocation $\gamma_1, \gamma_2, \cdots, \gamma_{N+1}$. So this must be the one maximizing R. Hence, optimum γ_1 is obtained by (37). Then, substituting the values of $\{a_k : k = 0, 1, \cdots, N\}$ and $d_k : k = 1, 2, \cdots, N+1$ in (36) and (37), we obtain the

values of $\gamma_1, \gamma_2, \cdots, \gamma_{N+1}$ as shown in Theorem 1.

Now under these optimal values of $\gamma_1, \gamma_2, \cdots, \gamma_{N+1}$, all terms in the argument of $\min\{\cdots\}$ in (2) are equal. So we can consider the first term alone. Thus we obtain the expression for R optimized over power allocation among all the nodes for fixed relay locations as : $R_{P_T}^{opt}(y_1, y_2, \cdots, y_N) = C\left(\frac{g_{0,1}P_{0,1}}{\sigma^2}\right) = C\left(\frac{g_{0,1}\gamma_1}{\sigma^2}\right)$. Substituting the expression for γ_1 from (5), we obtain the achievable rate formula (6). $\qquad\square$

APPENDIX C
PROOF OF THEOREM 2

Here we want to place the relay node at a distance r_1 from the source to minimize $\left\{\frac{1}{g_{0,1}} + \frac{g_{0,1}-g_{0,2}}{g_{0,2}(1+g_{0,1})}\right\}$ (see Equation (6)). Hence, our optimization problem becomes :

$$\min_{r_1 \in [0,L]} \left\{ e^{\rho r_1} + \frac{e^{-\rho r_1} - e^{-\rho L}}{e^{-\rho L}(1 + e^{-\rho r_1})} \right\}$$

Writing $z_1 = e^{\rho r_1}$, the problem becomes :

$$\min_{z_1 \in [1, e^{\rho L}]} \left\{ z_1 - 1 + \frac{e^{\rho L} + 1}{z_1 + 1} \right\}$$

This is a convex optimization problem. Equating the derivative of the objective function to zero, we obtain $1 - \frac{e^{\rho L}+1}{(z_1+1)^2} = 0$. Thus the derivative becomes zero at $z_1' = \sqrt{1 + e^{\rho L}} - 1 > 0$. Hence, the objective function is decreasing in z_1 for $z_1 \leq z_1'$ and increasing in $z_1 \geq z_1'$. So the minimizer is $z_1^* = \max\{z_1', 1\}$. So the optimum distance of the relay node from the source is $y_1^* = r_1^* = \max\{0, r_1'\}$, where $r_1' = \frac{1}{\rho}\log(\sqrt{1 + e^{\rho L}} - 1)$. Hence, $\frac{y_1^*}{L} = \max\{\frac{1}{\lambda}\log\left(\sqrt{e^\lambda + 1} - 1\right), 0\}$. Now $r_1' \geq 0$ if and only if $\lambda \geq \log 3$. Hence, $\frac{y_1^*}{L} = 0$ for $\lambda \leq \log 3$ and $\frac{y_1^*}{L} = \frac{1}{\lambda}\log\left(\sqrt{e^\lambda + 1} - 1\right)$ for $\lambda \geq \log 3$.

For $\lambda \leq \log 3$, the relay is placed at the source. Then $g_{0,1} = 1$ and $g_{0,2} = g_{1,2} = e^{-\lambda}$. Then $P_{0,1} = \gamma_1 = \frac{2P_T}{e^\lambda+1}$ (by Theorem 1) and $R^* = C\left(\frac{2P_T}{(e^\lambda+1)\sigma^2}\right)$. Also $\gamma_2 = \frac{e^\lambda-1}{e^\lambda+1}P_T$. Hence, $P_{0,2} = P_{1,2} = \frac{e^\lambda-1}{e^\lambda+1}\frac{P_T}{2}$.

for $\lambda \geq \log 3$, the relay is placed at r_1'. Substituting the value of r_1' into Equation (4), we obtain $P_{0,1} = \gamma_1 = \frac{P_T}{2}$, $P_{0,2} = \frac{1}{\sqrt{e^\lambda+1}}\frac{P_T}{2}$, $P_{1,2} = \frac{\sqrt{e^\lambda+1}-1}{\sqrt{e^\lambda+1}}\frac{P_T}{2}$. So in this case $R^* = C\left(\frac{g_{0,1}P_{0,1}}{\sigma^2}\right)$. Since $P_{0,1} = \frac{P_T}{2}$, we have $R^* = C\left(\frac{1}{\sqrt{e^\lambda+1}-1}\frac{P_T}{2\sigma^2}\right)$. $\qquad\square$

APPENDIX D
PROOF OF THEOREM 3

For the N-relay problem, let the minimizer in (8) be $z_1^*, z_2^*, \cdots, z_N^*$ and let $y_k^* = \frac{1}{\rho}\log z_k^*$. Clearly, there exists $i \in \{0, 1, \cdots, N\}$ such that $y_{i+1}^* > y_i^*$. Let us insert a new relay at a distance y from the source such that

