Modulation and Estimation with a Helper

Anatoly Khina, *Member, IEEE*, and Neri Merhav, *Life Fellow, IEEE*

Abstract

The problem of transmitting a parameter value over an additive white Gaussian noise (AWGN) channel is considered, where, in addition to the transmitter and the receiver, there is a helper that observes the noise non-causally and provides a description of limited rate R_{h} to the transmitter and/or the receiver. We derive upper and lower bounds on the optimal achievable α-th moment of the estimation error and show that they coincide for small values of α and for low SNR values. The upper bound relies on a recently proposed channel-coding scheme that effectively conveys R_{h} bits essentially error-free and the rest of the rate—over the same AWGN channel without help, with the error-free bits allocated to the most significant bits of the quantized parameter. We then concentrate on the setting with a total transmit energy constraint, for which we derive achievability results for both channel coding and parameter modulation for several scenarios: when the helper assists only the transmitter or only the receiver and knows the noise, and when the helper assists the transmitter and/or the receiver and knows both the noise and the message. In particular, for the message-informed helper that assists both the receiver and the transmitter, it is shown that the error probability in the channel-coding task decays doubly exponentially. Finally, we translate these results to those for continuous-time power-limited AWGN channels with unconstrained bandwidth. As a byproduct, we show that the capacity with a message-informed helper that is available only at the transmitter can exceed the capacity of the same scenario when the helper knows only the noise but not the message.

Index Terms

AWGN channel, helper, parameter estimation, channel coding, PPM.

A. Khina is with the Department of Electrical Engineering–Systems, Tel Aviv University, Tel Aviv 6997801, Israel (e-mail: anatolyk@tau.ac.il).

N. Merhav is with the Viterby Faculty of Electrical and Computer Engineering, Technion–Israel Institute of Technology, Haifa 3200004, Israel (e-mail: merhav@ee.technion.ac.il).

The work of A. Khina was supported by the ISRAEL SCIENCE FOUNDATION (grant 2077/20).

I. Introduction

Conveying a parameter over an additive white Gaussian noise (AWGN) channel is a fundamental problem in both information theory and estimation theory, which has received much attention over the years [1]–[30]. In particular, for a scalar parameter that is transmitted over n AWGN channel uses, the minimum achievable mean quadratic moment of the error D decays exponentially in n (see [11]–[13], [17]):

$$D \doteq \exp\left\{-n\,\frac{S}{6}\right\}, \tag{1}$$

where S denotes the signal-to-noise ratio (SNR), and '$\doteq$' denotes equality up to sub-exponential multiplicative terms. The decay exponent of $S/6$ in (1) is significantly smaller than the exponent dictated by the data-processing theorem (DPT) [31, Problem 10.8, Theorem 10.4.1], [32, Problem 3.18, Chapter 3.9] by a factor of 6, to wit

$$D \doteq \exp\left\{-n\,S\right\}.$$

This gap in performance stems from the constraint of processing and transmitting a single sample in (1) as opposed to the large source block lengths used in the achievability proof of the DPT.

To facilitate the communication between the transmitter and the receiver, a third party that has a correlated description of the source and/or the channel noise can act as a helper by providing its description over a rate-limited channel to the transmitter and/or the receiver. Wyner [33], and Ahlswede and Körner [34] studied the source coding problem in which the helper has a correlated description of the source and can provide it over a noiseless rate-limited channel to the receiver (see also [32, Chapter 10.4]). An extension of this scenario to the case of a transmitter and a helper that transmit to the receiver over a multiple-access channel was studied by Ahlswede and Han [35].

Channel coding over a state-dependent channel where the state is known to the helper, and the helper can provide a rate-limited description of the state to the transmitter was studied in [36]–[38] (see also [39, Chapter 3.5], [32, Chapter 7.8]). An interesting special case of the latter scenario, which has been studied more recently [40]–[44], is that of an additive noise channel with a helper that knows the noise sequence and can share a rate-limited description of the channel noise with the transmitter and/or the receiver; this scenario is motivated by network scenarios, such as the following. Consider a scenario in which both the transmitter and the helper act as transmitters, where the transmit signal of the helper interferes with the communication between the transmitter and the receiver. To mitigate this interference, the helper provides information about its transmit signal via a rate-limited channel to the transmitter and/or the receiver. Lapidoth and Marti [41] (see also [42], [40, Chapter 5.2]) proved that the capacity of an AWGN channel with a helper of rate (*help rate*) R_{h} equals the capacity without assistance $C_0 \triangleq \frac{1}{2}\log\left(1+S\right)$ plus the help rate R_{h}:

$$C = C_0 + R_{\mathrm{h}}, \tag{2}$$

both when the noise is known causally or non-causally to the helper.[1] Merhav [43] derived upper and lower bounds on the error exponent at all rates $R < C$ for the non-causal helper setting; in particular, he showed that R_{h} nats may be conveyed essentially error-free, whereas the remaining $R - R_{\mathrm{h}}$ nats are conveyed over an AWGN channel without help.

Despite the rich literature on source and channel coding with a helper, little has been done for the problem of parameter transmission over a noisy channel with a helper that knows the noise realization and can provide a rate-limited description to the transmitter and/or the receiver.

In this work, we study the problem of conveying a parameter over an AWGN channel with a helper that knows the noise non-causally and can provide a description over a rate-limited channel to the transmitter and/or receiver. The problem is formalized in Section II. We provide the necessary background about channel coding over AWGN channels with a helper that knows the noise sequence in Section III. In Section IV, we derive upper and lower bounds on the achievable α-th moment of the absolute error—which we term mean power-α error (MPαE) following [17]—and show that these bounds coincide for low α values as well as for low SNR values (for a fixed help rate).

To derive the lower (impossibility) bounds on the MPαE, we use the DPT and the extension of the technique of Ziv and Zakai [8], and Chazan, Zakai and Ziv [9] with exponentially many hypotheses (rather than 2) [45], [46]. For the upper (achievability) bound, we judiciously employ the aforementioned achievability result of [43] for transmitter-assisted channel coding over an AWGN channel: We apply uniform quantization to the parameter value where the quantization outputs are naturally labeled and allocate the most significant bits (MSBs) of rate R_{h} nats of the quantized description essentially error-free.

In Section V, we concentrate on the setting where the input is subject to a total energy constraint. To derive upper bounds on the MPαE, we refine, in Section V-A, the analysis of the

[1] All logarithms and exponents are to the base 2 in this work unless otherwise specified.

channel-coding scheme with a helper with a fixed help rate of Merhav [43] for this setting, as the original scheme and analysis do not carry over straightforwardly.

In the remainder of Section V, we consider the energy-constrained setting in which the helper is cognizant of both the noise and the message; we refer to such a helper as being *cribbed*. When the cribbed helper assists only the transmitter, we show in Section V-B that a channel-coding scheme that is based on pulse position modulation (PPM), and is reminiscent of the dirty-paper coding scheme of Liu and Viswanath [47], outperforms the scheme with a helper that knows only the noise but not the message for certain help rates. In Section V-C, we consider a helper that assists both the transmitter and the receiver. In this scenario, the helper can "simulate" several PPM schemes (without help) and then send the index of the best one to the transmitter and the receiver, to indicate which one of them to use. We show that the error probability of this scheme decays doubly exponentially with the total help budget.

In Section VI, we translate the results of Sections IV and V to the continuous-time transmission-assisted AWGN channel that is subject to a power constraint but has unconstrained bandwidth and compare the resulting bounds. In particular, we prove that a helper that is cognizant of both the noise and the message outperforms the capacity of the same channel with a helper that knows only the noise but not the message (2), for certain help rates.

We conclude the paper with a summary and discussion of future directions in Section VII.

II. Problem Setup

The sets of the natural and real numbers are denoted by $\mathbb{N}$ and $\mathbb{R}$, respectively. We denote indexed sequences between $i \in \mathbb{N}$ and $j \in \mathbb{N}$ by boldface letters: $\mathbf{A}_{i:j} = \left(A_i, A_{i+1}, \ldots, A_j\right)$, and $\mathbf{A} = \mathbf{A}_{1:n}$ for conciseness. We denote random variables by uppercase letters and realizations thereof by lowercase letters. $\mathscr{B}_n(r) \triangleq \{x \in \mathbb{R}^n | \, ||x||_2 \leq r\}$ is the n-dimensional ball of radius r and $||\cdot||_2$ denotes the Euclidean norm. $\mathrm{Vol}\{g\}$ denotes the volume of $g \subset \mathbb{R}^n$. The limit and the limit superior are denoted by $\lim$ and $\overline{\lim}$, respectively. We further denote by $\lim\limits_{x \downarrow a} f(x)$ the limit of a function $f(\cdot)$ as x decreases in value approaching a (limit from the right).

We make use of small o notation, viz., $f(x) = o\left(g(x)\right)$ means that $\lim_{x\to\infty} f(x)/g(x) = 0$. mod and $*$ denote the modulo and convolution operations, respectively.

We now formalize the modulation–estimation setting of this work; see also Figure 1. We concentrate here on the transmitter-assisted setting, namely, the setting in which the helper assists

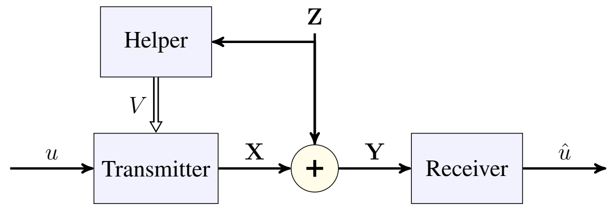


Fig. 1: Parameter estimation–modulation over an AWGN channel with an assisted transmitter.

only the transmitter. The setting of a helper that assists the receiver or both the transmitter and the receiver is defined similarly.

Parameter. The parameter to be conveyed is $u \in [-1/2, 1/2]$.

Transmitter. Maps the parameter u and the helper's description $V \in \{1, \ldots, M\}$ to a channel input sequence $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ that is subject to either a fixed (in n) power constraint P:

$$\sum_{t=1}^{n} \mathbb{E}\left[X_t^2\right] \le nP \qquad \forall u \in [-1/2, 1/2], \tag{3}$$

or to a fixed (in n) energy constraint E:

$$\sum_{t=1}^{n} \mathbb{E}\left[X_t^2\right] \le E \qquad \forall u \in [-1/2, 1/2]. \tag{4}$$

Channel. The input sequence $\mathbf{X}$ is transmitted over an AWGN channel:

$$Y_t = X_t + Z_t, \qquad t \in \{1, 2, \ldots, n\},$$

where $\mathbf{Y} \triangleq (Y_1, Y_2, \ldots, Y_n)$ is the output sequence, and $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ is the noise sequence whose entries are independent identically distributed (i.i.d.) Gaussian of zero mean and variance σ^2.

For the power-constrained setting, we define the signal-to-noise ratio (SNR) to be $S \triangleq P/\sigma^2$, and in the energy-constrained setting, we define $\gamma \triangleq E/\sigma^2$.

Helper. Knows (non-causally) the noise sequence $\mathbf{Z}$ and maps it into a finite-rate description $V \in \{1, 2, \ldots, M_\mathrm{h}\}$. V is revealed to the transmitter and/or receiver prior to the beginning of transmission. In the fixed-power setting, we will assume that $M_\mathrm{h} = \exp\{nR_\mathrm{h}\}$ where the help rate R_h is fixed, namely, M_h grows exponentially with n. In the energy-limited setting, we will consider both the fixed help rate case and the total nat budget $L_\mathrm{h} = \log M_\mathrm{h}$ which remains fixed, independently of n.

Receiver. Constructs an estimate $\hat{U}$ of the parameter u given the output $\mathbf{Y}$.

Objective. Achieving the minimal MPαE,

$$\epsilon(\alpha) \triangleq \sup_{u\in[-1/2,1/2)} \mathbb{E}\left[|\hat{U} - u|^{\alpha}\right], \tag{5}$$

where the expectation is with respect to the channel noise.

In this work, we will derive upper and lower bounds on the minimum achievable MPαE (5). In particular, for the power-limited scenario (3), we will concentrate on bounding the optimal achievable MPαE error exponent[2]

$$E(\alpha) \triangleq \overline{\lim_{n\to\infty}} -\frac{1}{n}\log\inf\epsilon(\alpha),$$

where the infimum is taken over all transmitter–receiver–helper triplets.

III. Background: Channel Coding over the AWGN Channel with a Helper

Consider the problem of reliable communication over an AWGN channel. The setting is as follows.

Message. A message W is uniformly distributed over $\{1, 2, \ldots, M\}$, with $M = \exp\{nR\}$.

Transmitter. Maps the message W and the helper's description $V \in \{1, \ldots, M_{\mathrm{h}}\}$ to a channel input sequence $\mathbf{X}$ of length n that is subject to the power constraint (3).

The channel and helper descriptions are the same as in Section II.

Receiver. Constructs an estimate $\hat{W}$ of W given the output $\mathbf{Y}$.

Objectives. The rate and the error probability are defined as

$$R \triangleq \frac{1}{n}\log M \qquad \text{and} \qquad P_e \triangleq \Pr\left(\hat{W} \neq W\right),$$

respectively. For a given rate R, we define the optimal channel-coding error exponent as

$$E(R) \triangleq \overline{\lim_{n\to\infty}} -\frac{1}{n}\log\inf P_e,$$

where the infimum is taken over all transmitter–receiver–helper triplets.

Denote by $C_0 = \frac{1}{2}\log(1+S)$ the capacity of an AWGN with SNR S without help. Then, the following results for the transmitter-assisted setting are known.

Theorem III.1 ([41]). *The capacity of an AWGN channel with SNR S and help rate R_{h} is*

$$C_{\mathrm{h}} = C_0 + R_{\mathrm{h}}.$$

[2]Both the energy in (4) and the error exponent are denoted by E. However, since the latter is a function, the distinction between them is clear.

Theorem III.2 ([43]). *The optimal achievable error exponent at rate R over an AWGN channel with SNR S and help rate R_{h} is bounded from below as*

$$E(R) \geq \begin{cases} \infty, & R < R_{\mathrm{h}}; \\ E_a\left(R - R_{\mathrm{h}}\right), & R_{\mathrm{h}} \leq R < C_0 + R_{\mathrm{h}}; \\ 0, & R \geq C_0 + R_{\mathrm{h}}, \end{cases}$$

where $E(R) = \infty$ for $R < R_{\mathrm{h}}$ means that an arbitrarily large exponent is achievable, and where $E_a(\cdot)$ is any achievable error-exponent function over the AWGN channel without help.

Moreover, for $R \in [R_{\mathrm{h}}, C_0 + R_{\mathrm{h}})$, $R' < R_{\mathrm{h}}$ nats per time step can be conveyed with an arbitrarily large error exponent and where $R_{\mathrm{h}} - R'$ is arbitrarily small for a sufficiently large n, whereas the remaining $R - R'$ nats per time step—with an error exponent that is arbitrarily close to $E_a\left(R - R_{\mathrm{h}}\right)$.

Theorem III.3 ([43]). *The optimal achievable error exponent at rate R over an AWGN channel with SNR S and help rate R_{h} is bounded from above by the weak sphere-packing bound*

$$E_{\mathrm{wsp}}(R) \triangleq \begin{cases} \infty, & R < R_{\mathrm{h}}; \\ \frac{1}{2}\left[\zeta(R) - \log\zeta(R) - 1\right], & R_{\mathrm{h}} < R < C_0 + R_{\mathrm{h}}; \\ 0, & R \geq C_0 + R_{\mathrm{h}}; \end{cases}$$

with

$$\zeta(R) \triangleq \frac{\exp\{2C_0\} - 1}{\exp\{2(R - R_{\mathrm{h}})\} - 1} = \frac{S}{\exp\{2(R - R_{\mathrm{h}})\} - 1}.$$

As the name suggests, the weak sphere-packing bound is weaker (larger upper bound) than the ordinary sphere-packing bound [5, Eq. (7.4.33)] (see also [48, Example 1], [49]), which is given by

$$\begin{aligned} E_{\mathrm{sp}}(R) \triangleq \frac{S}{4\beta}&\left[\beta + 1 - (\beta - 1)\sqrt{1 + \frac{4\beta}{S(\beta - 1)}}\right] \\ &+ \frac{1}{2}\log\left(\beta - \frac{S(\beta - 1)}{2}\left[\sqrt{1 + \frac{4\beta}{S(\beta - 1)}} - 1\right]\right) \end{aligned} \tag{6}$$

in the range $R_{\mathrm{h}} \leq R \leq R_{\mathrm{h}} + C_0$, where $\beta = \exp\{2(R - R_{\mathrm{h}})\}$. In particular, $\lim_{R \downarrow R_{\mathrm{h}}} E_{\mathrm{wsp}} = \infty$, whereas $\lim_{R \downarrow R_{\mathrm{h}}} E_{\mathrm{sp}} = S/2 < \infty$.

Indeed, whereas the ordinary sphere-packing bound relies on the independence between the input $\mathbf{X}$ and the noise $\mathbf{Z}$, in our setting of interest, the two may be correlated via the helper description V, which is a function of the noise $\mathbf{Z}$ and is provided to the transmitter to generate the input $\mathbf{X}$. This, in turn, suggests a larger upper bound; see [43] for further details and discussion.

We further note that the weak sphere-packing bound of Theorem III.3 shares important properties with the achievability bound of Theorem III.2: both are infinite for $R < R_{\mathrm{h}}$, strictly positive for $R < C_0 + R_{\mathrm{h}}$, and zero above the capacity, i.e., for $R > C_0 + R_{\mathrm{h}}$.

Remark III.1. While channel coding under an average error probability was considered in [43], discarding half of all the codewords that have the highest error probabilities achieves the same error exponent with respect to the maximal error probability; see [31, Chapter 7.7].

