

Nobody Expects a Differential Equation: Minimum Energy-Per-Bit for the Gaussian Relay Channel with Rank-1 Linear Relaying

Oliver Kosut
School of Electrical, Computer
and Energy Engineering
Arizona State University
Tempe, AZ, USA
Email: okosut@asu.edu

Michelle Effros
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA, USA
Email: effros@caltech.edu

Michael Langberg
Department of Electrical Engineering
University at Buffalo
Buffalo, NY, USA
email: mikel@buffalo.edu

Abstract—Motivated by the design of low-complexity low-power coding solutions for the Gaussian relay channel, this work presents an upper bound on the minimum energy-per-bit achievable on the Gaussian relay channel using rank-1 linear relaying. Our study addresses high-dimensional relay codes and presents bounds that outperform prior known bounds using 2-dimensional schemes. A novelty of our analysis ties the optimization problem at hand to the solution of a certain differential equation which, in turn, leads to a low energy-per-bit achievable scheme.

I. INTRODUCTION

As communication capabilities are built into an increasingly diverse array of technologies, coding strategies are needed for more (and often more constrained) communication scenarios. In some cases, the constraints are not new. Worries about power limitations, for example, and characterizations of the minimal energy required, on average, to reliably deliver each bit of information across a noisy channel (the so called “minimum energy-per-bit”) date all the way back to Shannon [1]. Given the proliferation of low-power wireless devices with built-in communication capabilities, such characterizations are, arguably, even more relevant today than they were in 1949 when Shannon proposed the question and derived an asymptotic solution of the minimum energy-per-bit for a point-to-point channel with additive, white Gaussian noise. Wireless communication devices motivate much of the ongoing work on this problem. Examples of results in this area include, among many others, characterizations of the minimum energy-per-bit in the finite blocklength domain with and without feedback [2], in broadcast and interference channels with correlated information [3], and in relay channels [4].

Motivated by wireless communication scenarios where power constraints are especially restrictive, this work provides a bound on the minimum energy-per-bit for the relay channel under a constrained family of codes. We choose the relay channel since it is well matched to low-power devices that may be unable to meet their communication goals without the aid of a helper, here described as a relay node. We maintain low coding complexity by using a constrained family of codes

— namely, where the relay performs only linear operations, and where the transmitter sends a Gaussian vector whose covariance matrix has rank 1. Our desire to limit coding complexity is motivated by the observation that devices with tight power constraints are often also extremely constrained in computational resources. In some cases, limits on computation can be traced back to power constraints; in particular, while the energy costs for using a code are typically far smaller than those for transmission, at extremely low power the two can become comparable, leaving the code operation to compete for resources with the cost of transmission.¹ Complexity constraints can also arise when constraints on price, weight, or size result in limitations on the computational resources on board a given communication device.

The main result of this work is an upper bound on the minimum energy-per-bit achievable on the Gaussian relay channel using the family of low-complexity rank-1 linear relay codes, described in Section II. The minimum energy-per-bit achievable by this family of codes can be posed as a non-convex optimization problem. We analyze the necessary optimality conditions for this problem and find that taking a limit of these optimality conditions leads to a system of differential equations.² Solutions to this system gives rise to our achievable bound, presented in Section III and proved via a number of lemmas in Sections IV and V.

II. PROBLEM DESCRIPTION

Notation: We use $\log$ to denote log base-2, and $\ln$ for natural log. For any $p \geq 1$, $\|x\|_p$ denotes the p -norm of x ; when the subscript is left off we mean 2-norm. The set of real numbers is denoted $\mathbb{R}$ and the set of positive integers $\mathbb{N}$. For an integer M , $[M]$ denotes $\{1, 2, \dots, M\}$.

¹The costs of code operation are device specific and are not included in any of the prior minimum energy-per-bit characterizations cited above or in our characterization.

²The title is a reference to the Monty Python sketch [5]; at the outset the authors were certainly not expecting to find a differential equation.

The Gaussian relay channel has two gain parameters a and b . At time instance i , the transmitted signals at the transmitter and the relay are denoted X_i and X_{ri} respectively. The received signals at the relay and receiver are

$$Y_{ri} = aX_i + Z_{ri}, \quad Y_i = X_i + bX_{ri} + Z_i \quad (1)$$

where $\{Z_i\}$ and $\{Z_{ri}\}$ are independent white Gaussian noise sequences with mean 0 and variance 1. We assume without loss of generality that a and b are non-negative. In addition, we assume that both are strictly positive, since if either is 0, the relay is effectively removed from the network.