$y_i^* < y < y_{i+1}^*$. Now we find that we can easily reach (38) (see next page) just by simple comparison. For example,

$$\frac{e^{\rho y} - z_i^*}{1 + z_1^* + \cdots + z_i^*} + \frac{z_{i+1}^* - e^{\rho y}}{1 + z_1^* + \cdots + z_i^* + e^{\rho y}}$$

$$< \frac{e^{\rho y} - z_i^*}{1 + z_1^* + \cdots + z_i^*} + \frac{z_{i+1}^* - e^{\rho y}}{1 + z_1^* + \cdots + z_i^*}$$

$$= \frac{z_{i+1}^* - z_i^*}{1 + z_1^* + \cdots + z_i^*}$$

First i terms in the summations of L.H.S (left hand side) and R.H.S (right hand side) of (38) are identical. Also sum of the remaining terms in L.H.S is smaller than that of the R.H.S since there is an additional $e^{\rho y}$ in the denominator of each fraction for the L.H.S. Hence, we can justify (38). Now R.H.S is precisely the optimum objective function for the N-relay placement problem (see (8)). On the other hand, L.H.S is a particular value of the objective in (8), for $(N+1)$-relay placement problem. This clearly implies that by adding one additional relay we can strictly improve from R^* of the N relay channel. Hence, $R^*(N+1) > R^*(N)$. $\qquad\square$

APPENDIX E
PROOF OF THEOREM 4

Consider the optimization problem as shown in (8). Let us consider λ_1, λ_2, with $\lambda_1 < \lambda_2$, the respective minimizers being $(z_1^*, \cdots, z_N^*)$ and $(z_1', \cdots, z_N')$. Clearly,

$$G(N, \lambda_1) = \frac{e^{\lambda_1}}{z_1^* + \sum_{k=2}^{N+1} \frac{z_k^* - z_{k-1}^*}{\sum_{l=0}^{k-1} z_l^*}} \qquad (39)$$

with $z_{N+1}^* = e^{\lambda_1}$ and $z_0^* = 1$. With $N \geq 1$, note that $z_1^* - \frac{z_N^*}{1+z_1^*+\cdots+z_N^*} \geq 0$, since $z_1^* \geq 1$ and $\frac{z_N^*}{1+z_1^*+\cdots+z_N^*} \leq 1$. Hence, it is easy to see that $\frac{e^\lambda}{z_1^* - \frac{z_N^*}{1+z_1^*+\cdots+z_N^*} + \sum_{k=2}^N \frac{z_k^* - z_{k-1}^*}{\sum_{l=0}^{k-1} z_l^*} + \frac{e^\lambda}{1+z_1^*+\cdots+z_N^*}}$ is increasing in λ where $(z_1^*, \cdots, z_N^*)$ is the optimal solution of (8) with $\lambda = \lambda_1$. Hence,

$$G(N, \lambda_1) = \frac{e^{\lambda_1}}{z_1^* + \sum_{k=2}^N \frac{z_k^* - z_{k-1}^*}{\sum_{l=0}^{k-1} z_l^*} + \frac{e^{\lambda_1} - z_N^*}{\sum_{l=0}^N z_l^*}}$$

$$\leq \frac{e^{\lambda_2}}{z_1^* + \sum_{k=2}^N \frac{z_k^* - z_{k-1}^*}{\sum_{l=0}^{k-1} z_l^*} + \frac{e^{\lambda_2} - z_N^*}{1+z_1^*+\cdots+z_N^*}}$$

$$\leq \frac{e^{\lambda_2}}{z_1' + \sum_{k=2}^N \frac{z_k' - z_{k-1}'}{\sum_{l=0}^{k-1} z_l'} + \frac{e^{\lambda_2} - z_N'}{1+z_1'+\cdots+z_N'}}$$

$$= G(N, \lambda_2) \qquad (40)$$

The second inequality follows from the fact that $(z_1', \cdots, z_N')$ minimizes $z_1 + \sum_{k=2}^{N+1} \frac{z_k - z_{k-1}}{\sum_{l=0}^{k-1} z_l}$ subject to the constraint $1 \leq z_1 \leq z_2 \leq \cdots \leq z_N \leq z_{N+1} = e^{\lambda_2}$.