IV. Power-Limited Input

In this section, we derive lower and upper bounds on the achievable MPαE (5) in Section IV-A and Section IV-B, respectively, for the power-limited setting (3), and extend them for vector parameters in Section IV-D. We further compare these bounds in Section IV-B and show that they coincide when α is small, the SNR is small, or $R_{\mathrm{h}} \gg C_0$. While we focus in this section on the setting of a helper that assists the transmitter only, the results readily apply to the setting of a helper that assists the receiver only, *mutatis mutandis*.

A. MPαE Lower Bounds

We first introduce a bound that is based on the DPT. We then refine it using the bounding technique of Ziv and Zakai [8], and Chazan, Zakai, and Ziv [9], but with a number of hypotheses that grows exponentially with n (rather than 2) [45], [46].

The DPT yields the following bound, whose proof is available in Appendix A.

Theorem IV.1. *The α-th moment of the estimation error* (5) *is bounded from below as*

$$\epsilon(\alpha) \geq \frac{\alpha^{\alpha-1} \exp\left\{-\alpha(C_0 + R_{\mathrm{h}})n\right\}}{[2\Gamma(1/\alpha)]^{\alpha} \mathrm{e}}.$$

Remark IV.1. The proof of Theorem IV.1 in Appendix A holds even if the receiver has access also to (the same) V.[3] If, however, the transmitter and the receiver obtain two different descriptions, at rates R_{ht} and R_{hr}, respectively, then R_{h} should be replaced by $R_{\mathrm{hr}} + R_{\mathrm{ht}}$.

[3] As shown in Marti's M.Sc. thesis [40], the capacity of the AWGN channel with decoder help is also $C_0 + R_{\mathrm{h}}$. Therefore, the converse in [43] continues to apply provided that, using the notations of [43], $\mathcal{S}_m$ is redefined as $\{z^n : \psi(\phi(m) + z^n, T(z^n)) \neq m\}$ or $\{z^n : \psi(\phi(m, T(z^n)) + z^n, T(z^n)) \neq m\}$ rather than $\{z^n : \psi(\phi(m, T(z^n)) + z^n) \neq m\}$.

We now tighten the DPT-based lower bound using an extension of the Ziv–Zakai bound [8] and the Chazan–Zakai–Ziv bound [9] for M hypotheses (rather than 2), where M is exponentially large [45], [46]; the proof of this theorem is available in Appendix B.

Theorem IV.2. *The α-th moment of the estimation error* (5) *is bounded from below as*

$$\epsilon(\alpha) \geq \exp\left\{-\alpha\left(R_{\mathrm{h}} + C_0^*\right)n + o(n)\right\},$$

where

$$C_0^* \triangleq \frac{1}{2}\left(\log\left(1+\frac{S}{s^*}\right) + \frac{s^* - \log s^* - 1}{\alpha}\right), \tag{7}$$

$$s^* \triangleq \frac{2(\alpha+1)S}{\sqrt{(S+1)^2 + 4\alpha S} + S - 1}.$$

Remark IV.2 (Comparison of lower bounds). The bound in Theorem IV.2 is valid for any choice of $s \geq 1$ (see Appendix B) in lieu of s^*. In particular, the choice $s = 1$ yields the bound of Theorem IV.1, implying that the bound of Theorem IV.2 cannot be worse than that of Theorem IV.1.

Remark IV.3 (Lower bound for assisted receiver). To prove Theorem IV.2, we used the (impossibility) *weak sphere-packing bound* of Theorem III.3, which was derived for transmitter-assisted communication. As is explained Section III, this bound is weaker than the ordinary sphere-packing bound (6) but when the receiver obtains help in lieu of the transmitter, the channel input and the noise are independent, and hence the weak sphere-packing bound can be replaced by the ordinary sphere-packing bound (6).

B. MPαE Upper Bounds

We first derive an upper bound by constructing a separation-based scheme that quantizes the parameter and communicates the quantization index using a transmitter-assisted code. We then improve this scheme by proper labeling of the quantization index and utilizing the fact that R_{h} nats may be conveyed essentially error-free [43].

We start by presenting the first transmission scheme for a predefined $M \in \mathbb{N}$. To that end, we make use of a uniform scalar quantizer of the unit interval that takes one of M possible values, with the quantized value of $u \in [-1/2, 1/2)$ being

$$Q_M(u) \triangleq \frac{1}{M}\left(\lfloor Mu \rfloor + \frac{1}{2}\right). \tag{8}$$

Scheme IV.1.

Transmitter.

- Quantize u using a uniform quantizer, resulting in a quantized value $\hat{u}_w = Q_M(u)$ and a corresponding index $w \in \{1, \ldots, M\}$, where $Q_M(\cdot)$ was defined in (8).
- Encode w into a transmit sequence $\mathbf{X}$ using the channel-coding scheme of [43] with a helper V of rate R_{h}.
- Transmit $\mathbf{X}$.

Receiver.

- Receive $\mathbf{Y}$.
- Using $\mathbf{Y}$, decode $\hat{W} \in \{1, \ldots, M\}$—the reconstruction of w.
- Construct the estimate $\hat{U} = \hat{u}_{\hat{W}}$ of u.

This scheme gives rise to the following upper bound on the achievable MPαE.

Theorem IV.3. *The α-th moment of the estimation error* (5) *is bounded from above as*

$$\epsilon(\alpha) \leq \exp\left\{-n \max_{R \in [R_{\mathrm{h}}, C_0 + R_{\mathrm{h}}]} \min\left\{\alpha R, E_a\left(R - R_{\mathrm{h}}\right)\right\} + o(n)\right\}, \tag{9}$$

where $E_a(\cdot)$ denotes any achievable coding error exponent for coding over the same transmitter-assisted AWGN channel, and the maximization is attained for the solution $R \in [R_{\mathrm{h}}, C_0 + R_{\mathrm{h}}]$ of the equation

$$\alpha R = E_a\left(R - R_{\mathrm{h}}\right). \tag{10}$$

Proof: Let $M = \lceil \exp\{nR\} \rceil$ for $R > 0$. Then, the MPαE of Scheme IV.3 is bounded for any $u \in [-1/2, 1/2)$ as follows.

$$\mathbb{E}\left[\left|\hat{U} - u\right|^\alpha\right] = \mathbb{E}\left[\left|\hat{U} - u\right|^\alpha \middle| \hat{W} \neq w\right] \Pr\left(\hat{W} \neq w\right) + \mathbb{E}\left[\left|\hat{U} - u\right|^\alpha \middle| \hat{W} = w\right] \Pr\left(\hat{W} = w\right) \tag{11a}$$

$$\leq \Pr\left(\hat{W} \neq w\right) + \mathbb{E}\left[\left|\hat{U} - u\right|^\alpha \middle| \hat{W} = w\right] \tag{11b}$$

$$\leq \exp\left\{-n\left(E_a\left(R - R_{\mathrm{h}}\right) - \epsilon\right) + o(n)\right\} + \exp\left\{-nR\alpha\right\} \tag{11c}$$

$$= \exp\left\{-n \min\left\{E_a\left(R - R_{\mathrm{h}}\right) - \epsilon, \alpha R\right\} + o(n)\right\}, \tag{11d}$$

where (11a) follows from the law of total expectation; (11b) holds since the estimation error and the probability of correct decoding of w are bounded by 1; and (11c) holds for any $\epsilon > 0$, however small, for a sufficiently large n [43] for $R < C_0 + R_{\mathrm{h}}$, and since the estimation error given correct decoding of w is bounded by $1/M = \exp\{-nR\}$.

Optimization of (11) over R concludes the proof. ■

Further scrutiny of the result of [43] reveals that R_{h} nats can be conveyed essentially error-free. Hence, Scheme IV.1 can be improved as follows, where $M', M \in \mathbb{N}$ and M' is a divisor of M, namely, $\frac{M}{M'} \in \mathbb{N}$. Denote the corresponding rates by $R = \frac{1}{n}\log M$ and $R' = \frac{1}{n}\log M'$ respectively, where we use $R' = R_{\mathrm{h}} - \epsilon$ with $\epsilon > 0$ that can be chosen to be arbitrarily small.

Scheme IV.2.

Transmitter.

- Quantize u using a uniform quantizer (8), resulting in a quantized value $\hat{u}_w = Q_M(u)$, and a corresponding index $w \in \{1, \ldots, M\}$ that is assigned using natural labeling, namely, $\hat{u}_1 < \hat{u}_2 < \cdots < \hat{u}_M$. Decompose w as
$$w = \frac{M}{M'} \cdot (w_{\mathrm{m}} - 1) + w_\ell,$$
for $w_{\mathrm{m}} \in \left\{1, \ldots, M'\right\}$ and $w_\ell \in \left\{1, \ldots, \frac{M}{M'}\right\}$.
- Encode w into a transmit signal $\mathbf{X}$ using the channel-coding scheme of [43] with a helper V of rate R_{h}, such that the sub-message w_h is conveyed error-free (more precisely, with arbitrarily large error exponent) and the sub-message w_ℓ is conveyed with an error exponent of (close to) $E(R - R_{\mathrm{h}})$.

Receiver.

- Receive $\mathbf{Y}$.
- Using $\mathbf{Y}$, decode $\hat{W}_{\mathrm{m}}$ and $\hat{W}_\ell$, which are the reconstructions of w_{m} and w_ℓ, respectively.
- Construct $\hat{W} = \frac{M}{M'} \cdot (\hat{W}_{\mathrm{m}} - 1) + \hat{W}_\ell$.
- Construct the estimate $\hat{u} = \hat{u}_{\hat{W}}$ of u.

This scheme attains the following MPαE.

Theorem IV.4. *The α-th moment of the estimation error* (5) *is bounded from above as*
$$\begin{aligned}\epsilon(\alpha) &\leq \exp\left\{-n \max_{R\in[R_{\mathrm{h}}, C_0+R_{\mathrm{h}}]} \min\left\{\alpha R, E_a\left(R - R_{\mathrm{h}}\right) + \alpha R_{\mathrm{h}}\right\} + o(n)\right\} \\ &= \exp\left\{-n\left(\alpha R_{\mathrm{h}} + \max_{R_0\in[0,C_0]} \min\left\{\alpha R_0, E_a\left(R_0\right)\right\}\right) + o(n)\right\},\end{aligned} \tag{12}$$
where $E_a(\cdot)$ denotes any achievable channel-coding error exponent for coding over the same transmitter-assisted AWGN channel, and the maximization is attained for $R_0 \in [0, C_0]$ that satisfies
$$\alpha R_0 = E_a\left(R_0\right). \tag{13}$$

Proof: Since according to the communication scheme in [43], w_{m} is conveyed essentially without error, the MPαE in conveying u reduces to that of transmitting an effective parameter whose support is an interval of size $1/M' = \exp\left\{-nR'\right\} = \exp\left\{-n\left(R_{\mathrm{h}} - \epsilon\right)\right\}$ over an AWGN channel at rate $R - R_{\mathrm{h}} \in [0, C_0)$ without help. Therefore, the MPαE of the scheme is bounded for any $u \in [-1/2, 1/2)$ as follows.

$$\begin{aligned}
\mathbb{E}\left[\left|\hat{U} - u\right|^{\alpha}\right] &= \mathbb{E}\left[\left|\hat{U} - u\right|^{\alpha}\middle|\hat{W} = w\right]\Pr\left(\hat{W} = w\right) + \mathbb{E}\left[\left|\hat{U} - u\right|^{\alpha}\middle|\hat{W}_{\mathrm{m}} \neq w_{\mathrm{m}}\right]\Pr\left(\hat{W}_{\mathrm{m}} \neq w_{\mathrm{m}}\right) \\
&\quad + \mathbb{E}\left[\left|\hat{U} - u\right|^{\alpha}\middle|\hat{W}_{\ell} \neq w_{\ell}, \hat{W}_{\mathrm{m}} = w_{\mathrm{m}}\right]\Pr\left(\hat{W}_{\ell} \neq w_{\ell}, \hat{W}_{\mathrm{m}} = w_{\mathrm{m}}\right) && \text{(14a)} \\
&\leq \frac{1}{M^{\alpha}} + \Pr\left(\hat{W}_{\mathrm{m}} \neq w_{\mathrm{m}}\right) + \frac{1}{(M')^{\alpha}}\Pr\left(\hat{W}_{\ell} \neq w_{\ell}\right) && \text{(14b)} \\
&\leq \exp\left\{-nR\alpha\right\} + \exp\left\{-nE_{\infty}\right\} \\
&\quad + \exp\left\{-n\left(R_{\mathrm{h}} - \epsilon\right)\alpha\right\} \cdot \exp\left\{-n\left(E_a\left(R - R_{\mathrm{h}}\right) - \epsilon\right) + o(n)\right\} && \text{(14c)} \\
&= \exp\left\{-n\left[\min\left\{\alpha R, E_a\left(R - R_{\mathrm{h}}\right) + \alpha R_{\mathrm{h}}\right\} - \epsilon\right] + o(n)\right\}, && \text{(14d)}
\end{aligned}$$

where (14a) follows from the law of total expectation by noting that the events

$$\begin{aligned}
&\left\{\hat{W} = w\right\} \equiv \left\{\hat{W}_{\mathrm{m}} = w_{\mathrm{m}}, \hat{W}_{\ell} = w_{\ell}\right\}, \\
&\left\{\hat{W}_{\mathrm{m}} \neq w_{\mathrm{m}}\right\} \equiv \left\{\hat{W}_{\mathrm{m}} \neq w_{\mathrm{m}}, \hat{W}_{\ell} = w_{\ell}\right\} \cup \left\{\hat{W}_{\mathrm{m}} \neq w_{\mathrm{m}}, \hat{W}_{\ell} \neq w_{\ell}\right\}, \\
&\text{and } \left\{\hat{W}_{\mathrm{m}} = w_{\mathrm{m}}, \hat{W}_{\ell} \neq w_{\ell}\right\}
\end{aligned}$$

are disjoint and their union is the entire sample space; (14b) holds since the estimation error and the probability of correct decoding of w are bounded by 1 and since the estimation error given (correct) w_{m} is bounded by $1/M'$; and (14c) holds for any $\epsilon > 0$, however small, for a sufficiently large n [43] for $R < C_0 + R_{\mathrm{h}}$, with E_{∞} arbitrarily large for a sufficiently large n, since the estimation error given correct decoding of w is bounded by $\exp\left\{-nR\right\}$, and since the estimation error given correct decoding of w_{m} is bounded by $\exp\left\{-nR'\right\} = \exp\left\{-n\left(R_{\mathrm{h}} - \epsilon\right)\right\}$; (14d) holds since E_{∞} is arbitrarily large for a sufficiently large n.

Optimization of (14) over R concludes the proof. ■

C. Vector Parameter

Consider now the case where the parameter to be modulated is a vector $\boldsymbol{u} = (u_1, \ldots, u_d) \in [-1/2, 1/2)^d$; the results of the scalar-parameter case carry over to the vector-parameter setting with α replaced by α/d.

Corollary IV.1. *The α-th moment of the estimation error*

$$\begin{aligned}\epsilon_d(\alpha) &\triangleq \sup_{u \in [-1/2,1/2)^d} \mathbb{E}\left[\left\|\hat{U} - u\right\|_\alpha^\alpha\right] \\ &= \sum_{i=1}^{d} \mathbb{E}\left[\left|\hat{U}_i - u_i\right|^\alpha\right]\end{aligned}$$

is bounded from below as

$$\epsilon_d(\alpha) \geq \exp\left\{-\frac{\alpha}{d}\left(R_{\mathrm{h}} + C_0^*\right) n + o(n)\right\},$$

where C_0^* is defined in (7)*, and from above as*

$$\epsilon_d(\alpha) \leq \exp\left\{-n\left(\frac{\alpha}{d}R_{\mathrm{h}} + \max_{R_0 \in [0, C_0]} \min\left\{\frac{\alpha}{d}R_0, E\left(R_0\right)\right\}\right) + o(n)\right\}$$

with the minimum achieved by the solution of

$$\frac{\alpha}{d}R_0 = E\left(R_0\right).$$

The proof of Corollary IV.1 is available in Appendix C.