An (n, M) -code with strict causality at the relay is given by encoding functions at the transmitter and relay

$$x^n : [M] \rightarrow \mathbb{R}^n, \quad x_{ri} : \mathbb{R}^{i-1} \rightarrow \mathbb{R}, \quad i = 1, \dots, n \quad (2)$$

and a decoding function $\hat{m} : \mathbb{R}^n \rightarrow [M]$. The probability of error is

$$P_e^{(n)} = \frac{1}{M} \sum_{m=1}^M \mathbb{P}\{\hat{m}(Y^n) \neq m | X^n = x^n(m), X_{ri} = x_{ri}(Y_r^{i-1})\} \quad (3)$$

where Y_{ri}, Y_i are related to X_i, X_{ri} by (1). The energy-per-bit is denoted

$$\mathcal{E}^{(n)} = \frac{\max_m \|x^n(m)\|^2 + \sup_{y_r^n} \sum_{i=1}^n x_{ri}(y_r^{i-1})^2}{\log M}. \quad (4)$$

We say energy-per-bit $\mathcal{E}$ is *achievable* if there exists a sequence of codes where $P_e^{(n)} \rightarrow 0$ and $\limsup \mathcal{E}^{(n)} \leq \mathcal{E}$. The *minimum energy-per-bit* $\mathcal{E}^*$ is the infimum of all achievable energies-per-bit.

A linear relay code is one where, at every time instant, the relay sends a linear combination of its received signals. This relaying can be described by a strictly lower-triangular matrix D , to ensure that coding at the relay is strictly causal. Under this assumption, it is optimal that the transmitter send a Gaussian vector. Thus, the code can be described by the matrix D and the covariance matrix Σ for the transmitted vector. Let $\mathcal{E}_{\text{LR}}^*$ be the minimum energy-per-bit achievable by linear codes. From results in [4], it follows that $\mathcal{E}_{\text{LR}}^*$ is equal to

$$\inf_{\substack{k \in \mathbb{N}, \\ \Sigma \in \mathcal{S}_+^k, \\ D \in \mathcal{L}^k}} \frac{1}{2} \log \frac{\text{tr}(\Sigma + a^2 D \Sigma D^T + D D^T) \det((I + abD)\Sigma(I + abD^T) + I + b^2 D D^T)}{\det(I + b^2 D D^T)} \quad (5)$$

where $\mathcal{S}_+^k$ is the set of positive semi-definite $k \times k$ matrices, and $\mathcal{L}^k$ is the set of strictly lower triangular $k \times k$ matrices. In addition, let $\mathcal{E}_{\text{LR}}^*$ be the value of (5) where Σ is constrained to be rank-1. Obviously $\mathcal{E}^* \leq \mathcal{E}_{\text{LR}}^* \leq \mathcal{E}_{\text{ILR}}^*$. Our main result will be an upper bound on $\mathcal{E}_{\text{ILR}}^*$.

A. Prior Bounds

In [4] it is shown that the block-Markov achievable scheme (i.e., partial decode-forward) from [6] leads to the upper bound on the minimum energy-per-bit

$$\mathcal{E}^* \leq 2 \ln 2 \min \left\{ 1, \frac{a^2 + b^2}{a^2(1 + b^2)} \right\}. \quad (6)$$

It is also shown in [4] that linear relaying can outperform this bound, using a $k = 2$ -dimensional linear scheme given by

$$\Sigma = 2P_1 \begin{bmatrix} \beta & \sqrt{\beta(1-\beta)} \\ \sqrt{\beta(1-\beta)} & 1-\beta \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ d & 0 \end{bmatrix} \quad (7)$$

where $\beta \in [0, 1]$ and $d = \sqrt{\frac{2P_2}{2a^2\beta P_1 + 1}}$, and where P_1, P_2 are the powers used at the transmitter and relay respectively. Note that Σ in (7) has rank 1.