Hence, $G(N, \lambda)$ is increasing in λ for fixed N. $\qquad\square$

$$z_1^* + \frac{z_2^* - z_1^*}{1 + z_1^*} + \cdots + \frac{e^{\rho y} - z_i^*}{1 + z_1^* + \cdots + z_i^*} + \frac{z_{i+1}^* - e^{\rho y}}{1 + z_1^* + \cdots + z_i^* + e^{\rho y}} + \cdots + \frac{e^{\rho L} - z_N^*}{1 + z_1^* + \cdots + z_i^* + e^{\rho y} + z_{i+1}^* + \cdots + z_N^*}$$

$$< z_1^* + \frac{z_2^* - z_1^*}{1 + z_1^*} + \cdots + \frac{e^{\rho L} - z_N^*}{1 + z_1^* + \cdots + z_i^* + z_{i+1}^* + \cdots + z_N^*} \tag{38}$$

APPENDIX F
PROOF OF THEOREM 5

When N relay nodes are uniformly placed along a line, we will have $y_k = \frac{kL}{N+1}$. Then our formula for achievable rate $R_{P_T}^{opt}(y_1, y_2, \cdots, y_N)$ for sum power constraint becomes: $R_N = C(\frac{P_T}{\sigma^2} \frac{1}{f(N)})$ where $f(N) = a_N + \sum_{k=2}^{N+1} \frac{a_N^k - a_N^{k-1}}{1 + a_N + \cdots + a_N^{k-1}}$ with $a_N = e^{\frac{\rho L}{N+1}} = e^{\frac{\lambda}{N+1}}$.

Since $a_N > 1$ for all $N < \infty$ and $\rho > 0$, we have $f(N) > a_N$ for all $N \geq 1$ and hence, $\liminf_N f(N) \geq \lim_{N \to \infty} a_N = 1$.

Now,

$$\begin{aligned}
f(N) &= a_N + \sum_{k=1}^{N} \frac{a_N^{k+1} - a_N^k}{1 + a_N + \cdots + a_N^k} \\
&= a_N + (a_N - 1)^2 \sum_{k=1}^{N} \frac{a_N^k}{a_N^{k+1} - 1} \\
&\leq a_N + (a_N - 1)^2 \sum_{k=1}^{N} \frac{a_N^k}{a_N^k - 1} \\
&= e^{\frac{\lambda}{N+1}} + (e^{\frac{\lambda}{N+1}} - 1)^2 \sum_{k=1}^{N} \frac{e^{\frac{k\lambda}{N+1}}}{e^{\frac{k\lambda}{N+1}} - 1} \\
&\leq e^{\frac{\lambda}{N+1}} + (e^{\frac{\lambda}{N+1}} - 1)^2 \sum_{k=1}^{N} \frac{e^{\frac{k\lambda}{N+1}}}{\frac{k\lambda}{N+1}} \\
&= e^{\frac{\lambda}{N+1}} + (e^{\frac{\lambda}{N+1}} - 1)^2 \frac{(N+1)}{\lambda} \sum_{k=1}^{N} \frac{e^{\frac{k\lambda}{N+1}}}{k} \tag{41}
\end{aligned}$$

where the first inequality follows from the fact that $a_N > 1$ and the second inequality follows from the fact that $e^{\frac{k\lambda}{N+1}} \geq 1 + \frac{k\lambda}{N+1}$.

Now, by Cauchy-Schwartz inequality,

$$\sum_{k=1}^{N} \frac{e^{\frac{k\lambda}{N+1}}}{k} \leq \sqrt{(\sum_{k=1}^{N} e^{\frac{2k\lambda}{N+1}})(\sum_{k=1}^{N} \frac{1}{k^2})} \tag{42}$$

Since $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$, we can write:

$$\sum_{k=1}^{N} \frac{e^{\frac{k\lambda}{N+1}}}{k} \leq \sqrt{(\sum_{k=1}^{N} e^{\frac{2k\lambda}{N+1}}) \frac{\pi^2}{6}} \tag{43}$$

Hence, by (43) and (41),

$$\begin{aligned}
f(N) &\leq e^{\frac{\lambda}{N+1}} + (e^{\frac{\lambda}{N+1}} - 1)^2 \frac{(N+1)\pi}{\sqrt{6}\lambda} \sqrt{\sum_{k=1}^{N} e^{\frac{2k\lambda}{N+1}}} \\
&= e^{\frac{\lambda}{N+1}} + (e^{\frac{\lambda}{N+1}} - 1)^2 \frac{(N+1)\pi}{\sqrt{6}\lambda} \sqrt{e^{\frac{2\lambda}{N+1}} \frac{(e^{\frac{2N\lambda}{N+1}} - 1)}{(e^{\frac{2\lambda}{N+1}} - 1)}} \tag{44}
\end{aligned}$$

Now, since $e^{\frac{2N\lambda}{N+1}} - 1 \leq e^{\frac{2N\lambda}{N+1}} \leq e^{2\lambda}$, we obtain:

$$\begin{aligned}
f(N) &\leq e^{\frac{\lambda}{N+1}} + (e^{\frac{\lambda}{N+1}} - 1)^2 \frac{(N+1)\pi}{\sqrt{6}\lambda} e^\lambda e^{\frac{\lambda}{N+1}} \sqrt{\frac{1}{(e^{\frac{2\lambda}{N+1}} - 1)}} \\
&= e^{\frac{\lambda}{N+1}} + (e^{\frac{\lambda}{N+1}} - 1)^{\frac{3}{2}} \frac{(N+1)\pi}{\sqrt{6}\lambda} e^\lambda e^{\frac{\lambda}{N+1}} \sqrt{\frac{1}{(e^{\frac{\lambda}{N+1}} + 1)}}
\end{aligned}$$