D. Comparison of Upper and Lower Bounds

For concreteness, consider first the case of a very noisy channel, namely, $S \ll 1$. For this channel [5], [14]:

$$E(R) \cong \begin{cases} \frac{C_0}{2} - R, & R < \frac{C_0}{4}; \\ (\sqrt{C_0} - \sqrt{R})^2, & \frac{C_0}{4} \leq R \leq C_0; \\ 0, & R \geq C_0. \end{cases}$$

The solution of (10) in Theorem IV.3, in this case, is

$$R^* = \begin{cases} \frac{\frac{C_0}{2} + R_{\mathrm{h}}}{1+\alpha}, & \alpha \geq \frac{C_0}{C_0 + 4R_{\mathrm{h}}}; \\ R_{\mathrm{h}} + \frac{\left[\sqrt{C_0} - \sqrt{C_0 - (1-\alpha)(C_0 - \alpha R_{\mathrm{h}})}\right]^2}{(1-\alpha)^2}, & \alpha \leq \frac{C_0}{C_0 + 4R_{\mathrm{h}}}; \end{cases}$$

and the corresponding upper bound on the MPαE (9) is

$$\epsilon(\alpha) \leq \exp\left\{-n\alpha R^* + o(n)\right\}.$$

Similarly, the solution of (13) of Theorem IV.4 for a very noisy channel is

$$R_0^* = \begin{cases} \frac{C_0}{2(1+\alpha)}, & \alpha \geq 1; \\ \frac{C_0}{(1+\sqrt{\alpha})^2}, & \alpha < 1; \end{cases} \tag{15}$$

and the corresponding upper bound on the MPαE (12) is

$$\epsilon(\alpha) \leq \exp\left\{-n\alpha(R_0^* + R_{\mathrm{h}}) + o(n)\right\}. \tag{16}$$

Relaxing the assumption of the very noisy channel, the more general formula for R_0^* of Theorem IV.4 is given by [50, Appendix B]

$$R_0^* = \sup_{0 \leq \rho \leq 1} \frac{E_0(\rho)}{\rho + \alpha},$$

where $E_0(\cdot)$ is the Gallager function of the Gaussian channel [5, Eq. (7.4.21)]:

$$E_0(\rho) = \sup_{0 \leq s < \frac{1}{2S}} \left[s(1+\rho)S + \frac{1}{2}\log(1 - 2sS) + \frac{\rho}{2}\log\left(1 - 2sS + \frac{S}{1+\rho}\right)\right].$$

This can be further improved by taking the maximum between R_0^* and the solution, R_0^{**}, to the equation $\alpha R = E_{\mathrm{ex}}(R)$, where $E_{\mathrm{ex}}(\cdot)$ is the expurgated exponent which, for the Gaussian channel considered here, is given by [5, Eq. (7.5.52)]:

$$E_{\mathrm{ex}}(R) = \frac{S}{4}\left(1 - \sqrt{1 - \exp\{-2R\}}\right).$$

This yields the following upper bound on the MPαE.

$$\epsilon(\alpha) \leq \exp\left\{-n\alpha\left(R_{\mathrm{h}} + \max\left\{R_0^*, R_0^{**}\right\}\right) + o(n)\right\}. \tag{17}$$

Finally, consider the limit of a very small $\alpha > 0$. The solution R^* of (10) in Theorem IV.3 in this limit is approximately,

$$R^* = C_0 + R_{\mathrm{h}},$$

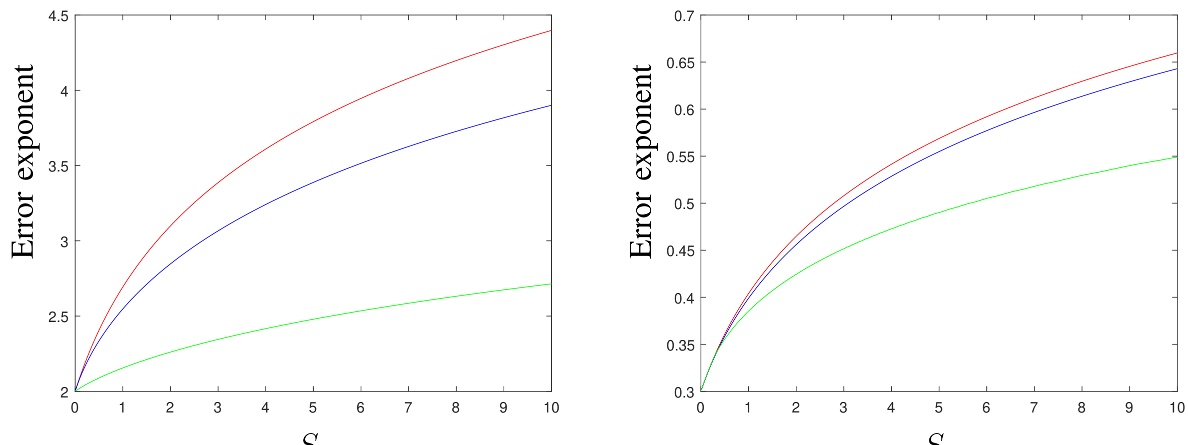
regardless of the assumption of the very noisy channel, and hence

$$\epsilon(S) \leq \exp\left\{-n\alpha R^* + o(n)\right\} = \exp\left\{-n\alpha(C_0 + R_{\mathrm{h}}) + o(n)\right\},$$

which meets the exponent of the DPT lower bound of Theorem IV.1.

Similarly, by comparing the mathematical expressions of the upper and lower bounds, the gap between them is small (and decays to zero in the limit) if at least one of the following conditions is satisfied: (1) α is small, (2) $R_{\mathrm{h}} \gg C_0$, (3) d is large.

In Figures 2 and 3, we display curves of the exponential error impossibility and achievability bounds as functions of S and α, respectively. We observe that the DPT-based bound of Theorem IV.1 is weaker than the Ziv–Zakai-based bound of Theorem IV.2, in agreement with Remark IV.2. These figures further demonstrate that while there is a gap between the lower and upper bounds, it diminishes in the limit of small α and in the limit of small S, as expected from the discussion above.



(a) $\alpha = 2$ (b) $\alpha = 0.3$

Fig. 2: Graphs of the converse and achievability bounds as functions of S, for $R_{\mathrm{h}} = 1$ and $\alpha = 0.3, 2$: The DPT impossibility bound, the Ziv–Zakai based impossibility bound, and the achievability bound (17), are depicted in red, blue, and green, respectively.

V. Energy-Limited Input

In this section, we consider several helper scenarios, where the transmitter is subject to a fixed energy constraint (4).

To that end, we start by reviewing a known PPM-based scheme and its analysis [12], [13], [27] in the absence of a helper.

Scheme V.1 (PPM-based)**.**

Transmitter.

1) Quantize u using a uniform quantizer, resulting in a quantized value $\hat{u}_w = Q_M(u)$, and a corresponding index $w \in \{1, \ldots, n\}$.
2) Send

$$x_t = \sqrt{E} \cdot \delta_{t,w}, \qquad\qquad t \in \{1, \ldots, n\},$$

where $n = M$, and δ is Kronecker's delta function, viz., it equals 1 if $t = w$, and 0 otherwise.

Receiver.

1) Receive $\mathbf{Y}$.
2) Decode $\hat{W} = \underset{t \in \{1,\ldots,n\}}{\arg\max}\, Y_t$.
3) Construct the estimate $\hat{U} = \hat{u}_{\hat{W}}$ of u.

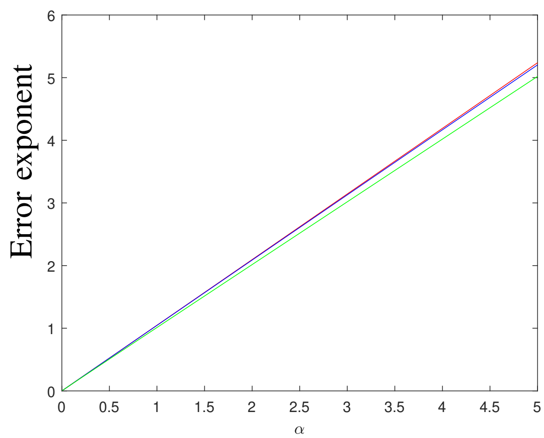


Fig. 3: Graphs of the converse and achievability bounds as functions of α, for $R_{\rm h} = 1$ and $S = 0.1$: The DPT impossibility bound, the Ziv–Zakai based impossibility bound, and the achievability bound (17), are depicted in red, blue and green, respectively.

Since $M = n$ in this scheme, the total budget L of this scheme is $L \triangleq \log M = \log n$.

We first present an upper bound on the error probability of the channel-coding scheme that is encapsulated in Scheme V.1, viz., with step 1 of the transmitter replaced with simply receiving a message $W \in \{1, 2, \ldots, M\}$, and with the receiver stopping at step 2. This scheme achieves the following performance [6, Ch. 8], [5, Ch. 8], [14, Ch. 2.5].

Theorem V.1. *The error probability in decoding $w \in (1, \ldots, \exp\{L\})$ in Scheme V.1, with a proper choice of parameters, is bounded from above as*[4]

$$\Pr\left(\hat{W} \neq w\right) \leq \begin{cases} \exp\left\{-\left(\frac{\gamma}{4} - L\right)\right\}, & L \leq \frac{\gamma}{8}; \\ \exp\left\{-\left(\sqrt{\frac{\gamma}{2}} - \sqrt{L}\right)^2\right\}, & \frac{\gamma}{8} < L \leq \frac{\gamma}{2}. \end{cases}$$

By applying a similar analysis to that of the proofs of Theorems IV.3 and IV.4 with the error probability bound of Theorem V.1, replacing nR with L, setting $nR_{\rm h} = 0$ (no helper), and optimizing over L, we arrive at the following result that was formerly derived in [12], [27] (see also the proof of Corollary V.3 in the sequel with $R_{\rm h}$ set to 0).

[4] Sub-exponential improvements are available in [5, Ch. 8] and [27].

Corollary V.1. *Scheme V.1 achieves an MPαE that is bounded from above as*

$$\mathbb{E}\left[\left|u - \hat{U}\right|^{\alpha}\right] \le \exp\left\{o(\gamma)\right\} \cdot \begin{cases} \exp\left\{-\frac{\alpha}{(1+\sqrt{\alpha})^2}\frac{\gamma}{2}\right\}, & \alpha < 1, \\ \exp\left\{-\frac{\alpha}{1+\alpha}\frac{\gamma}{4}\right\}, & \alpha \ge 1. \end{cases}$$

This bound is in fact tight up to sub-exponential terms in γ [13], [28].

In the remainder of this section, we address the following three scenarios:

1) In Section V-A, we study the setting of help rate R_{h} *per time step*. Namely, the helper conveys nR_{h} nats to the transmitter.
2) In Section V-B, we consider a cribbed helper, namely a helper that knows both the parameter u and the noise sequence $\mathbf{Z}$ and can convey a total budget of L_{h} nats to the transmitter (only), before transmission.
3) In Section V-C, we consider a cribbed helper that reveals the same helper message of total budget L_{h} to both the transmitter and the receiver.

A. Helper with a Fixed Rate

First, consider scenario 1, namely, that of R_{h} nats per time step that the helper can reveal to the transmitter prior to transmission.

The following lower bound on $\epsilon(\alpha)$ is easily obtained by replacing P by E/n in Theorem IV.1.

Corollary V.2. *The α-th moment of the estimation error* (5) *is bounded from below as*

$$\epsilon(\alpha) \ge \frac{\alpha^{\alpha-1}}{\left[2\Gamma\left\{\frac{1}{\alpha}\right\}\right]^{\alpha}\mathrm{e}} \cdot \exp\left\{-\frac{\alpha\gamma}{2}\right\} \cdot \exp\left\{-\alpha R_{\mathrm{h}} n\right\}.$$

Remark V.1. Replacing P by E/n in Theorem IV.2 and evaluating nC_0^* in the limit of $n \to \infty$ yields a similar lower bound to that of Corollary V.2 up to exponentially negligible terms.

Somewhat surprisingly, naïve replacement of P by E/n in the achievability analysis of Theorem IV.3 and Theorem IV.4 fails. Specifically, the proof of Theorem III.2 requires more intricate analysis since, in the original proof, the transmission power during the first sub-block in the scheme of [43, Section IV-A] needs to be larger than that of the noise. Yet, the result of Theorem III.2 still holds, as stated in the following theorem; the proof of this result is available in Appendix D for both receiver-only and transmitter-only helpers.

Theorem V.2. *The error probability in decoding w in Scheme V.1, with a proper choice of parameters, under a total energy constraint* (4) *is bounded from above as*

$$\Pr\left(\hat{W} \neq w\right) \leq \exp\{o(n)\} \cdot \begin{cases} \exp\{-nE_\infty\}, & L < nR_\mathrm{h}; \\ \exp\left\{-\left(\frac{\gamma}{4} - L + nR_\mathrm{h}\right)\right\}, & 0 \leq L - nR_\mathrm{h} < \frac{\gamma}{8}; \\ \exp\left\{-\left(\sqrt{\frac{\gamma}{2}} - \sqrt{L - nR_\mathrm{h}}\right)^2\right\}, & \frac{\gamma}{8} \leq L - nR_\mathrm{h} < \frac{\gamma}{2}; \end{cases} \tag{18}$$

where E_∞ is arbitrarily large for a sufficiently large n.

Moreover, for $L \in [nR_\mathrm{h}, \gamma/2 + nR_\mathrm{h})$, $L' < nR_\mathrm{h}$ nats can be conveyed with an arbitrarily large error exponent and where $nR_\mathrm{h} - L'$ is arbitrarily small for a sufficiently large n, whereas the remaining $nR - L'$ nats—with an error probability of (18).

Using Scheme IV.2 with $L = \log M$ (in lieu of $nR = \log M$) and following the proof of IV.4 with the error probability bound of Theorem V.1, yields the following upper bound on the MPαE of this scheme, which is formally proved in Appendix D.

Corollary V.3. *The α-th moment of the estimation error* (5) *under a total energy constraint* (4) *is bounded from above as*

$$\epsilon(\alpha) \leq \exp\{-n\alpha R_\mathrm{h} + o(n)\} \cdot \begin{cases} \exp\left\{-\alpha \frac{\gamma}{2(1+\sqrt{\alpha})^2}\right\}, & \alpha \leq 1; \\ \exp\left\{-\alpha \frac{\gamma}{4(1+\alpha)}\right\}, & \alpha > 1. \end{cases}$$

We note that the achievability result of Corollary V.3 asymptotically matches the exponential decay rate of $\exp\{-\alpha R_\mathrm{h} n\}$ of Corollary V.2. We further note that the proof of Theorem V.2 relies on the fact that a help rate of R_h nats *per time step* is available. It would be interesting to extend the result of Theorem V.2 (and consequently also of Corollary V.3) to work for a (possibly large) fixed total budget of L_h nats in lieu of nR_h.

B. Transmitter with a Cribbed Helper

Consider a helper that has access to both the parameter u and the entire noise sequence $\mathbf{Z}$, i.e., V is a function of both u and $\mathbf{Z}$. V of total budget $L_\mathrm{h} = \log M_\mathrm{h}$ is conveyed only to the transmitter prior to the beginning of the transmission.

We next construct a scheme that quantizes the parameter u and uses a variant of Scheme V.1 (the PPM-based scheme) to convey the quantized value. This variant of the scheme is reminiscent of the scheme of Liu and Viswanath [47] for writing on dirty paper: For each possible message

$w \in \{1, \ldots, M\}$, it allocates M_{h} time slots, and the helper provides to the transmitter the index of the slot with the maximal noise so that the transmitter can superimpose its entire transmit energy over that noise. This scheme is described as follows.

Scheme V.2 (Transmitter-assisted with cribbing)**.**

Helper.

i. Quantize u using a uniform quantizer, resulting in a quantized value $\hat{u}_w = Q_M(u)$, and a corresponding index $w \in \{1, \ldots, M\}$.

ii. Determine the most favorable transmit index $V \in \{1, \ldots, M_{\mathrm{h}}\}$:

$$V = \operatorname*{arg\,max}_{i \in \{1,\ldots,M_{\mathrm{h}}\}} Z_{w+(i-1)M}. \tag{19}$$

iii. Convey V to the transmitter.

Transmitter.

i. Receive V from the helper.

ii. Apply step i of the helper.

iii. Send

$$X_t = \sqrt{E} \cdot \delta_{t, w+(V-1)M}, \qquad\qquad t \in \{1, \ldots, n\},$$

where $n = M \cdot M_{\mathrm{h}}$.

Receiver.

i. Observe $\mathbf{Y}$.

ii. Decode $\breve{W} \triangleq \operatorname*{arg\,max}_{t \in \{1,\ldots,n\}} Y_t$, and $\hat{W} = \left(\breve{W} - 1 \mod M\right) + 1$.

iii. Construct the estimate $\hat{U} = \hat{u}_{\hat{W}}$ of u.

We first present an upper bound on the error probability of the channel-coding scheme that is encapsulated in Scheme V.2, viz., with step i of the helper replaced with simply receiving a message $w \in \{1, 2, \ldots, M\}$, and with the receiver stopping at step ii; the proof of this bound is available in Appendix E.

Theorem V.3. *The error probability in decoding w in Scheme V.2 is bounded from above as*

$$\Pr\left(\hat{W} \neq w\right) \leq \exp\left\{-\left(\frac{\gamma}{2} + \sqrt{2L_{\mathrm{h}}\gamma} - L\right) + o(1)\right\}$$

for $L < \frac{\gamma}{2} + \sqrt{2L_{\mathrm{h}}\gamma}$, where $o(1)$ is with respect to L_{h}.

Applying a similar analysis to that in the proof of Theorem IV.3 yields the following upper bound on the MPαE of this scheme; this result is formally proved in Appendix E.

Corollary V.4. *The* α*-th moment of the estimation error* (5) *with a cribbed helper is bounded from above as*

$$\epsilon(\alpha) \leq \exp\left\{-\frac{\alpha}{1+\alpha}\left(\frac{\gamma}{2}+\sqrt{2L_{\mathrm{h}}\gamma}\right)+o(1)\right\},$$

where $o(1)$ *is with respect to* L_{h}.

C. Two-Sided Cribbed Helper

Consider now a helper that has access to both the parameter u and the entire noise sequence $\mathbf{Z}$, as in Section V-B, i.e., V is a function of both U and $\mathbf{Z}$. But now, V of total nat budget L_{h} is conveyed to both the transmitter and the receiver prior to the beginning of the transmission. Consequently, the helper can now use its available nat budget L_{h} as a side channel (to simply convey noiselessly L_{h} nats of the message), assist directly the transmission over the noisy channel, or both. We next present these alternatives. Interestingly, when allocating the helper to assist the transmission over the noisy channel, a *double exponential* decay of the error probability in L_{h} is attained.

1) Side channel: Since the helper knows the message and conveys it to the receiver (and the transmitter), it may apply a similar scheme to Scheme IV.2, but convey w_{m}—the sub-message that corresponds to the $L_{\mathrm{h}} \log_2 \mathrm{e}$ MSBs of the quantization of u—using V without any noise. Since these $L_{\mathrm{h}} \log_2 \mathrm{e}$ MSBs are conveyed to the receiver without error, the support of the effective parameter is limited to $[-\exp\{-L_{\mathrm{h}}\}/2, \exp\{-L_{\mathrm{h}}\}/2)$. Applying Scheme V.1 to this effective parameter yields the following result, which is an immediate consequence of Corollary V.1.