In addition, [4] found that the cut-set bound from [6] leads to the lower bound on the minimum energy-per-bit given by

$$\mathcal{E}^* \geq 2 \ln 2 \frac{1 + a^2 + b^2}{(1 + a^2)(1 + b^2)}. \quad (8)$$

In [7], an LTI filtering approach was taken to coding for the relay channel. These also constitute linear codes, although for the strictly causal case for channels with constant gains, the achievability bounds do not appear to be any better than those from [4]. Linear codes for the Gaussian relay channel were also considered in [8], which derived methods to simplify the optimization problem for Σ and D for a given dimension k . This approach is interestingly complementary to ours, in that [8] finds optimal Σ for a given D , whereas we focus on finding optimal D for a given Σ . Even with these simplifications, finding optimal (Σ, D) pairs is challenging, particularly for large k , so we have chosen to omit comparisons in this paper.

III. MAIN RESULT

To state our result, we first need the following lemma, proved in [9].

Lemma 1: Fix any $A_f, B_f > 0$ where $A_f/B_f \leq a^2$. Let³

$$\phi = A_f B_f + \frac{1}{A_f} - \frac{1}{B_f}, \quad (9)$$

$$f(w) = \frac{\phi w - 1 + \sqrt{(\phi w - 1)^2 + 4w^3}}{2w^2}. \quad (10)$$

There exists a unique pair (A_0, ψ) such that $A_0 \geq A_f, \psi > 0$ and

$$\int_{A_f}^{A_0} \frac{f(w) dw}{1 + w f(w)^2} = \frac{A_0}{a\psi} - \frac{1}{B_f}, \quad (11)$$

$$\int_{A_f}^{A_0} \frac{f(w)^2 dw}{1 + w f(w)^2} = \ln \left(\frac{A_0^3 B_f}{a^4 \psi^2} \right). \quad (12)$$

Now we make some definitions that will be used to state our main result. Given any $A_f, B_f > 0$ satisfying $A_f/B_f \leq a^2$, define $\mathcal{E}(A_f, B_f)$ as follows. First let (A_0, ψ) be the pair from Lemma 1. Also let $f(w)$ be the function in (10). Then define

$$\mathcal{E}(A_f, B_f) = \frac{Q_1 + Q_2}{\frac{1}{2} \log \frac{A_0}{a^2} \left(\frac{1}{B_f} + A_0 B_0 - A_f B_f \right)} \quad (13)$$

where

$$B_0 = f(A_0), \quad Q_1 = -\frac{1}{a^2} + \frac{A_0^3 B_f}{a^6 \psi^2}, \quad (14)$$

³Note that f implicitly depends on A_f, B_f through ϕ .

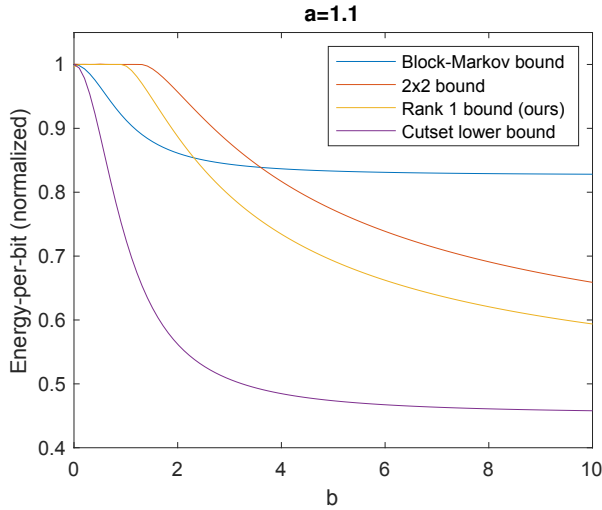


Fig. 1. Comparison between bounds on the minimum energy-per-bit for the Gaussian relay channel. Bounds are computed for channel parameters $a = 1.1$ and $b \in [0, 10]$. The normalized energy-per-bit $\mathcal{E}/(2 \ln 2)$ is plotted for each bound. The block-Markov achievable bound is (6), the 2×2 bound is from (7), the rank 1 bound is from Thm. 1, and the cut-set bound is (8).

$$Q_2 = -\frac{1}{b^2} + \frac{A_0^3}{a^5 b^2 \psi^3} + \frac{A_0^2 (A_f B_f^2 - 1)}{a^4 b^2 \psi^2 B_f}. \quad (15)$$