Hence,

$$\begin{aligned}
&\limsup_N f(N) \\
&\leq 1 + \frac{\pi e^\lambda}{\sqrt{12}} \lim_{N \to \infty} \frac{(N+1)}{\lambda} (e^{\frac{\lambda}{N+1}} - 1)^{\frac{3}{2}}
\end{aligned}$$

Putting $q = \frac{\lambda}{N+1}$,

$$\begin{aligned}
\limsup_N f(N) &\leq 1 + \frac{\pi e^\lambda}{\sqrt{12}} \lim_{q \to 0} \sqrt{q} \lim_{q \to 0} (\frac{e^q - 1}{q})^{\frac{3}{2}} \\
&= 1
\end{aligned}$$

Now, we have proved that $\limsup_N f(N) \leq 1 \leq \liminf_N f(N)$ and hence $\lim_{N \to \infty} f(N) = 1$. Hence, $\lim_{N \to \infty} R_N = C(\frac{P_T}{\sigma^2})$ and the theorem is proved. $\square$

APPENDIX G
PROOF OF THEOREM 7

As we have seen in Section 4, our problem is a negative dynamic programming problem (i.e., the N case of [29], where single-stage rewards are non-positive). It is to be noted that Schäl [29] discusses two other kind of problems as well: the P case (single-stage rewards are positive) and the D case (the reward at stage k is discounted by a factor α^k, where $0 < \alpha < 1$). In this appendix, we first state a general-purpose theorem for the value iteration (Theorem 11), prove it by some results of [29], and then we use this theorem to prove Theorem 7.

G.1 A General Result (Derived from [29])

Consider an infinite horizon total cost MDP whose state space $\mathcal{S}$ is an interval in $\mathbb{R}$ and the action space $\mathcal{A}$ is

$[0, \infty)$. Let the set of possible actions at state s be denoted by $\mathcal{A}(s)$. Let the single-stage cost be $c(s, a, w) \geq 0$ where s, a and w are the state, the action and the disturbance, respectively. Let us denote the optimal expected cost-to-go at state s by $V^*(s)$. Let the state of the system evolve as $s_{k+1} = h(s_k, a_k, w_k)$, where s_k, a_k and w_k are the state, the action and the disturbance at the k-th instant, respectively. Let $s^* \in \mathcal{S}$ be an absorbing state with $c(s^*, a, w) = 0$ for all a, w. Let us consider the value iteration for all $s \in \mathcal{S}$, with $V^{(0)}(\cdot) = 0$:

$$
\begin{aligned}
V^{(k+1)}(s) &= \inf_{a \in [0, \infty)} \mathbb{E}_w \Big(c(s, a, w) + V^{(k)}(h(s, a, w)) \Big), s \neq s^* \\
V^{(k+1)}(s^*) &= 0
\end{aligned}
\tag{45}
$$

We provide some results and concepts from [29], which will be used later to prove Theorem 7.

Theorem 8: [*Theorem 4.2 ([29])*] $V^{(k)}(s) \to V^{(\infty)}(s)$ for all $s \in \mathcal{S}$, i.e., the value iteration (45) converges. □

Let us recall that $\Gamma_k(s)$ is the set of minimizers of (45) at the k-th iteration at state s, if the infimum is achieved at some $a < \infty$. $\Gamma_\infty(s) := \{a \in \mathcal{A} : a$ is an accumulation point of some sequence $\{a_k\}$ where each $a_k \in \Gamma_k(s)\}$. $\Gamma^*(s)$ is the set of minimizers in the Bellman Equation.

Let $\mathcal{C}(\mathcal{A})$ be the set of nonempty compact subsets of $\mathcal{A}$. The Hausdorff metric d on $\mathcal{C}(\mathcal{A})$ is defined as follows:

$$
d(C_1, C_2) = \max\{ \sup_{c \in C_1} \rho(c, C_2), \ \sup_{c \in C_2} \rho(c, C_1) \}
$$

where $\rho(c, C)$ is the minimum distance between the point c and the compact set C.

Proposition 6: [*Proposition 9.1([29])*] $(\mathcal{C}(\mathcal{A}), d)$ is a separable metric space.

A mapping $\phi : \mathcal{S} \to \mathcal{C}(\mathcal{A})$ is called measurable if it is measurable with respect to the Borel σ-algebra of $(\mathcal{C}(\mathcal{A}), d)$.

$\hat{\mathcal{F}}(\mathcal{S} \times \mathcal{A})$ is the set of all measurable functions $v : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ which are bounded below and where every such $v(\cdot)$ is the limit of a non-decreasing sequence of measurable, bounded functions $v_k : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$.