Theorem V.4. *The* α*-th moment of the estimation error* (5) *is bounded from above as*

$$\epsilon(\alpha) \leq \exp\{-\alpha L_{\mathrm{h}} + o(\gamma)\} \cdot \begin{cases} \exp\left\{-\frac{\alpha}{(1+\sqrt{\alpha})^2}\frac{\gamma}{2}\right\}, & \alpha < 1; \\ \exp\left\{-\frac{\alpha}{1+\alpha}\frac{\gamma}{4}\right\}, & \alpha \geq 1. \end{cases}$$

This scheme attains the same performance as in Corollary V.3 with a simple scheme (thanks to cribbing) for any L_{h}, and not only for $L_{\mathrm{h}} = nR_{\mathrm{h}}$. Note that this scheme does not require knowing V at the transmitter and hence applies also to the setting of a receiver-only helper with cribbing.

2) Assisted noisy-channel coding: Since the helper knows both u and $\mathbf{Z}$ prior to transmission, it knows exactly what the transmitter and the receiver will send and receive. Therefore, it can simulate M_{h} instances of Scheme V.1 (the PPM-based scheme without assistance) one after the other in time, choose the most favorable one in terms of performance (an instance for which w is decoded correctly, if possible), and convey via V at what time interval to transmit (which of the M_{h} schemes to use) to both the transmitter and the receiver. Since the noise sequences in all of these M_{h} time intervals are independent, so are their respective error probabilities. This scheme and bounds on its performance are detailed next.

Scheme V.3 (Two-sided cribbed helper)**.**

Helper.

i. Quantize u using a uniform quantizer, resulting in a quantized value $\hat{u}_w = Q_M(u)$, and a corresponding index $w \in \{1, \ldots, M\}$
ii. Determine the most favorable transmit index $V \in \{1, \ldots, M_{\mathrm{h}}\}$ (an index of the instance of Scheme V.1 for which there will be no detection error of w, if possible):

$$V = \underset{i \in \{1,\ldots,M_{\mathrm{h}}\}}{\arg\max} \left(Z_{w+(i-1)M} - \max_{t \in \{1,\ldots,M\}, t \neq w} Z_{t+(i-1)M} \right). \tag{20}$$

iii. Convey V to the transmitter and the receiver.

Transmitter.

i. Receive V from the helper.
ii. Apply step i of the helper.
iii. Send

$$X_t = \sqrt{E} \cdot \delta_{t, w+(V-1)M}, \qquad\qquad t \in \{1, \ldots, n\},$$

where $n = M \cdot M_{\mathrm{h}}$.

Receiver.

i. Observe $\mathbf{Y}$ and receive V from the helper.
ii. Decode $\hat{W} = \underset{t \in \{1,\ldots,M\}}{\arg\max}\, Y_{t+(V-1)M}$.
iii. Construct the estimate $\hat{u} = \hat{u}_{\hat{W}}$ of u.

We first present a simple result that states that the error probability of the information transmission scheme that is encapsulated in Scheme V.3 decays *doubly exponentially* with L_{h}, namely, the error probability of the scheme with step i of the helper replaced with simply receiving a message $W \in \{1, 2, \ldots, M\}$, and with the receiver stopping at step ii. This improvement stems from the fact that a decoding error of w happens if and only if all the $M_\mathrm{h} = \exp\{L_\mathrm{h}\}$ independent instances of Scheme V.1 (the basic scheme without a helper) fail to correctly decode w. Since the instances are independent, the probability of this decoded error is raised to a power of $M_\mathrm{h} = \exp\{L_\mathrm{h}\}$ and hence the double exponential decay rate in L_h.

Proposition V.1. *Denote by $P_e(L)$ the error probability of Scheme V.1 (PPM without help) for some total nat budget L. Then, the overall error in decoding w, with the same total nat budget L and total help budget L_h, equals*

$$\Pr\left(\hat{W} \neq w\right) = \left[P_e(L)\right]^{\exp\{L_\mathrm{h}\}}.$$

In particular, $\Pr\left(\hat{W} \neq w\right)$ *decays doubly exponentially with L_h for any $L > 0$.*

Proof: Denote the channel error event if the transmitter and receiver (and helper) used instance $i \in \{1, \ldots, M_\mathrm{h}\}$ by $\mathscr{E}_i$. Since the noise sequences corresponding to the different scheme instance $i \in \{1, \ldots, M_\mathrm{h}\}$ are i.i.d.,

$$\Pr\left(\hat{W} \neq w\right) = \Pr\left(\bigcap_{i=1}^{M_\mathrm{h}} \mathscr{E}_i\right) = \left[P_e(L)\right]^{\exp\{L_\mathrm{h}\}}.$$

Since $P_e < 1$ for *any* $L > 0$, $\Pr\left(\hat{W} \neq w\right)$ decays doubly exponentially with L_h for any $L > 0$. ■

Remark V.2. The same result holds for any other scheme without help in lieu of Scheme V.1.

A straigtforward application of Proposition V.1 with the result of Corollary V.1 yields

$$\Pr\left(\hat{W} \neq w\right) \leq \begin{cases} \exp\left\{-\left(\frac{\gamma}{4} - L\right)\exp\{L_\mathrm{h}\}\right\}, & L < \frac{\gamma}{8}; \\ \exp\left\{-\left(\sqrt{\frac{\gamma}{2}} - \sqrt{L}\right)^2 \exp\{L_\mathrm{h}\}\right\}, & \frac{\gamma}{8} \leq L < \frac{\gamma}{2}. \end{cases} \tag{21}$$

Using this simple bound on the error probability (21) and optimizing over L yields the following loose upper bound on the achievable MPαE $\epsilon(\alpha)$ for a two-sided cribbed helper and transmission under a total power constraint (4), the proof of which is available in Appendix F.

Corollary V.5. *The α-th moment of the estimation error* (5) *with a cribbed helper that assists both the transmitter and the receiver is bounded from above as*

$$\epsilon(\alpha) \leq \begin{cases} \exp\left\{-\frac{\alpha\exp\{L_\mathrm{h}\}}{\exp\{L_\mathrm{h}\}+\alpha}\frac{\gamma}{4} + +o(1)\right\}, & L_\mathrm{h} \leq \log\alpha; \\ \exp\left\{-\frac{\alpha\exp\{L_\mathrm{h}\}}{\left(\exp\{L_\mathrm{h}/2\}+\sqrt{\alpha}\right)^2}\frac{\gamma}{2} + o(1)\right\}, & L_\mathrm{h} > \log\alpha; \end{cases} \tag{22}$$

where $o(1)$ is with respect to γ.

Remark V.3. The resulting bound of (22) is loose when L_{h} is large. In particular, in the limit of $L_{\mathrm{h}} \to \infty$, the bound of (22) saturates at $2\exp\{-\alpha\gamma/2\}$, instead of converging to zero. This is also suggested by the result of Theorem V.3 by noting that the decision rule (20) majorizes (19). The reason for (22) being loose for large values of L_{h} is that the upper bound on the error probability in decoding w that we have used (21) equals 1 at $L = \gamma/2$, whereas the error probability for any L is in fact strictly lower than 1. See Section VII for a further discussion.

3) Hybrid: By allocating a portion L_{m} of the total help budget L_{h} for side-channel and the remainder $L_{\mathrm{h}} - L_{\mathrm{m}}$ for assistance for noisy-channel coding, a more general achievable that subsumes (22) and the result of Theorem V.4 may be constructed as follows. Let $M_{\mathrm{m}}, M_\ell, M_{\mathrm{h}} \in \mathbb{N}$ where $M = M_{\mathrm{m}} \cdot M_\ell$ and such that $\frac{M_{\mathrm{h}}}{M_{\mathrm{m}}} \in \mathbb{N}$. Equivalently, $L_{\mathrm{m}} = \log M_{\mathrm{m}}$, $L_\ell = \log M_\ell$, $L = \log M$, and $L_{\mathrm{m}} + L_\ell = L$.

Scheme V.4 (two-sided cribbed helper [hybrid])**.**

Helper.

i. Quantize u using a uniform quantizer (8), resulting in a quantized value $\hat{u}_w = Q_M(u)$, and a corresponding index $w \in \{1, \ldots, M\}$ that is assigned using natural labeling, namely, $\hat{u}_1 < \hat{u}_2 < \cdots < \hat{u}_M$. Decompose w as

$$w = M_\ell \cdot (w_{\mathrm{m}} - 1) + w_\ell,$$

for $w_{\mathrm{m}} \in \{1, \ldots, M_{\mathrm{m}}\}$ and $w_\ell \in \{1, \ldots, M_\ell\}$.

ii. Determine the most favorable transmit index $V_\ell \in \{1, \ldots, M_{\mathrm{h}}/M_{\mathrm{m}}\}$ (an index of the instance of Scheme V.1 for which there will be no detection error of w_ℓ, if possible):

$$V_\ell = \operatorname*{arg\,max}_{i \in \{1,\ldots,M/M_{\mathrm{m}}\}} \left(Z_{w_\ell + (i-1)M_\ell} - \max_{t \in \{1,\ldots,M_\ell\}, t \neq w_\ell} Z_{t+(i-1)M_\ell} \right).$$

iii. Convey $V = M_{\mathrm{m}} \cdot (V_\ell - 1) + w_m$ to the transmitter and the receiver.

Transmitter.

i. Receive V from the helper.

ii. Construct $V_\ell = \lfloor V/M_{\mathrm{m}} \rfloor + 1$.

iii. Apply step i of the helper.

iv. Send

$$X_t = \sqrt{E} \cdot \delta_{t, w_\ell + (V_\ell - 1)M_\ell}, \qquad t \in \{1, \ldots, n\},$$

where $n = M_\ell \cdot M_{\mathrm{h}} / M_{\mathrm{m}}$.

Receiver.

i. Observe $\mathbf{Y}$ and receive V from the helper.
ii. Construct $V_\ell = \lfloor V / M_{\mathrm{m}} \rfloor + 1$ and $w_m = (V - 1 \mod M_{\mathrm{m}}) + 1$.
iii. Decode $\hat{W}_\ell = \underset{t \in \{1,\ldots,M_\ell\}}{\arg\max}\, Y_{t+(V_\ell - 1)M_\ell}$.
iv. Construct $\hat{W} = M_\ell (w_{\mathrm{m}} - 1) + \hat{W}_\ell$
v. Construct the estimate $\hat{u} = \hat{u}_{\hat{W}}$ of u.

Scheme V.4 transmits $L_{\mathrm{m}} \log 2$ MSBs essentially error-free (more precisely, with an error probability that decays exponentially fast with L_{h} with an arbitrarily large decay rate), resulting in an effective parameter support of $[-\exp\{-L_{\mathrm{m}}\}/2, \exp\{-L_{\mathrm{m}}\}/2)$. Applying the analysis the result of Corollary V.5 to this effective parameter yields the following upper bound on MPαE.

Corollary V.6. *The α-th moment of the estimation error* (5) *with a cribbed helper that assists both the transmitter and the receiver is bounded from above as*

$$\mathbb{E}\left[\left|u - \hat{U}\right|^\alpha\right] \leq \min_{L_{\mathrm{m}} \in [0, L_{\mathrm{h}}]} \exp\{-\alpha L_{\mathrm{m}}\} \begin{cases} \exp\left\{-\frac{\alpha}{4}\frac{\exp\{L_{\mathrm{h}}-L_{\mathrm{m}}\}}{\exp\{L_{\mathrm{h}}-L_{\mathrm{m}}\}+\alpha}\gamma + o(1)\right\}, & L_{\mathrm{h}} - L_{\mathrm{m}} \leq \log\alpha; \\ \exp\left\{-\frac{\alpha}{2}\frac{\exp\{L_{\mathrm{h}}-L_{\mathrm{m}}\}}{\left(\exp\left\{\frac{L_{\mathrm{h}}-L_{\mathrm{m}}}{2}\right\}+\sqrt{\alpha}\right)^2}\gamma + o(1)\right\}, & L_{\mathrm{h}} - L_{\mathrm{m}} > \log\alpha; \end{cases} \tag{23}$$

where the minimum in (23) *may be shown to be attained for one of three possible values:*

- $L_{\mathrm{m}} = L_{\mathrm{h}}$*: This value corresponds to the side-channel-only upper bound of Theorem V.4.*
- $L_{\mathrm{m}} = 0$*: This corresponds to the assisted-noisy-channel-only upper bound of Corollary V.6.*
- $L_{\mathrm{m}} = L_{\mathrm{h}} - \log\alpha$ *(when* $\log\alpha < L_{\mathrm{h}}$*).*

The proof that the minimum is attained for one of the three values of L_{m} that are stated in Corollary V.6 is technical and is therefore omitted.

Clearly, by improving the bound (21) for $R > C_0$, the bound of Corollary V.5, and consequently also that of Corollary V.6, may be improved as well.

VI. Continuous-Time AWGN Channel with Unconstrained Bandwidth and a Helper

Up until now, we have concentrated on the setting of transmission over discrete-time AWGN channels. We now explain how to transform the results of Sections and IV and V to apply to continuous-time AWGN channels with an input power constraint and unconstrained bandwidth. In particular, we show that for the transmitter-assisted setting, if the helper is cribbed, namely, it knows both the message and the noise, one can achieve rates that are higher than the transmitter-assisted capacity when the helper knows only the noise but not the message (2).

We note that the channel-coding rate of the PPM-based schemes of Section V is $R = \log(M)/n < \log(n)/n$; meaning that it decays to zero as n grows. On the other hand, we have seen in Section V that, to attain an error probability that decays to zero, one must increase the total transmit energy E to infinity. Consequently, under a power constraint P (3), increasing the total energy $E = nP$ amounts to increasing the number of utilized channel uses n, which in turn means a decaying coding rate R to zero. In contrast, in continuous time with unconstrained bandwidth, this coupling between the total energy and the rate can be lifted since the number of utilized channel uses for a fixed time period (and total energy) can be made arbitrarily large by taking a large enough bandwidth. Consequently, the results of Section V can be readily applied to the power-limited continuous-time AWGN channel with unconstrained bandwidth as we show in this section.

To avoid ambiguity, we denote by roman (upright) letters (P, R, C, E) continuous-time quantities that are normalized per second to distinguish them from parallel discrete-time quantities that are normalized per sample and are denoted by italicized (slanted) letters (P, R, C, E). We further denote the transmission time in seconds by T.

We follow the expositions in [1], [5, Chapter 8] and [31, Chapter 9] (cf. [6, Chapter 5], [14, Chapters 2 and 3]) of the problem of transmission over a continuous-time bandlimited AWGN channel.

Parameter. The parameter to be conveyed is $u \in [-1/2, 1/2)$.

Transmitter. Maps the message W and the helper's description $V \in \{1, \dots, M_{\mathrm{h}}\}$ to a channel input signal $\{X(t) | t \in [0, \mathrm{T}]\}$ that is subject to an average power constraint P watts:

$$\int_0^{\mathrm{T}} \mathbb{E}\left[X^2(t)\right] dt \leq \mathrm{PT}, \tag{24}$$

and T is the transmission time in seconds.

Channel. The signal X is transmitted over a bandlimited continuous-time AWGN channel:

$$Y(t) = (X(t) + Z(t)) * g(t), \qquad t \in [0, \mathrm{T}], \tag{25}$$

where Y and Z are the output and noise signals, respectively;

$$g(t) = 2\mathrm{B} \cdot \mathrm{sinc}\left(2\mathrm{B}t\right) \triangleq \begin{cases} \frac{\sin(2\pi \mathrm{B}t)}{\pi t}, & t \neq 0; \\ 2\mathrm{B}, & t = 0; \end{cases}$$

is the impulse response of an ideal low-pass filter

$$G(f) = \mathrm{rect}\left(\frac{f}{2\mathrm{B}}\right) \triangleq \begin{cases} 1, & |f| < \mathrm{B}; \\ 0, & \text{otherwise}; \end{cases}$$

and sinc is the normalized sinc function; and Z is white Gaussian noise with two-sided spectral density $\mathrm{N}_0/2$ watts per hertz.

Helper. Knows (non-causally) $\{Z(t) * g(t) | t \in [0, \mathrm{T}]\}$ and maps it into a finite-rate description $V \in \{1, 2, \ldots, M_{\mathrm{h}}\}$ with $M_{\mathrm{h}} = \exp\{\mathrm{T}\mathrm{R}_{\mathrm{h}}\}$ where R_{h} is the help rate in nats per second. V is revealed to the transmitter prior to the beginning of the transmission.

Receiver. Constructs an estimate $\hat{U}$ of the parameter u given the output signal $\{Y(t) | t \in [0, \mathrm{T}]\}$.

Objective. The same as in Section II, namely minimizing the MPαE $\epsilon(\alpha)$ of (5). The corresponding optimal achievable MPαE error exponent is defined now as

$$\mathrm{E}(\alpha) \triangleq \varlimsup_{\mathrm{T}\to\infty} -\frac{1}{\mathrm{T}} \log \inf \epsilon(\alpha),$$

where the infimum is taken over all transmitter–receiver–helper triplets.

As in Section III, for the parallel channel-coding problem, instead of a parameter, a message W that is uniformly distributed over $\{1, 2, \ldots, M\}$ is modulated and transmitted by the transmitter, with $M = \exp\{\mathrm{T}\mathrm{R}\}$ where R is the rate in nats per second, and decoded by the receiver from the output signal.

Discrete-time reduction. By the Nyquist–Shannon sampling theorem [1], this problem is essentially reduced to a discrete-time problem with

- $n = 2\mathrm{B}\mathrm{T}$ samples with sampling time interval $\frac{1}{2\mathrm{B}}$, equivalently, $2\mathrm{B}$ samples per second;
- $P = \frac{\mathrm{P}}{2\mathrm{B}}$ power per sample;
- $\sigma^2 = \mathrm{N}_0/2$ noise variance;
- Rates $\mathrm{R} = 2\mathrm{B}R$, $\mathrm{C} = 2\mathrm{B}C$, $\mathrm{C}_0 = 2\mathrm{B}C_0$, and $\mathrm{R}_{\mathrm{h}} = 2\mathrm{B}R_{\mathrm{h}}$ in nats per second.