Theorem 1: For any $a, b > 0$,

$$\mathcal{E}_{\text{ILR}}^* \leq \inf_{A_f, B_f > 0: A_f/B_f \leq a^2} \mathcal{E}(A_f, B_f). \quad (16)$$

Theorem 1 is derived from the optimality conditions for the optimization problem 5 when Σ is rank-1, then taking a limit as the energy per symbol goes to 0, and analyzing the resulting differential equation. For this reason, we suspect that the bound in Thm. 1 is optimal or close to optimal among rank-1 linear codes. The details of the proof of Theorem 1 make up the remainder of the paper. We have attempted to make the main steps of the proof readable in the body of the paper, while pushing more technical elements of the proof into Lemmas 2–5, which are proved in the [9].

Numerical results for the bound from Thm. 1 compared to bounds mentioned in Sec. II-A are shown in Fig. 1. As shown in the figure, our bound consistently outperforms the $k = 2$ bound in (7), and it sometimes outperforms the block-Markov bound in (6).

IV. TRANSFORMING THE OPTIMIZATION PROBLEM

The first step toward proving Thm. 1 is to transform the optimization problem in (5) under the assumption that Σ has rank 1. Specifically let $\Sigma = ss^T$ where $s \in \mathbb{R}^k$. Let the matrix G be such that $GG^T = (I + b^2 DD^T)^{-1}$, and let $v = G^T(I + abD)s$. Then the quantity inside the log in (5) is

$$\frac{\det((I + abD)ss^T(I + abD^T) + I + b^2 DD^T)}{\det(I + b^2 DD^T)} \quad (17)$$

$$= \det[G^T((I + abD)ss^T(I + abD^T) + I + b^2 DD^T)G] \quad (18)$$

$$= \det[G^T(I + abD)ss^T(I + abD^T)G + I] \quad (19)$$

$$= \det(I + vv^T) \quad (20)$$

$$= 1 + \|v\|^2 \quad (21)$$

$$= 1 + s^T(I + abD^T)GG^T(I + abD)s \quad (22)$$

$$= 1 + s^T(I + abD^T)(I + b^2 DD^T)^{-1}(I + abD)s \quad (23)$$

where (21) follows because the eigenvalues of $I + vv^T$ are $1 + \|v\|^2$ and 1, the latter with $k - 1$ repetitions. We thus define, for any $k, s \in \mathbb{R}^k, D \in \mathbb{R}^{k \times k}$,

$$\mathcal{E}_{\text{ILR}}(s, D) = \frac{\|s\|^2 + a^2 \|Ds\|^2 + \text{tr}(DD^T)}{\frac{1}{2} \log(1 + s^T(I + abD^T)(I + b^2 DD^T)^{-1}(I + abD)s)}. \quad (24)$$

Also define

$$\mathcal{E}_{\text{ILR}}^*(s) = \sup_{D \in \mathcal{L}^k} \mathcal{E}_{\text{ILR}}(s, D). \quad (25)$$

Thus

$$\mathcal{E}_{\text{ILR}}^* = \sup_{k \in \mathbb{N}} \sup_{s \in \mathbb{R}^k} \mathcal{E}_{\text{ILR}}^*(s). \quad (26)$$

We now make the following definitions, which will allow us to analyze the optimization problem over D in (25):

$$u = Ds, \quad (27)$$

$$z = (I + b^2 DD^T)^{-1}(s + abu), \quad (28)$$

$$r = abs - b^2 D^T z. \quad (29)$$

Given s and D , z is well-defined since $I + b^2 DD^T$ is positive definite, and so it is invertible. With these definitions, (23) becomes $1 + (s + abu)^T z$. In addition,

$$s + abu = (I + b^2 DD^T)z = z + D(b^2 D^T z) \quad (30)$$

$$= z + D(abs - r) = z + abu - Dr. \quad (31)$$

Thus, $Dr = z - s$. Now for any s , $\mathcal{E}_{\text{ILR}}^*(s)$ is the optimal value of the optimization problem

$$\begin{aligned} & \underset{u, z, r, D}{\text{minimize}} && \frac{\|s\|^2 + a^2 \|u\|^2 + \text{tr}(DD^T)}{\frac{1}{2} \log(1 + (s + abu)^T z)}, \\ & \text{subject to} && Ds = u, \\ & && Dr = z - s, \\ & && b^2 D^T z = abs - r, \\ & && D_{ij} = 0, \quad j \geq i. \end{aligned} \quad (32)$$

The next step is to write optimality conditions for this optimization problem. Since the problem is non-convex, these are necessary but not sufficient conditions. Specifically, in the following lemma, proved in [9], we find necessary conditions for optimal u, z, r, D given s . While this lemma is not actually necessary to the proof of Thm. 1, we include it as motivation for the system of differential equations that we study in Sec. V.