We will next present a condition followed by a theorem. The condition, if satisfied, implies the convergence of value iteration (45) to the optimal value function (according to the theorem).

Condition 1: [*Derived from Condition A in [29]*]

(i) $\mathcal{A}(s) \in \mathcal{C}(\mathcal{A})$ for all $s \in \mathcal{S}$ and $\mathcal{A} : \mathcal{S} \to \mathcal{C}(\mathcal{A})$ is measurable.

(ii) $\mathbb{E}_w(c(s, a, w) + V^{(k)}(h(s, a, w)))$ is in $\hat{\mathcal{F}}(\mathcal{S} \times \mathcal{A})$ for all $k \geq 0$. □

Theorem 9: [*Theorem 13.3, [29]*] If $c(s, a, w) \geq 0$ for all s, a, w and Condition 1 holds:

(i) $V^{(\infty)}(s) = V^*(s)$, $s \in \mathcal{S}$.

(ii) $\Gamma_\infty(s) \subset \Gamma^*(s)$.

(iii) There is a stationary optimal policy f^∞ where $f : \mathcal{S} \to \mathcal{A}$ and $f(s) \in \Gamma_\infty(s)$ for all $s \in \mathcal{S}$. □

The next condition and theorem deal with the situation where the action space is noncompact.

Condition 2: [*Condition B ([29])*] There is a measurable mapping $\underline{\mathcal{A}} : \mathcal{S} \to \mathcal{C}(\mathcal{A})$ such that:

(i) $\underline{\mathcal{A}}(s) \subset \mathcal{A}(s)$ for all $s \in \mathcal{S}$.

(ii) $\inf_{a \in \mathcal{A}(s) - \underline{\mathcal{A}}(s)} \mathbb{E}_w \Big(c(s, a, w) + V^{(k)}(h(s, a, w)) \Big) > \inf_{a \in \mathcal{A}(s)} \mathbb{E}_w \Big(c(s, a, w) + V^{(k)}(h(s, a, w)) \Big)$ for all $k \geq 0$. □

This condition requires that for each state s, there is a compact set $\underline{\mathcal{A}}(s)$ of actions such that no optimizer of the value iteration lies outside the set $\underline{\mathcal{A}}(s)$ at any stage $k \geq 0$.

Theorem 10: [*Theorem 17.1, [29]*] If Condition 2 is satisfied and if the three statements in Theorem 9 are valid for the modified problem having admissible set of actions $\underline{\mathcal{A}}(s)$ for each state $s \in \mathcal{S}$, then those statements are valid for the original problem as well. □

Now we will provide an important theorem which will be used to prove Theorem 7.

Theorem 11: If the value iteration (45) satisfies the following conditions:

(a) For each k, $\mathbb{E}_w \Big(c(s, a, w) + V^{(k)}(h(s, a, w)) \Big)$ is jointly continuous in a and s for $s \neq s^*$.

(b) The infimum in (45) is achieved in $[0, \infty)$ for all $s \neq s^*$.

(c) For each $s \in \mathcal{S}$, there exists $a(s) < \infty$ such that $a(s)$ is continuous in s for $s \neq s^*$, and no minimizer of (45) lies in $(a(s), \infty)$ for each $k \geq 0$.

Then the following hold:

(i) The value iteration converges, i.e., $V^{(k)}(s) \to V^{(\infty)}(s)$ for all $s \neq s^*$.

(ii) $V^{(\infty)}(s) = V^*(s)$ for all $s \neq s^*$.

(iii) $\Gamma_\infty(s) \subset \Gamma^*(s)$ for all $s \neq s^*$.

(iv) There is a stationary optimal policy f^∞ where $f : \mathcal{S} \setminus \{s^*\} \to \mathcal{A}$ and $f(s) \in \Gamma_\infty(s) \ \forall \ s \neq s^*$. □

Proof of Theorem 11: By Theorem 8, the value iteration converges, i.e., $V^{(k)}(s) \to V^{(\infty)}(s)$. Moreover, $V^{(k)}(s)$ is the optimal cost for a k-stage problem with zero terminal cost, and the cost at each stage is positive. Hence, $V^{(k)}(s)$ increases in k for every $s \in \mathcal{S}$. Thus, for all $s \in \mathcal{S}$, $V^{(k)}(s) \uparrow V^{(\infty)}(s)$.