Applying these parameters to (2), results in the capacity per second of transmitter-assisted channel with bandwidth B (without cribbing)

$$\mathrm{C}(\mathrm{B}) = \mathrm{C}_0(\mathrm{B}) + \mathrm{R}_{\mathrm{h}}$$

where $\mathrm{C}_0(\mathrm{B}) = \mathrm{B}\log\left(1+\frac{\mathrm{P}}{\mathrm{N}_0\mathrm{B}}\right)$ is the capacity of this channel without assistance.

For the unconstrained-bandwidth setting, the capacity reduces to

$$\mathrm{C} \triangleq \lim_{\mathrm{B}\to\infty} \mathrm{C}(\mathrm{B}) = \mathrm{C}_0 + \mathrm{R}_\mathrm{h},$$

where $\mathrm{C}_0 = \mathrm{P}/\mathrm{N}_0$ is the power-limited capacity with unconstrained bandwidth without assistance [31, Chapter 8.2], [5, Chapter 8.2], [6, Chapter 5.6], [14, Chapters 2.5 and 3.6.1], [31, Chapter 9.3].

Using the discrete-time reduction above and Theorem III.3, setting $B \to \infty$ and noting that in this limit the channel becomes very noisy (recall the exposition in Section IV-D), results in the following upper bound on the achievable error exponent $\mathrm{E}(\cdot)$:

$$\Pr\left(\hat{W} \neq W\right) \leq \exp\left\{-\mathrm{T}\,\mathrm{E}_a(\mathrm{R})\right\},$$

$$\mathrm{E}_a(\mathrm{R}) = \begin{cases} \infty, & \mathrm{R} < \mathrm{R}_\mathrm{h}; \\ \frac{\mathrm{C}_0}{2} - \mathrm{R} + \mathrm{R}_\mathrm{h}, & 0 \leq \mathrm{R} - \mathrm{R}_\mathrm{h} < \frac{\mathrm{C}_0}{4}; \\ \left(\sqrt{\mathrm{C}_0} - \sqrt{\mathrm{R} - \mathrm{R}_\mathrm{h}}\right)^2, & \frac{\mathrm{C}_0}{4} \leq \mathrm{R} - \mathrm{R}_\mathrm{h} < \mathrm{C}_0; \\ 0, & \mathrm{R} \geq \mathrm{C}_0 + \mathrm{R}_\mathrm{h}. \end{cases} \tag{26}$$

Unfortunately, Theorem III.2 does not provide a meaningful lower bound on the error exponent for $\mathrm{R} < \mathrm{C}$.

Returning to the problem of parameter transmission, we attain the following upper and lower bounds as a corollary of Theorem IV.4 (and Corollary V.3) and Theorem IV.1, respectively, using the discrete-time reduction above and by appealing to the results for a very noisy channel of Section IV-D.

Corollary VI.1. *The α-th moment of the estimation error* (5) *with help rate R_h in nats per second, unconstrained bandwidth (arbitrarily large bandwidth), power P, two-sided spectral density $\mathrm{N}_0/2$, and transmission time T, is bounded from above as*

$$\epsilon(\alpha) \leq \exp\left\{-\mathrm{T}\alpha\left(\mathrm{R}_0^* + \mathrm{R}_\mathrm{h}\right) + o(\mathrm{T})\right\},$$

$$\mathrm{R}_0^* = \begin{cases} \frac{\mathrm{C}_0}{2(1+\alpha)}, & \alpha \geq 1; \\ \frac{\mathrm{C}_0}{(1+\sqrt{\alpha})^2}, & \alpha < 1; \end{cases}$$

and from below as

$$\epsilon(\alpha) \geq \frac{\alpha^{\alpha-1}\exp\left\{-\mathrm{T}\alpha(\mathrm{C}_0 + \mathrm{R}_\mathrm{h})\right\}}{[2\Gamma(1/\alpha)]^\alpha \mathrm{e}};$$

where $\mathrm{C}_0 = \mathrm{P}/\mathrm{N}_0$.

In particular, as discussed in Section IV-D, the exponents of the upper and the lower bounds of Corollary VI.1 coincide in the limit of small α.

We now consider the cribbed-helper scenario, namely, when the helper knows both the message and the noise before transmission.

Using the discrete-time reduction, we obtain the following corollary to Theorem V.3.

Corollary VI.2. *The capacity* C_C *in nats per second of the transmitter-assisted continuous-time AWGN channel* (25) *with help rate* $\mathrm{R_h}$ *in nats per second, input power constraint* P*, unconstrained bandwidth (arbitrarily large bandwidth), and two-sided noise spectral density* $\mathrm{N}_0/2$*, where the helper knows both the message and the noise, is bounded from below as*

$$\mathrm{C}_C \geq \mathrm{C}_0 + 2\sqrt{\mathrm{R_h C_0}},$$

where $C_0 = P/\mathrm{N}_0$ *is the capacity of this channel without assistance. In particular,* $\mathrm{C}_C > \mathrm{C}$ *for* $\mathrm{R_h} < 4\mathrm{C}_0$.

Furthermore, the error probability is bounded from above by

$$\Pr\left(\hat{W} \neq W\right) \leq \exp\left\{-\mathrm{T}\left(\mathrm{C}_0 + 2\sqrt{\mathrm{R_h C_0}} - \mathrm{R}\right) + o(\mathrm{T})\right\}$$

for $\mathrm{R} < \mathrm{C}_0 + 2\sqrt{\mathrm{R_h C_0}}$*, where* T *is the total transmission time, namely, an error exponent of*

$$\mathrm{E}_C(\mathrm{R}) = \mathrm{C}_0 + 2\sqrt{\mathrm{R_h C_0}} - \mathrm{R}$$

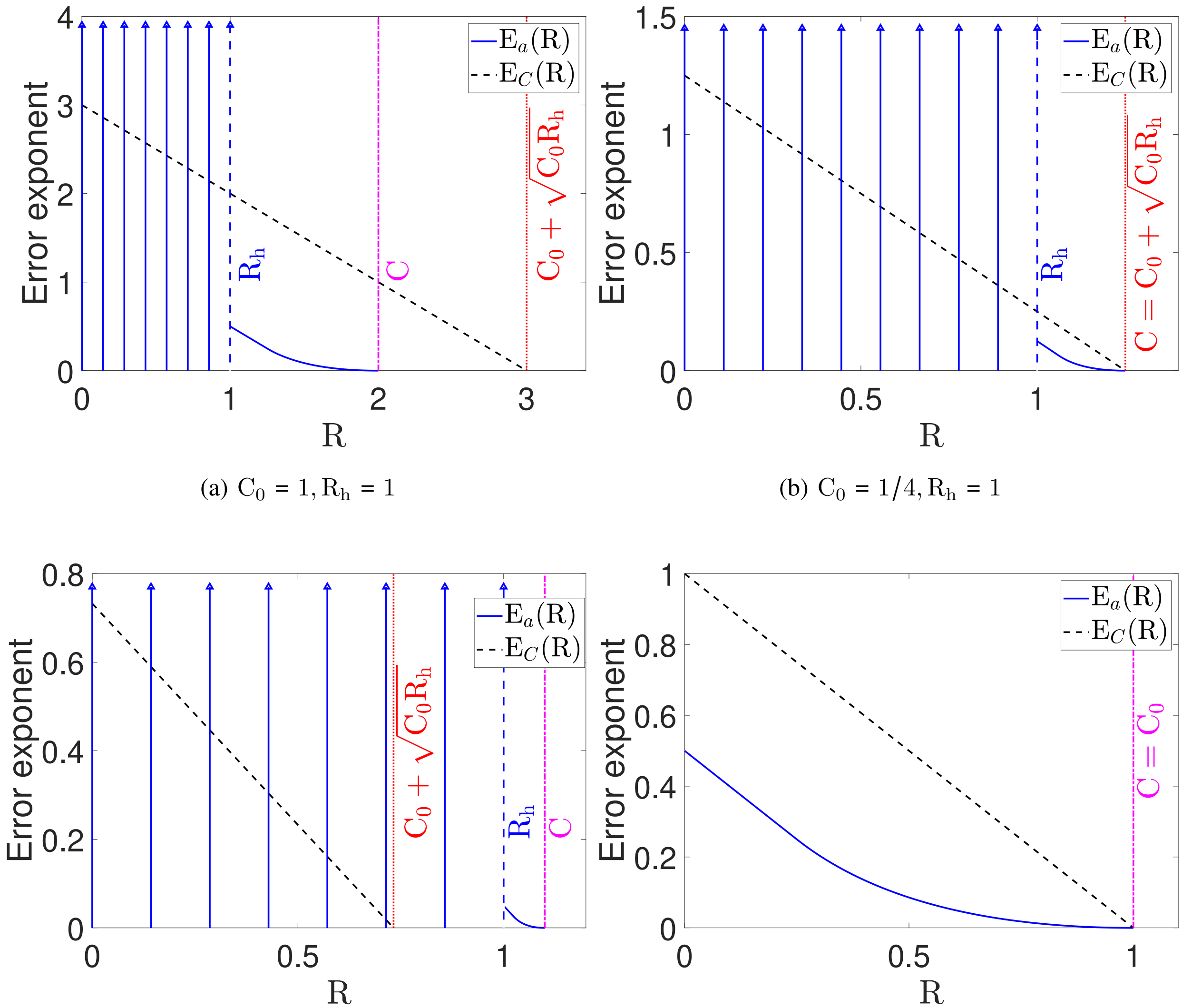
is achievable.

Proof: We follow [5, Chapter 8.2] (cf. [6, Chapter 5.6], [14, Chapters 2.5 and 3.6.1]) by applying Scheme V.2 the discrete-time reduction:

- $n = 2\mathrm{BT}$, $P = \frac{\mathrm{P}}{2B}$, $\sigma^2 = \mathrm{N}_0/2$;
- rates $\mathrm{R} = L/\mathrm{T} = \log(M)/\mathrm{T}$ and $\mathrm{R_h} = L_\mathrm{h}/\mathrm{T} = \log(M_\mathrm{h})/\mathrm{T}$ of nats per second.

For this choice

- $E = \mathrm{PT}$, where P is the transmit power in Joules per second (cf. P from Section II which was in Joules per channel use);
- $\gamma = E/\sigma^2 = 2\mathrm{PT}/\mathrm{N}_0 = 2\mathrm{TC}_0$.



(a) $C_0 = 1, R_h = 1$

(b) $C_0 = 1/4, R_h = 1$

(c) $C_0 = 1/10, R_h = 1$

(d) $C_0 = 1, R_h = 0$

Fig. 4: Achievable transmitter-assisted channel-coding error exponents as a function of the rate R for various values of the help rate R_h and the capacity without assistance C_0: $E_a(R)$—the achievable error exponent of (26) of a helper that knows only the noise, $E_C(R)$—the achievable of Corollary VI.1 of a helper that knows both the noise and the message. The blue vertical arrows denote an arbitrarily large error exponent $E_a(R)$ for $R < R_h$.

We note that for fixed rates R and R_h, the utilized bandwidth of the scheme with the above choice of parameters,

$$B = \frac{\exp\{(R + R_h)\,T\}}{2T},$$

grows exponentially with T.

Substituting these parameters in Theorem V.3 yields the desired result. ■

We compare the achievable error exponents $\mathrm{E}_a(\mathrm{R})$ of (26) and $\mathrm{E}_C(\mathrm{R})$ of Corollary VI.2 in Figure 4. First, note that for $\mathrm{R} < \mathrm{R_h}$, $\mathrm{E}_a(\mathrm{R})$ is arbitrarily large whereas $\mathrm{E}_a(\mathrm{R})$ is bounded. In Figure 4, the arbitrarily large error exponent $\mathrm{E}_a(\mathrm{R})$ for $\mathrm{R} < \mathrm{R_h}$ is illustrated by blue vertical arrows. For $\mathrm{R} > \mathrm{R_h}$, on the other hand, both achievable error exponents are bounded. Moreover, whenever $\mathrm{R_h} \le 4\mathrm{C_0}$, $\mathrm{E}_C(\mathrm{R})$ dominates $\mathrm{E}_a(\mathrm{R})$ for $\mathrm{R} \in (\mathrm{R_h}, \mathrm{C_0} + \mathrm{R_h})$. This is demonstrated in Figures 4a and 4b, wherein the former figure the condition $\mathrm{R_h} \le 4\mathrm{C_0}$ is strict and hence achievable rates that exceed C are attainable when the helper is cribbed, whereas in the latter figure this condition holds with equality and hence the maximal achievable rate of Corollary VI.2 coincides with C.

Figure 4c depicts the two exponents for the case of $\mathrm{R_h} \gg 4\mathrm{C_0}$. In this case, the maximal achievable rate of Corollary VI.2 is lower than R_h and hence $\mathrm{E}_a(\mathrm{R})$ dominates $E_C(\mathrm{R})$.

It is interesting to note that there is a discontinuity at $\mathrm{R_h} = 0$ in the error exponent: In the limit of small $\mathrm{R_h}$, the error exponent of Corollary VI.2 converges to

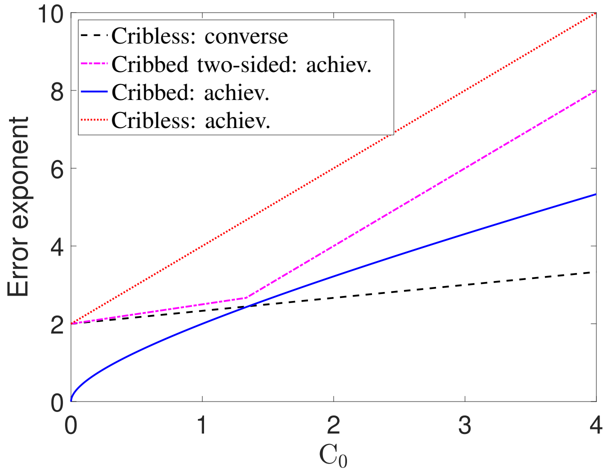
$$\lim_{\mathrm{R_h} \to 0} \mathrm{E}_C(\mathrm{R}, \mathrm{R_h}) = \mathrm{C_0} - \mathrm{R},$$

which is strictly higher, for all $\mathrm{R} < \mathrm{C_0}$, than the optimal error exponent without help over the continuous-time AWGN channel with unconstrained bandwidth equals [51][5]

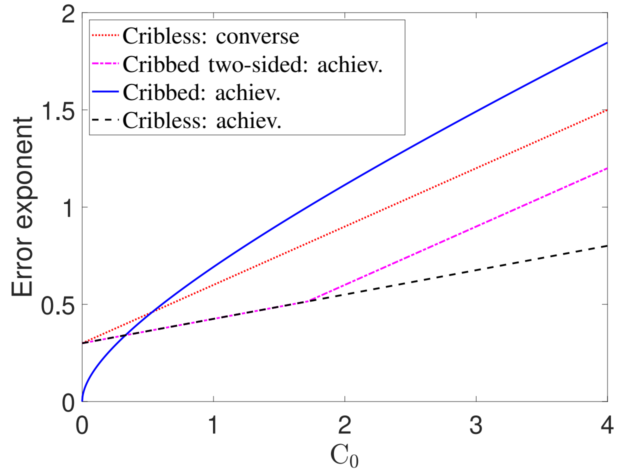
$$\mathrm{E}(\mathrm{R}) = \begin{cases} \frac{\mathrm{C_0}}{2} - \mathrm{R}, & \mathrm{R} < \frac{\mathrm{C_0}}{4}; \\ \left(\sqrt{\mathrm{C_0}} - \sqrt{\mathrm{R}}\right)^2, & \frac{\mathrm{C_0}}{4} \le \mathrm{R} \le \mathrm{C_0}; \\ 0, & \mathrm{R} \ge \mathrm{C_0}, \end{cases} \tag{27}$$

which agrees with $\mathrm{E}_a(\mathrm{R})$ for $\mathrm{R_h} \to 0$. While surprising, we note that a similar phenomenon happens in the presence of feedback for discrete channels and random transmission duration time, where zero-rate feedback (e.g., stop-feedback) attains a similar achievable straight-line error exponent [55]–[58]; in fact, since the helper knows both the message and the noise, it can mimic feedback operation. This is depicted in Figure 4d, where clearly $E_C(\mathrm{R})$ dominates $\mathrm{E}(\mathrm{R})$ [which coincides with (27)].

[5]The optimality of this error exponent assumes that (24) holds for every message w and not on average. While, equal-energy codes (including simplex codes) are suboptimal in general [52] (see also [53], [54, Chapter 4]) they can be shown to achieve the optimal exponential decay with T by an expurgation argument with respect to the codeword energies.



(a) $\alpha = 2, R_h = 1$



(b) $\alpha = 0.3, R_h = 1$

Fig. 5: Achievable MPαE error exponents as a function of the capacity without assistance C_0 for help rate $R_h = 1$ and $\alpha = 2, 0.3$: “Cribless: converse” and “Cribless: achievable” denote the upper and lower bounds on the MPαE error exponent of Corollary VI.1 for the scenario of a helper that knows only the noise and assists only the transmitter, “Cribbed: achiev.” denotes the achievable error exponent of Corollary VI.3 for the scenario of a cribbed helper that assists only the transmitter, and “Cribbed two-sided: achiev.” denotes the achievable error exponent of Corollary VI.5 for the scenario of a cribbed helper that assists both the transmitter and the receiver.