Lemma 2: Fix $k \in \mathbb{N}$ and any vector $s \in \mathbb{R}^k$ such that $\|s\|^2 = Q_1 > 0$. Let u, z, r, D solve (32). Let

$$S_i = \sum_{j: j \leq i} s_j^2, \quad T_i = \sum_{j: j \leq i} r_j s_j, \quad R_i = \sum_{j: j \leq i} r_j^2, \quad (33)$$

$$V_i = \sum_{j:j>i} u_j z_j, \quad Z_i = \sum_{j:j>i} z_j^2. \quad (34)$$

There exists $\lambda > 0$ such that

$$D_{ij} = -a^2 u_i s_j + \frac{z_i r_j}{\lambda}, \quad j < i, \quad (35)$$

$$u_i = \frac{T_{i-1} s_i}{(1 + a^2 S_{i-1})(\lambda - R_{i-1}) + a^2 T_{i-1}^2}, \quad (36)$$

$$z_i = \frac{\lambda(1 + a^2 S_{i-1}) s_i}{(1 + a^2 S_{i-1})(\lambda - R_{i-1}) + a^2 T_{i-1}^2}, \quad (37)$$

$$r_i = \frac{\lambda(ab + a^2 b^2 V_i) s_i}{\lambda + b^2 Z_i}, \quad (38)$$

$$(s + abu)^T z = Z_0 + a^2 Q_1 - \frac{2aT_k}{b} + \frac{R_k}{b^2}, \quad (39)$$

$$a^2 \|u\|^2 + \text{tr}(DD^T) = \frac{aT_k}{b\lambda} - \frac{R_k}{b^2 \lambda}. \quad (40)$$

V. FROM OPTIMALITY CONDITIONS TO A DIFFERENTIAL EQUATION

Consider a fixed vector $s \in \mathbb{R}^k$ and the corresponding running sum S_i defined in (33). From the definitions in (33)–(34), the remaining sequences satisfy the difference equations

$$V_i = V_{i-1} - u_i z_i, \quad Z_i = Z_{i-1} - z_i^2, \quad (41)$$

$$T_i = T_{i-1} + r_i s_i, \quad R_i = R_{i-1} + r_i^2. \quad (42)$$

In turn, u_i and z_i depend on T_{i-1} and R_{i-1} , and r_i depends on V_i, Z_i . Thus, these sequences can be computed in order. Moreover, these computations can be viewed as a variation on Euler's method to approximate the solution to a differential equation. This differential equation emerges from Lemma 2 by taking a limit as $\max_i s_i^2 \rightarrow 0$. Since $S_0 = 0$ and $S_k = \|s\|^2 = Q_1$, we use S as the independent variable, where S goes from 0 to Q_1 . We use $(\cdot)'$ for $\frac{d}{dS}(\cdot)$. The system of differential equations is

$$u = \frac{T}{(1 + a^2 S)(\lambda - R) + a^2 T^2}, \quad (43)$$

$$z = \frac{\lambda(1 + a^2 S)}{(1 + a^2 S)(\lambda - R) + a^2 T^2}, \quad (44)$$

$$r = \frac{\lambda(ab + a^2 b^2 V)}{\lambda + b^2 Z}, \quad (45)$$

$$T' = r, \quad R' = r^2, \quad V' = -uz, \quad Z' = -z^2. \quad (46)$$

We also have the boundary conditions

$$T(0) = R(0) = V(Q_1) = Z(Q_1) = 0. \quad (47)$$

In order to show that this differential equation does in fact result from taking a limit of the optimality conditions in Lemma 2, we need a very general lemma regarding a variation on Euler's method for approximating a differential equation. Specifically, consider an arbitrary multi-dimensional differential equation