Now, Condition 2 and Theorem 10 say that if no optimizer of the value iteration in each stage k lies outside a compact subset $\underline{\mathcal{A}}(s)$ of $\mathcal{A}(s) \subset \mathcal{A}$, then we can deal with the modified problem having a new action space $\underline{\mathcal{A}}(s)$. If the value iteration converges to the optimal value in this modified problem, then it will converge to the optimal value in the original problem as well, provided that the mapping $\underline{\mathcal{A}} : \mathcal{S} \to \mathcal{C}(\mathcal{A})$ is measurable. Let us choose $\underline{\mathcal{A}}(s) := [0, a(s)]$, where $a(s)$ satisfies hypothesis (c) of Theorem 11. Since $a(s)$ is continuous at $s \neq s^*$, for any $\epsilon > 0$ we can find a $\delta_{s, \epsilon} > 0$ such that $|a(s) - a(s')| < \epsilon$ whenever $|s - s'| < \delta_{s, \epsilon}$, $s \neq s^*$, $s' \neq s^*$. Now, when $|a(s) - a(s')| < \epsilon$, we have $d([0, a(s)], [0, a(s')]) < \epsilon$. Hence, the mapping $\underline{\mathcal{A}} : \mathcal{S} \to \mathcal{C}(\mathcal{A})$ is continuous at all $s \neq s^*$, and thereby measurable in this case. Hence, the value iteration (45) satisfies Condition 2.

Thus, the value iteration for $s \neq s^*$ can be modified as:

$$V^{(k+1)}(s) = \inf_{a \in [0, a(s)]} \mathbb{E}_w(c(s, a, w) + V^{(k)}(h(s, a, w))) \qquad (46)$$

Now, $\mathbb{E}_w(c(s, a, w) + V^{(k)}(h(s, a, w)))$ is continuous (can be discontinuous at $s = s^*$, since this quantity is 0 at $s = s^*$) on $\mathcal{S} \times \mathcal{A}$ (by our hypothesis). Hence, $\mathbb{E}_w(c(s, a, w) + V^{(k)}(h(s, a, w)))$ is measurable on $\mathcal{S} \times \mathcal{A}$. Also, it is bounded below by 0. Hence, it can be approximated by an increasing sequence of bounded measurable functions $\{v_{n,k}\}_{n \geq 1}$ given by $v_{n,k}(s, a) = \min\{\mathbb{E}_w(c(s, a, w) + V^{(k)}(h(s, a, w))), n\}$. Hence, $\mathbb{E}_w(c(s, a, w) + V^{(k)}(h(s, a, w)))$ is in $\hat{\mathcal{F}}(\mathcal{S} \times \mathcal{A})$.

Thus, Condition 1 is satisfied for the modified problem and therefore, by Theorem 9, the modified value iteration in (46) converges to the optimal value function. Now, by Theorem 10, we can argue that the value iteration (45) converges to the optimal value function in the original problem and hence $V^{(\infty)}(s) = V^*(s)$ for all $s \in \mathcal{S} \setminus s^*$. Also, $\Gamma_\infty(s) \subset \Gamma^*(s)$ for all $s \in \mathcal{S} \setminus s^*$ and there exists a stationary optimal policy f^∞ where $f(s) \in \Gamma_\infty(s)$ for all $s \in \mathcal{S} \setminus s^*$ (by Theorem 9).

G.2 Proof of Theorem 7

This proof uses the results of Theorem 11 provided in this appendix. Remember that the state **EOL** is absorbing and $c(\mathbf{EOL}, a, w) = 0$ for all a, w. We can think of it as state 0 so that our state space becomes $[0, 1]$ which is a Borel set. We will see that the state 0 plays the role of the state s^* as mentioned in Theorem 11.

We need to check whether the conditions (a), (b), and (c) in Theorem 11 are satisfied for the value iteration (21). Of course, $J_\xi^{(0)}(s) = 0$ is concave, increasing in $s \in (0, 1]$. Suppose that $J_\xi^{(k)}(s)$ is concave, increasing in s for some $k \geq 0$. Also, for any fixed $a \geq 0$, $\frac{se^{\rho a}}{1 + se^{\rho a}}$ is concave and increasing in s. Thus, by the composition rule for the composition of a concave increasing function $J_\xi^{(k)}(\cdot)$ and a concave increasing function $\frac{se^{\rho a}}{1 + se^{\rho a}}$, for any $a \geq 0$ the term $J_\xi^{(k)}\left(\frac{se^{\rho a}}{1 + se^{\rho a}}\right)$ is concave, increasing over $s \in (0, 1]$. Hence, $\int_0^a \beta e^{-\beta z} s(e^{\rho z} - 1)dz + e^{-\beta a}\left(s(e^{\rho a} - 1) + \xi + J_\xi^{(k)}\left(\frac{se^{\rho a}}{1 + se^{\rho a}}\right)\right)$ (in (21)) is concave increasing over $s \in (0, 1]$. Since the infimization over a preserves concavity, we conclude that $J_\xi^{(k+1)}(s)$ is concave, increasing over $s \in (0, 1]$. Hence, for each k, $J_\xi^{(k)}(s)$ is continuous in s over $(0, 1)$, since otherwise concavity w.r.t. s will be violated. Now, we must have $J_\xi^{(k)}(1) \leq \lim_{s \uparrow 1} J_\xi^{(k)}(s)$, since otherwise the concavity of $J_\xi^{(k)}(s)$ will be violated. But since $J_\xi^{(k)}(s)$ is increasing in s, $J_\xi^{(k)}(1) \geq \lim_{s \uparrow 1} J_\xi^{(k)}(s)$. Hence, $J_\xi^{(k)}(1) = \lim_{s \uparrow 1} J_\xi^{(k)}(s)$. Thus, $J_\xi^{(k)}(s)$ is continuous in s over $(0, 1]$ for each k.