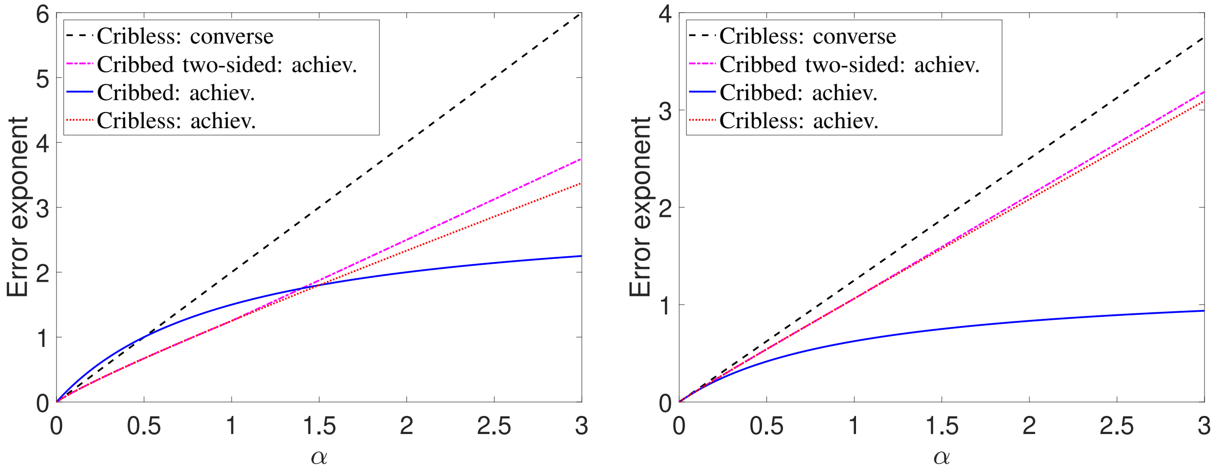
Returning to the parameter transmission problem with a cribbed helper, we attain the following bound on the achievable MPαE upon applying the discrete-time reduction to Corollary V.4.

Corollary VI.3. *The α-th moment of the estimation error* (5) *with help rate R_h in nats per second, unconstrained bandwidth (arbitrarily large bandwidth), power P, two-sided spectral density $N_0/2$, transmission time T, and a helper that knows both the message and the noise is bounded from above as*

$$\epsilon(\alpha) \leq \exp\left\{-\frac{\alpha}{1+\alpha}\left(C_0 + 2\sqrt{R_h C_0}\right)T + o(T)\right\},$$

where $C_0 = P/N_0$.

Comparing the upper and lower bounds on the MPαE error exponent of Corollary VI.1 (cribless helper) to the achievable MPαE error exponent of Corollary VI.3 (cribbed helper), we see that the achievable cribbed-helper MPαE error exponent outperforms the achievable error exponent without cribbing of Corollary VI.1 for a sufficiently large C_0 for a fixed value of



(a) $\mathrm{C}_0 = 1, \mathrm{R}_\mathrm{h} = 1$ (b) $\mathrm{C}_0 = 1/4, \mathrm{R}_\mathrm{h} = 1$

Fig. 6: Achievable MPαE error exponents as a function of α for help rate $\mathrm{R}_\mathrm{h} = 2, 0.3$ and two values of the capacity without assistance $\mathrm{C}_0 = 1, 1/4$: "Cribless: converse" and "Cribless: achievable" denote the upper and lower bounds on the MPαE error exponent of Corollary VI.1 for the scenario of a helper that knows only the noise and assists only the transmitter, "Cribbed: achiev." denotes the achievable error exponent of Corollary VI.3 for the scenario of a cribbed helper that assists only the transmitter, and "Cribbed two-sided: achiev." denotes the achievable error exponent of Corollary VI.5 for the scenario of a cribbed helper that assists both the transmitter and the receiver.

R_h. This behavior is illustrated in Figure 5 and agrees with the behavior observed for channel coding in Figure 4.

Interestingly, although the upper (converse) bound on the achievable MPαE error exponent of Corollary VI.1 is somewhat weak since it relies on the DPT, the achievable cribbed-helper MPαE error exponent of Corollary VI.3 exceeds it for low values of α and sufficiently large help-rate values, as is evident from Figures 5b and 6a.

For $\mathrm{C}_0 = 1/4$ and $\mathrm{R}_\mathrm{h} = 1$, we have seen that the channel-coding error exponent of a cribbed-helper dominates the optimal achievable error exponent without cribbing in Figure 4b for $\mathrm{R} > \mathrm{R}_\mathrm{h}$ (which is the relevant rate region for the derivation of the MPαE performance). Thus, one may expect the cribbed-helper MPαE error exponent to dominate the cribless-helper MPαE error exponent. Yet, Figure 6b clearly illustrates the opposite. The reason for that seeming contradiction is the unequal protection that is offered by Theorem III.2—where a rate of R_h nats per second can be conveyed essentially error-free—and is leveraged by Scheme IV.2, whereas the channel-coding error exponent of the message w in Scheme V.2 is finite for all R values.

When the cribbed helper is available both at the transmitter and the receiver, Proposition V.1 suggests an achievable doubly-exponential error probability decay rate with time T. This is formulated in the next corollary, whose proof is very similar to that of Corollary VI.2, and is therefore omitted.

Corollary VI.4. *Let* $P_e(\mathrm{R},\mathrm{T})$ *by the error probability of a channel-coding scheme with rate* R *nats per second and transmission time* T *seconds over the continuous-time AWGN channel* (25) *(without a helper) with input power constraint* P*, unconstrained bandwidth (arbitrarily large bandwidth), and two-sided noise spectral density* $\mathrm{N}_0/2$*. Then, there exists a channel-coding scheme with the same rate* R *and transmission time* T*, with a helper with help rate* $\mathrm{R_h}$ *that knows both the message and the noise, and assists both the transmitter and the receiver, that attains an error probability of*

$$\Pr\left(\hat{W} \neq w\right) = \left[P_e(\mathrm{R},\mathrm{T})\right]^{\exp\{L_\mathrm{h}\}}.$$

In particular, the error probability decays doubly exponentially with $\mathrm{R_h}$ *for any* R*, and it decays doubly exponentially with* T *for* $\mathrm{R} < \mathrm{C}_0$*.*

Applying the discrete-time reduction to Corollary V.6 yields the following result, whose proof is available in Appendix G.

Corollary VI.5. *The* α*-th moment of the estimation error* (5) *with help rate* $\mathrm{R_h}$ *in nats per second, unconstrained bandwidth (arbitrarily large bandwidth), power* P*, two-sided spectral density* $\mathrm{N}_0/2$*, transmission time* T*, and a helper that knows both the message and the noise and assists both the transmitter and the receiver is bounded from above as*

$$\epsilon(\alpha) \leq \exp\left\{-\mathrm{T}\alpha \max\left\{\mathrm{C}_0, \mathrm{R_h} + \frac{\mathrm{C}_0}{\min\left(4, \{1+\sqrt{\alpha}\}^2\right)}\right\} + o(1)\right\},$$

where $\mathrm{C}_0 = \mathrm{P}/\mathrm{N}_0$ *and* $o(1)$ *is with respect to* T*.*

The lower bound on the MPαE error exponent of Corollary VI.5 for the two-sided cribbed helper scenario coincides with the lower bound of Corollary VI.1 for a cribless helper that is available only at the transmitter for $\alpha \leq 1$ and $\mathrm{R_h} \geq 0.75\mathrm{C}_0$, and is strictly better otherwise. This is also illustrated in Figures 5 and 6.

However, when comparing the lower bound on the MPαE error exponent of Corollary VI.5 for the two-sided cribbed helper scenario with that of Corollary VI.3 for the scenario of a cribbed helper that assists only the transmitter, we see in Figures 5 and 6 that the latter exceeds the former when α is low and C_0 is large enough compared to R_h. Since Scheme V.3 (all the more so Scheme V.4) majorizes Scheme V.2, this means that bound of Corollary VI.5 is loose in these regimes as the bound of Corollary VI.3 serves also as a lower bound on the MPαE error exponent for the two-sided cirbbed helper scenario.

VII. Summary and Discussion

In this work, we derived exponential upper and lower bounds on the optimal achievable α-th moment absolute error in transmitting a parameter over an AWGN channel with a helper that knows the noise before transmission begins. The lower (impossibility) bound utilized the Ziv–Zakai bounding technique, whereas the upper (achievability) bound used judiciously the unequal error protection property of the transmitter-assisted channel coding technique of [43]. These bounds coincide for low values of α, but a gap remains for other values.

We further considered the setting of energy-limited channel coding and parameter transmission over this channel, for which the previously proposed schemes do not work as is, and refined the achievability technique [41], [43] and its analysis. For the energy-constrained setting where the helper knows both the message ("cribbing") and the noise before transmission begins and reveals it to the transmitter, we proposed a PPM-based scheme that outperforms these results under a condition between the help rate and the available energy. When such a helper assists both the transmitter and the receiver, we showed that a doubly-exponential error probability decay rate with the help rate is achievable.

Finally, we translated the results for both the power- and energy-limited settings to results for the setting of transmission over a continuous-time AWGN noise with an input power constraint and unconstrained bandwidth. In particular, for transmitter-assisted channel coding, we found achievable rates when the helper knows both the message and the noise that are higher than $\mathrm{C}_0 + \mathrm{R}_\mathrm{h}$—the capacity of the same channel when the helper knows only the noise but not the message [41]. Moreover, an achievable error exponent in the limit of zero help rate, $R_\mathrm{h} \to 0$, was shown to equal $C_0 - R$, which is reminiscent of Burnashev's error exponent [55].

The results of Proposition V.1 and Corollary VI.4 rely on error probability results for channel coding without a helper, with most upper and lower bounds becoming trivial for $R \geq C_0$. To attain meaningful results for $R \geq C_0$, more delicate bounds for this rate region need to be employed, such as the bounds on correct decoding above capacity [49], [59], [60]. This is left for future research.

APPENDIX A
PROOF OF THEOREM IV.1

Note first that

$$\epsilon(\alpha) \triangleq \sup_{u\in[-1/2,1/2)} \mathbb{E}\left[\left|\hat{U}-u\right|^{\alpha}\right]$$
$$\geq \mathbb{E}\left[\left|\hat{U}-U\right|^{\alpha}\right]$$

where U is uniformly distrubted over $[-1/2, 1/2]$.

Given the noise description V, $U \to \mathbf{X} \to \mathbf{Y}$ forms a Markov chain, and so,

$$I(U;\mathbf{Y}|V) \leq I(\mathbf{X};\mathbf{Y}|V). \tag{28}$$

Now, the right-hand side of (28) is bounded from above as

$$I(\mathbf{X};\mathbf{Y}|V) = h(\mathbf{Y}|V) - h(\mathbf{Y}|\mathbf{X},V) \tag{29a}$$
$$\leq h(\mathbf{Y}|V) - \frac{n}{2}\log(2\pi e\sigma^2) + nR_{\mathrm{h}} \tag{29b}$$
$$\leq \frac{n}{2}\log\left[2\pi e\left(P+\sigma^2\right)\right] - \frac{n}{2}\log\left(2\pi e\sigma^2\right) + nR_{\mathrm{h}} \tag{29c}$$
$$= n(C_0 + R_{\mathrm{h}}),$$

where (29b) and (29c) are supported, respectively, by the chains (32)–(36) and (37)–(46) in [41].

The right-hand side of (28), on the other hand, is bounded according to the Shannon lower bound as follows.

$$I(U;\mathbf{Y}|V) = h(U|V) - h(U|V,\mathbf{Y}) \tag{30a}$$
$$= h(U) - h(U|V,\mathbf{Y}) \tag{30b}$$
$$= h(U) - h(U-\hat{U}|V,\mathbf{Y}) \tag{30c}$$
$$\geq 0 - h(U-\hat{U}) \tag{30d}$$
$$\geq -\frac{1}{\alpha}\log\left\{\frac{[2\Gamma(1/\alpha)]^{\alpha}e}{\alpha^{\alpha-1}}\cdot\mathbf{E}|U-\hat{U}|^{\alpha}\right\}, \tag{30e}$$

where (30b) holds since U is independent of V, (30d) holds since conditioning cannot increase the differential entropy and since U is uniformly distributed over an interval of unit length, and

(30e) follows from maximizing the differential entropy under an α-th moment constraint [31, Chapter 12].

The desired result then follows by substituting (29) and (30) in (28). ■

Appendix B
Proof of Theorem IV.2

Define U to be uniformly distributed over the finite set $\mathcal{U}_M \triangleq \left\{\frac{1}{M}\left(k-\frac{1}{2}\right)-\frac{1}{2}\middle|k=1,\ldots,M\right\}$ such that all the points in this set belong to $[-1/2,1/2)$. Then,

$$\begin{aligned}
\epsilon(\alpha) &\triangleq \sup_{u\in[-1/2,1/2)} \mathbb{E}\left[\left|\hat{U}-u\right|^{\alpha}\right] &\text{(31a)}\\
&\geq \mathbb{E}\left[\left|\hat{U}-U\right|^{\alpha}\right] &\text{(31b)}\\
&\geq \frac{1}{(2M)^{\alpha}}\Pr\left(\left|\hat{U}-U\right|^{\alpha} > \frac{1}{(2M)^{\alpha}}\right) &\text{(31c)}\\
&\geq \frac{1}{(2M)^{\alpha}}\inf_{\hat{U}}\Pr\left(\hat{U}\neq U\right) &\text{(31d)}\\
&\geq \exp\left\{-\alpha R n\right\}\exp\left\{-n\left[E_{\text{wsp}}(R-\varepsilon)+o(n)\right]\right\} &\text{(31e)}
\end{aligned}$$

where (31a) holds by definition (5); (31c) follows from the Markov inequality; the infimum in (31d) is over all estimators of U given $\mathbf{Y}$, and holds since a point $\hat{U}\in\mathbb{R}$ can be within a distance $\frac{1}{2M}$ of at most a single point in $\mathcal{U}_M$; (31e) follows from Theorem III.3.

To tighten the bound in (31), let us optimize the bound in (31e) over R, namely,

$$\begin{aligned}
&\min_{R\in\mathbb{R}}\left\{\alpha R+E_{\text{wsp}}(R)\right\}\\
&= \min_{R\in[0,C_0+R_{\text{h}}]}\left\{\alpha R+\frac{1}{2}\left[\zeta(R)-\log\zeta(R)-1\right]\right\} &\text{(32a)}\\
&= \alpha R_{\text{h}}+\frac{1}{2}\min_{s\geq 1}\left\{\alpha\log\left(1+\frac{S}{s}\right)+s-\log s-1\right\} &\text{(32b)}\\
&= \alpha R_{\text{h}}+\frac{1}{2}\left\{\alpha\log\left(1+\frac{S}{s^*}\right)+s^*-\log s^*-1\right\} &\text{(32c)}
\end{aligned}$$

where (32a) follows by the definition of E_{wsp} in Theorem III.3, (32b) follows by the definition of $\zeta(R)$ in Theorem III.3 by substituting $s=\zeta(R)$, and (32c) holds by solving the minimization in (32b), resulting in

$$\begin{aligned}
s^* &\triangleq \frac{\sqrt{(1+S)^2+4\alpha S}}{2}-\frac{S-1}{2}\\
&= \frac{2(\alpha+1)S}{\sqrt{(S+1)^2+4\alpha S}+S-1}
\end{aligned}$$

and by noting that $s^* \geq 1$ for all $S \geq 0$.

Substituting (32) in (31) completes the proof. ■

Appendix C
Proof of Corollary IV.1

Impossibility. Define U to be uniformly distributed over the discrete finite set

$$\left\{u + \frac{i}{M} \middle| i \in \{0, \ldots, M-1\}\right\}^d$$

such that all the points in this set belong to $[-1/2, 1/2)^d$. Then, the MPαE is bounded from below as follows.

$$\sum_{i=1}^{d} \mathbb{E}\left[|\hat{U}_i - u_i|^\alpha\right] \geq \sum_{i=1}^{d} \frac{1}{M^\alpha} \Pr\left(|\hat{U}_i - U_i| \geq \frac{1}{M}\right) \tag{33a}$$

$$\geq \exp\{-n\alpha r\} \Pr\left(\bigcup_{i=1}^{d} \left\{|\hat{U}_i - U_i| \geq \exp\{-nr\}\right\}\right) \tag{33b}$$

$$\geq \exp\{-n\alpha r\} \exp\{-nE_{\text{wsp}}(r \cdot d - \varepsilon) + o(n)\} \tag{33c}$$

$$= \exp\left\{-n\frac{\alpha}{d}R\right\} \exp\{-nE_{\text{wsp}}(R - \epsilon) + o(n)\} \tag{33d}$$

where (33a) follows from the Markov inequality; (33b) follows from the union bound by taking $M = \exp\{r\}$; (33c) follows from Theorem III.2; and (33d) holds for $R \triangleq r \cdot d$. Noting that (33d) is identical to (31e) with α/d in lieu of α, completes the proof of the lower bound.