$$\theta'(t) = F(\theta(t), t) \quad (48)$$

where $\theta(t) \in \mathbb{R}^m$. The usual Euler's method is to fix a step size $\Delta > 0$, and form a sequence by initializing $\theta_0 = \theta(0)$, and then proceeding forward by

$$\theta_{i+1} = \theta_i + \Delta F(\theta_i, i\Delta). \quad (49)$$

A careful consideration of (41)–(42) reveals that this is not quite what is happening there, because the updates of T_i and R_i use r_i , which is a function of the just-calculated values of V_i, Z_i , rather than those from the previous round, V_{i-1}, Z_{i-1} . Thus, we need to analyze a slight variation on Euler's method, wherein the entries of θ are updated one-by-one, each using the most recent value of the other entries when calculating the next step. Specifically, for any $j \in \{0, \dots, m\}$, let

$$\theta_i^{(j)} = (\theta_{i+1,1}, \dots, \theta_{i+1,j-1}, \theta_{i,j}, \dots, \theta_{i,m}). \quad (50)$$

Now we replace the update rule in (49) by

$$\theta_{i+1,j} = \theta_{i,j} + \Delta F_j(\theta_i^{(j)}, i\Delta), \quad j = 1, \dots, m. \quad (51)$$

The following lemma, proved in [9], shows that, for any differential equation that is sufficiently well-behaved, this modified Euler's method approaches the true solution as the step size goes to zero. The proof of the lemma is similar to the proof that the global error of the standard Euler's method vanishes with step size, as in [10, Thm. 12.2].

Lemma 3: Let $\theta(t) \in \mathbb{R}^m$ for $t \in [0, T]$ satisfy the differential equation in (48), and let $\theta_0, \theta_1, \dots$ be derived from the update rule (51) with initial condition $\theta_0 = \theta(0)$ and step size Δ . Assume there exist sets $\mathcal{A}(t) \subset \mathbb{R}^m$ for $t \in [0, T]$ such that

$$\inf\{\|\theta(t) - \theta\| : t \in [0, T], \theta \in \mathbb{R}^m \setminus \mathcal{A}(t)\} > 0. \quad (52)$$

Also, assume there exist constants K_1, K_2, K_3 such that, for all $t \in [0, T]$, $\theta, \tilde{\theta} \in \mathcal{A}(t)$, and all $j \in \{1, \dots, m\}$

$$|F_j(\theta, t) - F_j(\tilde{\theta}, t)| \leq K_1 \|\theta - \tilde{\theta}\|, \quad (53)$$

$$\left| \frac{\partial}{\partial t} F_j(\theta, t) \right| \leq K_2, \quad (54)$$

$$\|F(\theta, t)\| \leq K_3. \quad (55)$$

Then

$$\lim_{\Delta \rightarrow 0} \max_{i=0, \dots, \lfloor T/\Delta \rfloor} \|\theta(i\Delta) - \theta_i\| = 0. \quad (56)$$

The following lemma, proved in [9], shows that our differential equation of interest in (43)–(47) does in fact satisfy the assumptions of Lemma 3, at least under certain conditions, and thus the sequences defined by the optimality conditions in Lemma 2 do approach the differential equation solutions, which leads to an achievable energy-per-bit.

Lemma 4: Fix any $\lambda > 0$ and $Q_1 > 0$. Let u, z, r, T, R, V, Z be functions that solve the differential equation system (43)–(47) where

$$\inf_{S \in [0, Q_1]} (1 + a^2 S)(\lambda - R(S)) + a^2 T(S)^2 > 0, \quad (57)$$

$$\sup_{S \in [0, Q_1]} \max\{|T(S)|, |R(S)|, |V(S)|, |Z(S)|\} < \infty. \quad (58)$$

Then

$$\mathcal{E}_{\text{ILR}}^* \leq \frac{Q_1 + \frac{aT(Q_1)}{b\lambda} - \frac{R(Q_1)}{b^2\lambda}}{\frac{1}{2} \log \left(1 + Z(0) + a^2 Q_1 - \frac{2aT(Q_1)}{b} + \frac{R(Q_1)}{b^2} \right)}. \quad (59)$$