Hence, $\int_0^a \beta e^{-\beta z} s(e^{\rho z} - 1)dz + e^{-\beta a}(s(e^{\rho a} - 1) + \xi + J_\xi^{(k)}\left(\frac{se^{\rho a}}{1 + se^{\rho a}}\right))$ is continuous in s, a for $s \neq 0$. Hence, condition (a) in Theorem 11 is satisfied.

Now, we will check condition (c) in Theorem 11.

By Theorem 8, the value iteration converges, i.e., $J_\xi^{(k)}(s) \to J_\xi^{(\infty)}(s)$. Also, $J_\xi^{(\infty)}(s)$ is concave, increasing in $s \in (0, 1]$ and hence continuous. Moreover, $J_\xi^{(k)}(s)$ is the optimal cost for a k-stage problem with zero terminal cost, and the cost at each stage is positive. Hence, $J_\xi^{(k)}(s)$ increases in k for every $s \in (0, 1]$. Thus, for all $s \in (0, 1]$, $J_\xi^{(k)}(s) \uparrow J_\xi^{(\infty)}(s)$.

Again, $J_\xi^{(k)}(s)$ is the optimal cost for a k-stage problem with zero terminal cost. Hence, it is less than or equal to the optimal cost for the infinite horizon problem with the same transition law and cost structure. Hence, $J_\xi^{(k)}(s) \leq J_\xi(s)$ for all $k \geq 1$. Since $J_\xi^{(k)}(s) \uparrow J_\xi^{(\infty)}(s)$, we have $J_\xi^{(\infty)}(s) \leq J_\xi(s)$.

Now, consider the following two cases:

G.2.1 $\beta > \rho$

Let us define a function $\psi : (0, 1] \to \mathbb{R}$ by $\psi(s) = \frac{J_\xi^{(\infty)}(s) + \theta s}{2}$. By Proposition 1, $J_\xi(s) < \theta s$ for all $s \in (0, 1]$. Hence, $J_\xi^{(\infty)}(s) < \psi(s) < \theta s$ and $\psi(s)$ is continuous over $s \in (0, 1]$. Since $\beta > \rho$ and $J_\xi^{(k)}(s) \in [0, \theta]$ for any s in $(0, 1]$, the expression $\theta s + e^{-\beta a}\left(-\theta se^{\rho a} + \xi + J_\xi^{(k)}\left(\frac{se^{\rho a}}{1 + se^{\rho a}}\right)\right)$ obtained from the R.H.S of (21) converges to θs as $a \to \infty$. A lower bound to this expression is $\theta s + e^{-\beta a}(-\theta se^{\rho a})$. With $\beta > \rho$, for each s, there exists $a(s) < \infty$ such that $\theta s + e^{-\beta a}(-\theta se^{\rho a}) > \psi(s)$ for all $a > a(s)$. But $\theta s + \inf_{a \geq 0} e^{-\beta a}\left(-\theta se^{\rho a} + \xi + J_\xi^{(k)}\left(\frac{se^{\rho a}}{1 + se^{\rho a}}\right)\right)$ is equal to $J_\xi^{(k+1)}(s) < \psi(s)$. Hence, for any $s \in (0, 1]$, the minimizers for (18) always lie in the compact interval $[0, a(s)]$ for all $k \geq 1$. Since $\psi(s)$ is continuous in s, we can choose $a(s)$ as a continuous function of s on $(0, 1]$.

G.2.2 $\beta \leq \rho$

Fix A, $0 < A < \infty$. Let $K := \frac{1}{\beta A}\left(\xi + (e^{\rho A} - 1)\right) + (e^{\rho A} - 1)$. Then, by Proposition 2, $J_\xi(s) \leq K$ for all $s \in (0, 1]$. Now, we observe that the objective function (for minimization over a) in the R.H.S of (21) is lower bounded by $\int_0^a \beta e^{-\beta z} s(e^{\rho z} - 1)dz$, which is continuous in s, a and goes to ∞ as $a \to \infty$ for each $s \in (0, 1]$. Hence, for each $s \in (0, 1]$, there exists $0 < a(s) < \infty$ such that $\int_0^a \beta se^{-\beta z}(e^{\rho z} - 1)dz > 2K$ for all $a > a(s)$ and $a(s)$ is continuous over $s \in (0, 1]$. But $J_\xi^{(k+1)}(s) \leq J_\xi(s) \leq K$ for all k. Hence, the minimizers in (21) always lie in $[0, a(s)]$ where $a(s)$ is independent of k and continuous over $s \in (0, 1]$.