Achievability. The achievability part is the same as before except that the quantization stage is carried out by applying uniform quantization of each parameter dimension at rate R/d. ■

Appendix D
Proofs for Section V-A

Proof of Theorem V.2: We refine the analysis in the proof of Theorem III.2 in [43, Section IV-A] with a refined analysis and choice of design parameters that depend on n. Consider first the case of a receiver-only helper, in which the quantized noise is subtracted by the receiver. In particular, instead of the fixed in n value τ, let $\tau_n \in [0, 1]$ be a sequence with the following two properties: (i) $\lim_{n\to\infty} n\tau_n = \infty$; (ii) $\lim_{n\to\infty} \tau_n \log n = 0$. In other words, τ_n decays faster than $1/\log n$ but slower than $1/n$, e.g., $\tau_n = 1/\log^2 n$ or $\tau_n = 1/\sqrt{n}$. Divide the block of length n into two sub-blocks: a sub-block of length $n\tau_n$, for which the receiver receives a quantized noise description from the helper at rate nR_{h}, whereas over the remaining part of the block, of length $n(1-\tau_n)$, no help is provided at all. A uniform scalar quantizer is used for all $t \in \{1,\ldots,\tau_n n\}$. In the limit $n \to \infty$, $\tau_n \to 0$, $R_{\mathrm{h}}/\tau_n \to \infty$, meaning that the quantizer operates in the high-resolution regime. More precisely, consider the $\tau_n n$-dimensional ball of radius $\sqrt{\tau_n n\sigma^2\left(1+\frac{B}{\tau_n}\right)}$, centered at the origin, in the space of noise sequences, $\{\mathbf{Z}_{1:\tau_n n}\}$, where $B>0$ is a design parameter, to be chosen later. The helper's lossy compression scheme is based on partitioning this ball into hypercubes of size $\Delta_n > 0$ and quantizing $\mathbf{Z}_{1:\tau_n n}$ into the center of the hypercube to which it belongs. If $\mathbf{Z}_{1:\tau_n n}$ falls outside the ball, then the compression fails and an overload error occurs. Accordingly, the step size Δ_n of the uniform scalar quantizer is chosen such that the logarithm of the number of hypercubes whose union fully covers the ball would not exceed nR_{h}. In other words, Δ_n should be chosen such that $\exp\{nR_{\mathrm{h}}\}$ hypercubes must fully cover the ball. This amounts to the following relationship:

$$nR_{\mathrm{h}} = \left\lceil \log \frac{\mathrm{Vol}\left\{\mathscr{B}_{\tau_n n}\left(\sqrt{\tau_n n\sigma^2\left(1+\frac{B}{\tau_n}\right)}+\sqrt{t}\cdot\Delta_n\right)\right\}}{\Delta_n^{\tau_n n}} \right\rceil \tag{34a}$$

$$= \frac{n\tau_n}{2}\log\frac{2\pi\mathrm{e}\left(\sigma\sqrt{1+\frac{B}{\tau_n}}+\Delta_n\right)^2}{\Delta_n^2} + o(n), \tag{34b}$$

where the right-hand side of (34a) is an upper bound on the number of covering hypercubes, with $\sqrt{\tau_n n\sigma^2(1+B/\tau_n)}+\sqrt{\tau_n n}\cdot\Delta_n$ being an upper bound on the radius of a larger ball that contains (bounds from above) all hypercubes whose union completely covers[6] the original ball of radius $\sqrt{\tau_n n\sigma^2(1+B/\tau_n)}$. By demanding that Δ_n would be large enough such that the first equality holds, we guarantee that the rate budget of the helper is enough to support the implementability of this scheme. Equivalently, from the second equality, we find Δ_n to be given by

$$\Delta_n = \frac{\sqrt{2\pi e\sigma^2\left(1+\frac{B}{\tau_n}\right)}}{\exp\left\{\frac{R_{\mathrm{h}}-o(1)}{\tau_n}\right\}-\sqrt{2\pi\mathrm{e}}}, \tag{35}$$

where for sufficiently small $\tau_n > 0$, the denominator is positive. Let $\mathbf{Z}'_{1:\tau_n n} = q\left(\mathbf{Z}_{1:\tau_n n}\right)$ be the quantized version of $\mathbf{Z}_{1:\tau_n n}$ using this quantizer.

[6]The second term, $\sqrt{\tau_n n}\cdot\Delta_n$, is the length of the main diagonal of a $\tau_n n$-dimensional hypercube of size Δ_n, which is the largest excess radius that each covering hypercube can add to the original ball.

During the first sub-block of length $\tau_n n$, the transmitter sends $\mathbf{X}_{1:\tau_n n}$ that depends only on the message. The corresponding segment of the received signal, after subtracting $\mathbf{Z}'_{1:\tau_n n}$, is then

$$\begin{aligned}\mathbf{Y}_{1:\tau_n n} - \mathbf{Z}'_{1:\tau_n n} &= \mathbf{X}_{1:\tau_n n} + \mathbf{Z}_{1:\tau_n n} - \mathbf{Z}'_{1:\tau_n n} \\ &= \mathbf{X}_{1:\tau_n n} + \tilde{b}'' Z_{1:\tau_n n},\end{aligned}$$

where $\tilde{\mathbf{Z}}_{1:\tau_n n} \triangleq \mathbf{Z}_{1:\tau_n n} - \mathbf{Z}'_{1:\tau_n n}$ is the residual quantization noise. As long as the norm of $\mathbf{Z}_{1:\tau_n n}$ is less than $\sqrt{\tau_n n \sigma^2 (1 + B/\tau_n)}$, the quantization error sequence $\tilde{\mathbf{Z}}_{1:\tau_n n}$, lies within the hypercube $[-\Delta_n/2, \Delta_n/2]^{\tau_n n}$. Therefore, if the transmitter uses a quantizer with step size Δ_n for each coordinate, the transmission in this segment will be error-free (beyond the overload error event), as the residual noise sequence cannot cause a passage to the hypercube of any other codeword provided it falls within the ball. Such a lattice code can therefore support essentially error-free transmission of nR' information nats, where

$$\begin{aligned}nR' &\geq \frac{n\tau_n}{2} \log\left(\frac{2\pi \mathrm{e} E}{n\tau_n \Delta_n^2}\right) - o(n) - n\tau\epsilon(\Delta_n) && \text{(36a)}\\ &= \frac{n\tau_n}{2} \log \frac{2\pi \mathrm{e} E \cdot (\exp\{(R_\mathrm{h} - o(1))/\tau_n\} - \sqrt{2\pi e})^2}{2\pi n \tau_n \mathrm{e} \sigma^2 (1 + B/\tau_n)} - o(n) - n\tau_n \epsilon(\Delta_n) && \text{(36b)}\\ &= n[R_\mathrm{h} - o(1)] + \frac{n\tau_n}{2} \log \frac{E}{n\tau_n \sigma^2 (1 + B/\tau_n)} \\ &\quad + n\tau_n \log\left(1 - \sqrt{2\pi \mathrm{e}} \cdot \exp\{-(R_\mathrm{h} - o(1))/\tau_n\}\right) - o(n) - n\tau_n \epsilon(\Delta_n), && \text{(36c)}\end{aligned}$$

where $\epsilon(\Delta_n) > 0$ is a function with the property $\lim_{\Delta_n \to 0} \epsilon(\Delta_n)/\Delta_n < \infty$, (36a) is proved in the appendix of [43], and (36b) follows from substituting (35) in the main term of the resulting expression. Clearly, since τ_n tends to zero such that $\tau_n \log n \to 0$, then R' approaches a limit that is not less than R_h. In other words, we can transmit at any rate, arbitrarily close to R_h nats per channel use, error-free, provided that $\mathbf{Z}_{1:\tau_n n} \in \mathscr{B}_{\tau_n n}\left(\tau_n n \sigma^2 (1 + B/\tau_n)\right)$. An error will occur, in this segment, only if $\mathbf{Z}_{1:\tau_n n}$ falls outside the aforementioned union of hypercubes (in the space of noise sequences), which in turn implies that $\mathbf{Z}_{1:\tau_n n} \notin \mathscr{B}_{\tau_n n}\left(\tau_n n \sigma^2 (1 + B/\tau_n)\right)$ as those hypercubes cover the sphere of radius $\sqrt{\tau_n n \sigma^2 (1 + B/\tau_n)}$; the probability of this event, namely, $\{\sum_{i=1}^{\tau_n n} Z_i^2 > \tau_n n \sigma^2 (1 + B/\tau_n)\}$, is easily upper-bounded by the Chernoff bound:

$$\Pr\left\{\sum_{i=1}^{\tau_n n} Z_i^2 > \tau_n n \sigma^2 \left(1 + \frac{B}{\tau_n}\right)\right\} \leq \exp\left\{-\frac{n}{2}\left[B - \tau_n \log\left(1 + \frac{B}{\tau_n}\right)\right]\right\},$$

which behaves like $\exp\{-nB/2\}$ as $\tau_n \to 0$. Hence, we can make the exponential decay of this probability as fast as desired by choosing B sufficiently large.

The number of information nats that we can encode in the first sub-block (of length $n\tau_n$) is therefore

$$nR' \geq nR_\mathrm{h} + \frac{n\tau_n}{2}\log\frac{E}{n\tau_n\sigma^2(1+B/\tau_n)} + n\tau_n\log\left(1-\sqrt{2\pi\mathrm{e}}\cdot\exp\left\{-\frac{R_\mathrm{h}-o(1)}{\tau_n}\right\}\right)$$
$$- o(n) - n\tau_n\epsilon(\Delta_n) \tag{38a}$$
$$= nR_\mathrm{h} + \frac{n\tau_n}{2}\log\frac{E}{n\sigma^2(B+\tau_n)} + n\tau_n\log\left(1-\sqrt{2\pi e}\cdot\exp\left\{-\frac{R_\mathrm{h}-o(1)}{\tau_n}\right\}\right)$$
$$- o(n) - n\tau_n\epsilon(\Delta_n), \tag{38b}$$

which is arbitrarily close nR_h for a sufficiently small τ_n, as the second and the third terms in (38b) tend to zero as $\tau_n \log n \to 0$. Thus, we can transmit at a rate arbitrarily close to R_h nats with an error exponent that is as large as desired.

Finally, we use this channel code in cascade with uniform quantization of U, as before. The overall α-th moment of the estimation error is of the order of $\exp\{-n\alpha R_\mathrm{h}\} + \exp\{-nB/2\}$, where the second term can be made exponentially negligible as n grows by selecting $B > 2\alpha R_\mathrm{h}$.

Finally, if only the transmitter receives help, we can still approach the DPT converse bound by letting τ_n tend to zero such that $n\tau_n$ is a constant, as long as it is a reasonably large constant. In this case, the quantized noise sequence, $q\left(\mathbf{Z}_{1:\tau_n n}\right)$, can be subtracted from the codeword, provided that the allowed energy, E, is at least as large as $\sigma^2\tau_n n$. All the above calculations continue to hold with a constant $\tau_n n$ in mind, as the Stirling approximation applies reasonably well to $\Gamma(t+1/2)$ (which plays a role as a factor in the formula of the volume of a $\tau_n n$-dimensional ball). ■

Proof of Corollary V.3: The resulting MPαE of Scheme IV.2 with $L = \log M$ is bounded for any $u \in [-1/2, 1/2)$ as follows.

$$\mathbb{E}\left[\left|\hat{U}-u\right|^\alpha\right] \leq \frac{1}{M^\alpha} + \Pr\left(\hat{W}_\mathrm{m} \neq w_\mathrm{m}\right) + \frac{1}{(M')^\alpha}\Pr\left(\hat{W}_\ell \neq w_\ell\right) \tag{39a}$$
$$\leq \exp\{-\alpha L\} + \exp\{-nE_\infty\}$$
$$+ \exp\{-n\alpha(R_\mathrm{h}-\epsilon)+o(n)\}\cdot\begin{cases}\exp\{-nE_\infty\}, & L < nR_\mathrm{h};\\ \exp\left\{-\left(\frac{\gamma}{4}-L+nR_\mathrm{h}\right)\right\}, & 0 \leq L-nR_\mathrm{h} < \frac{\gamma}{8};\\ \exp\left\{-\left(\sqrt{\frac{\gamma}{2}}-\sqrt{L-nR_\mathrm{h}}\right)^2\right\}, & \frac{\gamma}{8} \leq L-nR_\mathrm{h} < \frac{\gamma}{2};\end{cases} \tag{39b}$$

$$\leq \exp\{n\epsilon + o(n)\} \cdot \begin{cases} \exp\{-\alpha L\}, & L < nR_\mathrm{h}; \\ \exp\left\{-\min\left\{\alpha L, \frac{\gamma}{4} - L + (1+\alpha)\, nR_\mathrm{h}\right\}\right\}, & 0 \leq L - nR_\mathrm{h} < \frac{\gamma}{8}; \\ \exp\left\{-\min\left\{\alpha L, \frac{\gamma}{2} - \sqrt{2\gamma\,(L - nR_\mathrm{h})} + L\right\}\right\}, & \frac{\gamma}{8} \leq L - nR_\mathrm{h} < \frac{\gamma}{2}; \end{cases} \tag{39c}$$

where (39a) follows from (14b); and (39b) follows from (18).

Optimization of (39c) over L, which is essentially the same as that in (15) and (16), concludes the proof. ■

Appendix E
Proofs for Section V-B

Proof of Theorem V.3: Without loss of generality, assume $\sigma^2 = 1$; consequently, $\gamma = E$. Denote

$$\mathscr{I}_w \triangleq \{w, w+M, w+2M, \dots, w + (M_\mathrm{h} - 1)\, M\}$$

and recall that $n = M \cdot M_\mathrm{h}$. Denote further $G_w = \max_{\tau \in \mathscr{I}_w} Z_\tau$. Clearly, the cumulative distribution function (c.d.f.), F_G, and probability distribution function (p.d.f.), f_G, of G_w equal

$$F_G(x) = F_Z^{M_\mathrm{h}}(x) = [1 - Q(x)]^{M_\mathrm{h}}, \tag{40a}$$

$$f_G(x) = M_\mathrm{h} F_Z^{M_\mathrm{h}-1}(x) f_Z(x) = M_\mathrm{h}\,[1 - Q(x)]^{M_\mathrm{h}-1} f_Z(x), \tag{40b}$$

respectively. Hence, the error probability may be bounded as follows.

$$\Pr\left(\hat{W} \neq w\right) = \Pr\left(\max_{t \notin \mathscr{I}_w} Z_t \geq \sqrt{E} + \max_{\tau \in \mathscr{I}_w} Z_\tau\right) \tag{41a}$$

$$= \Pr\left(\max_{t \notin \mathscr{I}_w} Z_t \geq \sqrt{E} + G_w\right) \tag{41b}$$

$$= \int_{-\infty}^{\infty} \Pr\left(\max_{t \notin \mathscr{I}_w} Z_t \geq \sqrt{E} + x\right) f_G(x) dx \tag{41c}$$

$$= \int_{-\infty}^{\infty} \Pr\left(\bigcup_{t \notin \mathscr{I}_w} Z_t \geq \sqrt{E} + x\right) M_\mathrm{h}\,[1 - Q(x)]^{M_\mathrm{h}-1} f_Z(x) dx \tag{41d}$$

$$\leq \int_{-\infty}^{\infty} \sum_{t \notin \mathscr{I}_w} \Pr\left(Z_t \geq \sqrt{E} + x\right) M_\mathrm{h}\,[1 - Q(x)]^{M_\mathrm{h}-1} f_Z(x) dx \tag{41e}$$

$$\leq M_\mathrm{h}^2 M \int_{-\infty}^{\infty} Q\left(\sqrt{E} + x\right) \exp\{-(M_\mathrm{h} - 1)\, Q(x)\} f_Z(x) dx \tag{41f}$$

$$\leq [1 + o(1)]\, M_\mathrm{h}^2 M \int_{\sqrt{2(1-\epsilon) L_\mathrm{h}}}^{\infty} Q\left(\sqrt{E} + x\right) \frac{1}{\sqrt{2\pi\mathrm{e}}} \exp\left\{-\frac{x^2}{2}\right\} dx \tag{41g}$$

$$\leq [1+o(1)]\, M_{\mathrm{h}}^2 M \int_{\sqrt{2(1-\epsilon)L_{\mathrm{h}}}}^{\infty} \exp\left\{-\frac{\left(\sqrt{E}+x\right)^2}{2}\right\}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}dx \tag{41h}$$

$$= [1+o(1)]\,\frac{M_{\mathrm{h}}^2 M}{\sqrt{2}}\exp\left\{-\frac{E}{4}\right\}\int_{\sqrt{2(1-\epsilon)L_{\mathrm{h}}}}^{\infty}\frac{1}{\sqrt{\pi}}\exp\left\{-\left(x+\frac{\sqrt{E}}{2}\right)^2\right\}dx \tag{41i}$$

$$= [1+o(1)]\,\frac{M_{\mathrm{h}}^2 M}{\sqrt{2}}\exp\left\{-\frac{E}{4}\right\}Q\left(\sqrt{2}\left[\sqrt{2(1-\epsilon)L_{\mathrm{h}}}+\frac{\sqrt{E}}{2}\right]\right) \tag{41j}$$

$$\leq [1+o(1)]\exp\{2L_{\mathrm{h}}+L\}\exp\left\{-\frac{E}{4}\right\}\exp\left\{-\left(\sqrt{2(1-\epsilon)L_{\mathrm{h}}}+\frac{\sqrt{E}}{2}\right)^2\right\} \tag{41k}$$

$$= \exp\left\{-\left(\frac{E}{2}+\sqrt{2(1-\epsilon)EL_{\mathrm{h}}}-(1-\epsilon)L_{\mathrm{h}}\right)+o(1)\right\} \tag{41l}$$

where (41b) holds by the definition of G_w; (41c) holds since the entries of $\mathbf{Z}$ are i.i.d.; (41d) follows from (40b) (41e) follows from the union bound; (41f) follows from the inequality $1-x \leq \exp\{-x\}$ and since the entries of $\mathbf{Z}$ are i.i.d. standard Gaussian; (41g) holds for any $\epsilon > 0$, however small, for a sufficiently large n, and follows from

$$(M_{\mathrm{h}}-1)Q(x) \geq (\exp\{L_{\mathrm{h}}\}-1)\,Q\left(\sqrt{2(1-\epsilon)L_{\mathrm{h}}}\right) \tag{42a}$$

$$\geq (\exp\{L_{\mathrm{h}}\}-1)\,\frac{\sqrt{2(1-\epsilon)L_{\mathrm{h}}}}{\sqrt{2\pi}\left(1+\sqrt{2(1-\epsilon)L_{\mathrm{h}}}\right)^2}\exp\{-(L_{\mathrm{h}}-\epsilon)\} \tag{42b}$$

$$\geq \exp\{\epsilon L_{\mathrm{h}}/2\} \tag{42c}$$

for all $x \leq \sqrt{2(1-\epsilon)L_{\mathrm{h}}}$ for any $\epsilon > 0$ for a sufficiently large n, where (42a) follows from the monotonicity of the Q function, and (42b) follows from the lower bound on the Q function [61] $Q(x) \geq \frac{x}{\sqrt{2\pi}(1+x)^2}\exp\left\{-x^2/2\right\}$ for $x > 0$, meaning that

$$\begin{aligned} & M_{\mathrm{h}}^2 M \int_{-\infty}^{\sqrt{2(1-\epsilon)L_{\mathrm{h}}}} Q\left(\sqrt{E}+x\right)\exp\{-(M_{\mathrm{h}}-1)\,Q(x)\}f_Z(x)dx \\ &\quad \leq \exp\{2L_{\mathrm{h}}+L\}\exp\{-\exp\{\epsilon L_{\mathrm{h}}/2\}\}\int_{-\infty}^{\sqrt{2(1-\epsilon)L_{\mathrm{h}}}} Q\left(\sqrt{E}+x\right)f_Z(x)dx \\ &\quad \leq \exp\{2L_{\mathrm{h}}+L\}\exp\{-\exp\{\epsilon L_{\mathrm{h}}/2\}\} \end{aligned}$$

decays *doubly exponentially* with L_{h}; and (41h) and (41k) follow from the Chernoff bound [61] $Q(x) \leq \exp\left\{-x^2/2\right\}$. ∎

Proof of Corollary V.4: The resulting MPαE of Scheme V.2 is bounded for any $u \in [-1/2, 1/2)$ as follows.

$$\mathbb{E}\left[\left|\hat{U} - u\right|^{\alpha}\right] \le \Pr\left(\hat{W} \neq w\right) + \mathbb{E}\left[\left|\hat{U} - u\right|^{\alpha} \middle| \hat{W} = w\right] \tag{43a}$$

$$\le \exp\left\{-\left(\frac{\gamma}{2} + \sqrt{2L_{\mathrm{h}}\gamma} - L\right) + o(1)\right\} + \exp\{-\alpha L\} \tag{43b}$$

$$= \exp\left\{-\min\left\{\frac{\gamma}{2} + \sqrt{2L_{\mathrm{h}}\gamma} - L, \alpha L\right\} + o(1)\right\} \tag{43c}$$

where (43a) follows from (11b), and (43b) follows from Theorem V.3.