To proceed, we analyze the differential equation system in (43)–(46). Define the following variations on T, R, Z, V :

$$\bar{T} = \frac{T}{\lambda}, \quad \bar{R} = \frac{R - \lambda}{\lambda^2}, \quad (60)$$

$$\bar{V} = a^2 \lambda^2 \left(\frac{1}{ab} + V \right), \quad \bar{Z} = \lambda^2 \left(\frac{\lambda}{b^2} + Z \right). \quad (61)$$

Furthermore, let $\bar{S} = 1/a^2 + S$ (so that derivatives with respect to S are equivalent to derivatives with respect to $\bar{S}$). Then the system simplifies to

$$\bar{T}' = \frac{\bar{V}}{\bar{Z}}, \quad \bar{R}' = (\bar{T}')^2, \quad (62)$$

$$\bar{Z}' = - \left(\frac{\bar{S}}{-\bar{S}\bar{R} + \bar{T}^2} \right)^2, \quad \bar{V}' = \frac{\bar{T}\bar{Z}'}{\bar{S}}, \quad (63)$$

with boundary conditions

$$\bar{T}(0) = 0, \quad \bar{R}(0) = -\frac{1}{\lambda}, \quad (64)$$

$$\bar{Z}(Q_1) = \frac{\lambda^3}{b^2}, \quad \bar{V}(Q_1) = \frac{a\lambda^2}{b}. \quad (65)$$

Observe that

$$(\bar{S}\bar{V} - \bar{T}\bar{Z})' = \bar{V} + \bar{S}\bar{V}' - \bar{T}'\bar{Z} - \bar{T}\bar{Z}' \quad (66)$$

$$= \bar{V} + \bar{T}\bar{Z}' - \bar{V} - \bar{T}\bar{Z}' = 0. \quad (67)$$

Thus, for any solution, there exists a constant c_1 where⁴

$$\bar{S}\bar{V} - \bar{T}\bar{Z} = c_1^3. \quad (68)$$

The above analysis, resulting in (68), effectively reduces the differential equation system from 4 variables to 3. The following definitions will allow us to further reduce the system to 2 variables:

$$A = c_1^2 \frac{\bar{T}^2 - \bar{S}\bar{R}}{\bar{S}^2}, \quad B = c_1^{-4} \bar{S}\bar{Z}. \quad (69)$$

We have

$$A' = c_1^2 \left(\frac{2\bar{T}\bar{T}' - \bar{R} - \bar{S}\bar{R}'}{\bar{S}^2} - \frac{2(\bar{T}^2 - \bar{S}\bar{R})}{\bar{S}^3} \right) \quad (70)$$

$$= c_1^2 \frac{2\bar{S}\bar{T}\bar{T}' - \bar{S}\bar{R} - \bar{S}^2(\bar{T}')^2 - 2\bar{T}^2 + 2\bar{S}\bar{R}}{\bar{S}^3} \quad (71)$$

$$= c_1^2 \frac{-(\bar{T} - \bar{S}\bar{T}')^2 - \bar{T}^2 + \bar{S}\bar{R}}{\bar{S}^3} \quad (72)$$

$$= -\frac{c_1^2}{\bar{S}} \left(\frac{\bar{T}}{\bar{S}} - \frac{\bar{V}}{\bar{Z}} \right)^2 - \frac{A}{\bar{S}} \quad (73)$$

$$= -\frac{c_1^2}{\bar{S}} \left(\frac{\bar{T}\bar{Z} - \bar{S}\bar{V}}{\bar{S}\bar{Z}} \right)^2 - \frac{A}{\bar{S}} \quad (74)$$

$$= -\frac{c_1^8}{\bar{S}^3 \bar{Z}^2} - \frac{A}{\bar{S}} = -\frac{1}{\bar{S}B^2} - \frac{A}{\bar{S}}. \quad (75)$$

⁴We write c_1^3 instead of c_1 because later on it will make things simpler.