Let us set $a(0) = a(1)$.[8] Then, the chosen function $a(s)$ is continuous over $s \in (0, 1]$ and can be discontinuous only at $s = 0$. Thus, condition (c) of Theorem 11 has been verified for the value iteration (21). Condition (b)

---

8. Remember that at state 0 (i.e., state **EOL**), the single stage cost is 0 irrespective of the action, and that this state is absorbing. Hence, any action at state 0 can be optimal.

of Theorem 11 is obviously satisfied since a continuous function over a compact set always has a minimizer. $\square$

Remark: Observe that in our value iteration (21) it is always sufficient to deal with compact action spaces, and the objective functions to be minimized at each stage of the value iteration are continuous in s, a. Hence, $\Gamma_k(s)$ is nonempty for each $s \in (0,1]$, $k \geq 0$. Also, since there exists $K > 0$ such that $J_\xi^{(k)}(s) \leq K$ for all $k \geq 0$, $s \in (0,1]$, it is sufficient to restrict the action space in (21) to a set $[0, a(s)]$ for any $s \in (0,1]$, $k \geq 0$. Hence, $\Gamma_k(s) \subset [0, a(s)]$ for all $s \in (0,1]$, $k \geq 0$. Now, for a fixed $s \in (0,1]$, any sequence $\{a_k\}_{k \geq 0}$ with $a_k \in \Gamma_k(s)$, in bounded. Hence, the sequence must have a limit point. Hence, $\Gamma_\infty(s)$ is nonempty for each $s \in (0,1]$. Since $\Gamma_\infty(s) \subset \Gamma^*(s)$, $\Gamma^*(s)$ is nonempty for each $s \in (0,1]$.

APPENDIX H
PROOFS OF PROPOSITIONS 3, 4 AND 5

H.1 Proof of Proposition 3

Fix ξ. Consider the value iteration (21). Let us start with $J_\xi^{(0)}(s) := 0$ for all $s \in (0,1]$. Clearly, $J_\xi^{(1)}(s)$ is concave and increasing in s, since pointwise infimum of linear functions is concave. Now let us assume that $J_\xi^{(k)}(s)$ is concave and increasing in s. Then, by the composition rule, it is easy to show that $J_\xi^{(k)}\left(\frac{se^{\rho a}}{1+se^{\rho a}}\right)$ is concave and increasing in s for any fixed $a \geq 0$. Hence, $J_\xi^{(k+1)}(s)$ is concave and increasing, since pointwise infimum of a set of concave and increasing functions is concave and increasing. By Theorem 7, $J_\xi^{(k)}(s) \to J_\xi(s)$. Hence, $J_\xi(s)$ is concave and increasing in s. $\square$

H.2 Proof of Proposition 4

Consider the value iteration (21). Since $J_\xi^{(0)}(s) := 0$ for all $s \in (0,1]$, $J_\xi^{(1)}(s)$ is obtained by taking infimum (over a) of a linear, increasing function of ξ. Hence, $J_\xi^{(1)}(s)$ is concave, increasing over $\xi \in (0, \infty)$. If we assume that $J_\xi^{(k)}(s)$ is concave and increasing in ξ, then $J_\xi^{(k)}\left(\frac{se^{\rho a}}{1+se^{\rho a}}\right)$ is also concave and increasing in ξ for fixed s and a. Thus, $J_\xi^{(k+1)}(s)$ is also concave and increasing in ξ. Now, $J_\xi^{(k)}(s) \to J_\xi(s)$ for all $s \in \mathcal{S}$, and $J_\xi^{(k)}(s)$ is concave, increasing in ξ for all $k \geq 0$, $s \in \mathcal{S}$. Hence, $J_\xi(s)$ is concave and increasing in ξ. $\square$

H.3 Proof of Proposition 5

Clearly, $J_\xi(s)$ is continuous in s over $(0,1)$, since otherwise concavity w.r.t. s will be violated. Now, since $J_\xi(s)$ is concave in s over $(0,1]$, we must have $J_\xi(1) \leq \lim_{s \uparrow 1} J_\xi(s)$. But since $J_\xi(s)$ is increasing in s, $J_\xi(1) \geq \lim_{s \uparrow 1} J_\xi(s)$. Hence, $J_\xi(1) = \lim_{s \uparrow 1} J_\xi(s)$. Thus, $J_\xi(s)$ is continuous in s over $(0,1]$.

Again, for a fixed $s \in (0,1]$, $J_\xi(s)$ is concave and increasing in ξ. Hence, $J_\xi(s)$ is continuous in ξ over $\xi \in (0, c)$, $\forall c > 0$. Hence, $J_\xi(s)$ is continuous in ξ over $(0, \infty)$. $\square$