Optimization of (39c) is achieved for the value L for which the two operands in the minimum in (43c) are equal; its substitution concludes the proof. ■

APPENDIX F
PROOF OF COROLLARY V.6

$$\epsilon(\alpha) \le \min_{L:\ \exp\{L\} \in \mathbb{N}} \mathbb{E}\left[\left|u - \hat{U}\right|^{\alpha}\right] \tag{44a}$$

$$\le \min_{L:\ \exp\{L\} \in \mathbb{N}} \left\{\mathbb{E}\left[\left|u - \hat{U}\right|^{\alpha} \middle| \hat{W} = w\right] + \Pr\left(\hat{W} \neq w\right)\right\} \tag{44b}$$

$$\le \min_{L:\ \exp\{L\} \in \mathbb{N}} \left\{\exp\{-\alpha L\} + \begin{cases} \exp\left\{-\left(\frac{\gamma}{4} - L\right)\exp\{L_{\mathrm{h}}\}\right\}, & L \le \frac{\gamma}{8}; \\ \exp\left\{-\left(\sqrt{\frac{\gamma}{2}} - \sqrt{L}\right)^2 \exp\{L_{\mathrm{h}}\}\right\}, & \frac{\gamma}{8} < L \le \frac{\gamma}{2}. \end{cases}\right\} \tag{44c}$$

$$\le \min_{L \ge 0} \left\{\exp\{-\alpha L\} + \begin{cases} \exp\left\{-\left(\frac{\gamma}{4} - L\right)\exp\{L_{\mathrm{h}}\} + o(1)\right\}, & L \le \frac{\gamma}{8}; \\ \exp\left\{-\left(\sqrt{\frac{\gamma}{2}} - \sqrt{L}\right)^2 \exp\{L_{\mathrm{h}}\} + o(1)\right\}, & \frac{\gamma}{8} < L \le \frac{\gamma}{2}. \end{cases}\right\} \tag{44d}$$

$$= \begin{cases} \exp\left\{-\frac{\alpha \exp\{L_{\mathrm{h}}\}}{\exp\{L_{\mathrm{h}}\} + \alpha}\frac{\gamma}{4} + o(1)\right\}, & L_{\mathrm{h}} \le \log\alpha; \\ \exp\left\{-\frac{\alpha \exp\{L_{\mathrm{h}}\}}{(\exp\{L_{\mathrm{h}}/2\} + \sqrt{\alpha})^2}\frac{\gamma}{2} + o(1)\right\}, & L_{\mathrm{h}} > \log\alpha; \end{cases} \tag{44e}$$

(44b) follows from (11b), (44c) follows from (21) and since the estimation error given correct decoding of w is bounded by $1/M = \exp\{-L\}$, (44d) follows from relaxing the minimization domain where $o(1)$ is with respect to L, and (44e) follows from solving the optimization with respect to L by equating the exponents where where $o(1)$ is with respect to γ. ■

APPENDIX G

PROOF OF COROLLARY VI.5

Applying the discrete-time reduction of Section VI with $L_{\mathrm{h}} = \mathrm{TR_h}$ and $\gamma = 2\mathrm{TC_0}$ to (23) of Corollary V.6 results in the following upper bound on the MPαE.

$$\epsilon(\alpha) \leq \min_{\mathrm{R_m}\in[0,\mathrm{R_h}]} \begin{cases} \exp\left\{-\mathrm{T}\alpha\left[\mathrm{R_m} + \frac{\exp\{\mathrm{T(R_h-R_m)}\}}{\exp\{\mathrm{T(R_h-R_m)}\}+\alpha}\frac{\mathrm{C_0}}{2}\right] + o(1)\right\}, & \mathrm{R_h} - \mathrm{R_m} \leq \frac{\log\alpha}{\mathrm{T}}; \\ \exp\left\{-\mathrm{T}\alpha\left[\mathrm{R_m} + \frac{\exp\{\mathrm{T(R_h-R_m)}\}}{\left(\exp\left\{\frac{\mathrm{T(R_h-R_m)}}{2}\right\}+\sqrt{\alpha}\right)^2}\mathrm{C_0}\right] + o(1)\right\}, & \mathrm{R_h} - \mathrm{R_m} > \frac{\log\alpha}{\mathrm{T}}; \end{cases} \tag{45}$$

where $o(1)$ is with respect to T, and the minimum is attained at $\mathrm{R_m} = \mathrm{R_h}$, $\mathrm{R_m} = 0$, or $\mathrm{R_m} = \mathrm{R_h} - \log(\alpha)/\mathrm{T}$. Next, we evaluate the bound of (45) for these three values.

- $\underline{\mathrm{R_m} = \mathrm{R_h}}$: The conditions of the first and second cases on the right-hand side of (45) reduce to $\alpha \geq 1$ and $\alpha < 1$, respectively. Therefore, the resulting achievable exponent is

$$\begin{cases} \alpha\left[\mathrm{R_h} + \frac{\mathrm{C_0}}{2(1+\alpha)}\right], & \alpha \geq 1; \\ \alpha\left[\mathrm{R_h} + \frac{\mathrm{C_0}}{(1+\sqrt{\alpha})^2}\right], & \alpha < 1. \end{cases}$$

- $\underline{\mathrm{R_m} = 0}$: Now, $\mathrm{R_h} - \mathrm{R_m} = \mathrm{R_h} > \frac{\log\alpha}{\mathrm{T}}$ for a sufficiently large T. Therefore, the resulting achievable exponent is $\alpha\mathrm{C_0}$, since

$$\lim_{\mathrm{T}\to\infty} \frac{\exp\{\mathrm{T}(\mathrm{R_h} - \mathrm{R_m})\}}{\left(\exp\left\{\frac{\mathrm{T(R_h-R_m)}}{2}\right\} + \sqrt{\alpha}\right)^2} = 1$$

  for any fixed $\alpha > 0$.

- $\underline{\mathrm{R_m} = \mathrm{R_h} - \log(\alpha)/\mathrm{T}}$: In this case both of the cases on the right-hand side of (45) yield the same achievable error exponent $\alpha(\mathrm{R_h} + \mathrm{C_0}/4)$.

Noting that $2(1+\alpha), (1+\sqrt{\alpha})^2 \geq 4$ for $\alpha \geq 1$ and that $2(1+\alpha), (1+\sqrt{\alpha})^2 < 4$ for $\alpha < 1$ concludes the proof.

REFERENCES

[1] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[2] V. A. Kotel'nikov, *The Theory of Optimum Noise Immunity*. New York: McGraw-Hill, 1959.

[3] P. Elias, "Channel capacity without coding," in *Proceedings of the IRE*, vol. 45, no. 3, Jan. 1957, pp. 381–381.

[4] T. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Transactions on Information Theory*, vol. 11, pp. 558–567, 1965.

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, 1968.

[6] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York, NY, USA: John Wiley & Sons, 1965.

[7] L. P. Seidman, “Performance limitations and error calculations for parameter estimation,” *Proceedings of the IEEE*, vol. 58, no. 5, pp. 644–652, 1970.

[8] J. Ziv and M. Zakai, “Some lower bounds on signal parameter estimation,” *IEEE Transactions on Information Theory*, vol. 15, no. 3, pp. 386–391, 1969.

[9] D. Chazan, M. Zakai, and J. Ziv, “Improved lower bounds on signal parameter estimation,” *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 90–93, 1975.

[10] E. Zehavi, “Estimation of time of arrival for rectangular pulses,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 20, no. 6, pp. 742–747, 1984.

[11] M. Burnashev, “A new lower bound for the a-mean error of parameter transmission over the white Gaussian channel,” *IEEE Transactions on Information Theory*, vol. 30, no. 1, pp. 23–34, 1984.

[12] ——, “A new lower bound for the a-mean error of parameter transmission over the white Gaussian channel,” *IEEE Transactions on Information Theory*, vol. 30, no. 1, pp. 23–34, 1984.

[13] M. Burnashev, “Minimum attainable mean-square error in transmission of a parameter over a channel with white Gaussian noise,” (in Russian) *Problemy Pered. Inf. (Problems of Inf. Trans.)*, vol. 21, 10 1985.

[14] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

[15] A. Weiss and E. Weinstein, “A lower bound on the mean-square error in random parameter estimation (corresp.),” *IEEE Transactions on Information Theory*, vol. 31, no. 5, pp. 680–682, 1985.

[16] N. Merhav, “On optimum parameter modulation–estimation from a large deviations perspective,” *IEEE Transactions on Information Theory*, vol. 58, no. 12, pp. 7215–7225, 2012.

[17] N. Weinberger and N. Merhav, “Lower bounds on parameter modulation–estimation under bandwidth constraints,” *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 3854–3874, 2017.

[18] S. Y. Chung, “On the construction of some capacity-approaching coding schemes,” Ph.D. dissertation, Dept. EECS, Massachusetts Institute of Technology, Cambridge, MA, USA, 2000.

[19] F. Hekland, P. A. Floor, and T. A. Ramstad, “Shannon–Kotel’nikov mappings in joint source–channel coding,” *IEEE Transactions on Communications*, vol. 57, no. 1, pp. 94–105, Jan. 2009.

[20] Y. Hu, J. Garcia-Frias, and M. Lamarca, “Analog joint source–channel coding using non-linear curves and MMSE decoding,” *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3016–3026, Nov. 2011.

[21] N. Merhav, “Tradeoffs between weak-noise estimation performance and outage exponents in nonlinear modulation,” *IEEE Transactions on Information Theory*, vol. 65, no. 8, pp. 5189–5196, 2019.

[22] ——, “Weak–noise modulation–estimation of vector parameters,” *IEEE Transactions on Information Theory*, vol. 66, no. 5, pp. 3268–3276, 2020.

[23] ——, “Tradeoffs between weak-noise estimation performance and outage exponents in nonlinear modulation,” *IEEE Transactions on Information Theory*, vol. 65, no. 8, pp. 5189–5196, 2019.

[24] S. Tridenski, R. Zamir, and A. Ingber, “The Ziv–Zakai–Rényi bound for joint source–channel coding,” *IEEE Transactions on Information Theory*, vol. 61, no. 8, pp. 4293–4315, Aug. 2015.

[25] E. Köken, D. Gündüz, and E. Tuncel, “Energy–distortion exponents in lossy transmission of Gaussian sources over Gaussian channels,” *IEEE Transactions on Information Theory*, vol. PP, pp. 1–1, 12 2016.

[26] E. Köken and E. Tuncel, “On minimum energy for robust Gaussian joint source-channel coding with a distortion-noise profile,” in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 1668–1672.

[27] C. Sevinç and E. Tuncel, “On asymptotic analysis of energy–distortion tradeoff for low-delay transmission over Gaussian channels,” *IEEE Transactions on Communications*, vol. 69, no. 7, pp. 4448–4460, 2021.

[28] O. Lev and A. Khina, “Energy-limited joint source–channel coding via analog pulse position modulation,” *IEEE Transactions on Communications*, vol. 70, no. 8, pp. 5140–5150, Aug. 2022.

[29] ——, “Universal joint source–channel coding under an input energy constraint,” in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Espoo, Finland, June 2022, pp. 240–245.

[30] M. Baniasadi, E. Köken, and E. Tuncel, “Minimum energy analysis for robust Gaussian joint source–channel coding with a distortion–noise profile,” *IEEE Transactions on Information Theory*, vol. 68, no. 12, pp. 7702–7713, Dec. 2022.

[31] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[32] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[33] A. Wyner, “On source coding with side information at the decoder,” *IEEE Transactions on Information Theory*, vol. IT-21, pp. 294–300, May 1975.

[34] R. Ahlswede and J. Korner, “Source coding with side information and a converse for the degraded broadcast channel,” *IEEE Transactions on Information Theory*, vol. 21, pp. 629–637, November 1975.

[35] R. Ahlswede and T. Han, “On source coding with side information via a multiple-access channel and related problems in multi-user information theory,” *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 396–412, May 1983.

[36] C. Heegard and A. El Gamal, “On the capacity of computer memory with defects,” *IEEE Transactions on Information Theory*, vol. IT-29, pp. 731–739, Sep. 1983.

[37] A. Rosenzweig, Y. Steinberg, and S. Shamai, “On channels with partial channel state information at the transmitter,” *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1817–1830, 2005.

[38] A. Lapidoth and P. Narayan, “Reliable communication under channel uncertainty,” *IEEE Transactions on Information Theory*, vol. 44, pp. 2148–2177, 1998.

[39] G. Keshet, Y. Steinberg, and N. Merhav, “Channel coding in the presence of side information: Subject review,” *Foundations and Trends® in Communications and Information Theory*, vol. 4, no. 6, pp. 445–586, 2008.

[40] G. Marti, “Channels with a helper,” Master’s thesis, Signal and Information Processing Laboratory (ISI), Swiss Federal Institute of Technology Zurich (ETH), Sep. 2019.

[41] A. Lapidoth and G. Marti, “Encoder-assisted communications over additive noise channels,” *IEEE Transactions on Information Theory*, vol. 66, no. 11, pp. 6607–6616, 2020.

[42] S. I. Bross, A. Lapidoth, and G. Marti, “Decoder-assisted communications over additive noise channels,” *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4150–4161, 2020.

[43] N. Merhav, “On error exponents of encoder-assisted communication systems,” *IEEE Transactions on Information Theory*, vol. 67, no. 11, pp. 7019–7029, 2021.

[44] L. Wang and A. Lapidoth, “State-dependent DMC with a helper,” *CoRR*, 2023. [Online]. Available: http://arxiv.org/abs/2304.12687

[45] L. D. Brown and R. C. Liu, “Bounds on the Bayes and minimax risk for signal parameter estimation,” *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1386–1394, 1993.

[46] K. L. Bell, “Performance bounds in parameter estimation with application to bearing estimation,” Ph.D. dissertation, Department of Electrical and Computer Engineering, George Mason University, Fairfax, VA, USA, 1995.

[47] T. Liu and P. Viswanath, “Opportunistic orthogonal writing on dirty paper,” *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 1828–1846, April 2006.

[48] B. Nakiboğlu, “The sphere packing bound for memoryless channels,” (in Russian) *Problemy Pered. Inf. (Problems of Inf. Trans.)*, vol. 56, pp. 201–244, 2020.

[49] S. Tridenski and A. Somekh-Baruch, “The method of types for the AWGN channel,” *arXiv preprint arXiv:2307.13322*, July 2023.

[50] N. Merhav, "Exponential error bounds on parameter modulation–estimation for discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 832–841, 2013.

[51] A. Wyner, "On the probability of error for communication in white Gaussian noise," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 86–90, January 1967.

[52] M. Steiner, "The strong simplex conjecture is false," *IEEE Transactions on Information Theory*, vol. 40, no. 3, pp. 721–731, May 1994.

[53] B. Dunbridge, "Asymmetric signal design for the coherent Gaussian channel," *IEEE Transactions on Information Theory*, vol. 13, no. 3, pp. 422–431, July 1967.

[54] ——, "Optimal signal design for the coherent Gaussian channel," Ph.D. dissertation, School of Engineering, University of Southern California, January 1966.

[55] M. V. Burnashev, "Information transmission over a discrete channel with feedback. Random transmission time (in Russian)." *Problems of Information Transmission*, vol. 12, pp. 10–30, 1976.

[56] A. Tchamkerten and E. Telatar, "A feedback strategy for binary symmetric channels," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Lausanne, Switzerland, 2002, p. 362.

[57] A. Tchamkerten and I. E. Telatar, "Variable length coding over an unknown channel," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2126–2145, 2006.

[58] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Feedback in the non-asymptotic regime," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4903–4925, 2011.

[59] Y. Oohama, "The optimal exponent function for the additive white Gaussian noise channel at rates above the capacity," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Aachen, Germany, June 2017, pp. 1053–1057.

[60] ——, "Equivalence of two exponent functions for discrete memoryless channels with input cost at rates above the capacity," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 101, no. 12, pp. 2199–2204, December 2018.

[61] P. Borjesson and C.-E. Sundberg, "Simple approximations of the error function $Q(x)$ for communications applications," *IEEE Transactions on Communications*, vol. 27, no. 3, pp. 639–643, Mar. 1979.