Also

$$B' = c_1^{-4} (\bar{S}\bar{Z}' + \bar{Z}) \quad (76)$$

$$= c_1^{-4} \left(-\bar{S} \left(\frac{\bar{S}}{-\bar{S}\bar{R} + \bar{T}^2} \right)^2 + \bar{Z} \right) \quad (77)$$

$$= -\frac{c_1^{-4} \bar{S}^3}{(-\bar{S}\bar{R} + \bar{T}^2)^2} + c_1^{-4} \bar{Z} = -\frac{1}{\bar{S}A^2} + \frac{B}{\bar{S}}. \quad (78)$$

That is, we have the self-contained two-dimensional system

$$A' = -\frac{1}{\bar{S}B^2} - \frac{A}{\bar{S}}, \quad B' = -\frac{1}{\bar{S}A^2} + \frac{B}{\bar{S}}. \quad (79)$$

To solve the two-dimensional system, we observe that

$$\left(AB + \frac{1}{A} - \frac{1}{B} \right)' = \left(B - \frac{1}{A^2} \right) A' + \left(A + \frac{1}{B^2} \right) B' \quad (80)$$

$$= \left(B - \frac{1}{A^2} \right) \left(-\frac{1}{\bar{S}B^2} - \frac{A}{\bar{S}} \right) + \left(A + \frac{1}{B^2} \right) \left(-\frac{1}{\bar{S}A^2} + \frac{B}{\bar{S}} \right) = 0. \quad (81)$$

Thus, there exists a constant c_2 where

$$AB + \frac{1}{A} - \frac{1}{B} = c_2. \quad (82)$$

Note that $B = f(A)$ solves (82), where f is defined in (10) with $\phi = c_2$. Assuming $B = f(A)$, we can reduce to a single-variable differential equation

$$A' = \frac{1}{1/a^2 + S} \left(-\frac{1}{f(A)^2} - A \right). \quad (83)$$

To complete the achievability proof in Thm. 1, we fix $A_f, B_f > 0$, and construct a solution to the two-dimensional system where $A(Q_1) = A_f, B(Q_1) = B_f$, with $c_1 = b\psi$ and $c_2 = \phi$, which allow us to satisfy the boundary constraints. The following lemma, proved in [9], provides the details, showing that these differential equation solutions exist, that they satisfy the assumptions of Lemma 4, and giving the result in the theorem.

Lemma 5: Given any $A_f, B_f > 0$ satisfying $A_f/B_f \leq a^2$, let (A_0, c_1) be the pair from Lemma 1. Then there exist $\lambda > 0$ and functions that solve the differential equation system (43)–(47), satisfying the conditions (57)–(58), where Q_1 is given in (14),

$$\frac{aT(Q_1)}{b\lambda} - \frac{R(Q_1)}{b^2\lambda} = Q_2 \quad (84)$$

where Q_2 is given in (15), and

$$1 + Z(0) + a^2 Q_1 - \frac{2aT(Q_1)}{b} + \frac{R(Q_1)}{b^2} \quad (85)$$

$$= \frac{A_0}{a^2} \left(\frac{1}{B_f} + A_0 B_0 - A_f B_f \right). \quad (86)$$

ACKNOWLEDGMENTS

The authors would like to thank Cynthia Keeler for her help solving the 4-dimensional system of differential equations. This work is supported in part by NSF grants CCF-1817241, CCF-1908725, CCF-1909451, CCF-2107526, and CCF-2245204.

REFERENCES

- [1] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.
- [2] P. Polyanskiy, H. V. P. Poor, and S. Verdú, "Minimum energy to send k bits through the Gaussian channel with and without feedback," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4880 – 4902, 2011.
- [3] A. Host-Madsen, "Minimum energy per bit in broadcast and interference channels with correlated information," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3796–3810, 2013.
- [4] A. El Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1545–1561, 2006.
- [5] Monty Python, "The Spanish inquisition," 1970. [Online]. Available: <https://www.youtube.com/watch?v=D5Df191WJ3o>
- [6] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, 1979.
- [7] C. Kim, Y. Sung, and Y. H. Lee, "A joint time-invariant filtering approach to the linear Gaussian relay problem," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4360–4375, 2012.
- [8] R. H. Gohary and H. Yanikomeroglu, "Joint optimization of the transmit covariance and relay precoder in general Gaussian amplify-and-forward relay channels," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5331–5351, 2013.
- [9] O. Kosut, M. Effros, and M. Langberg, "Nobody expects a differential equation: Minimum energy-per-bit for the Gaussian relay channel with rank-1 linear relaying," [Online] arXiv:2401.15495, 2024.
- [10] E. Süli and D. F. Mayers, *An Introduction to Numerical Analysis*. Cambridge University Press, 2003.