Transmitting k samples over the Gaussian channel: Energy-distortion tradeoff

Transmitting k samples over the Gaussian channel: energy-distortion tradeoff

Victoria Kostina
California Institute of Technology
Pasadena, CA 91125
vkostina@caltech.edu

Yury Polyanskiy
Massachusetts Institute of Technology
Cambridge, MA 02139
yp@mit.edu

Sergio Verdú
Princeton University
Princeton, NJ 08544
verdu@princeton.edu

***Abstract*—We investigate the minimum transmitted energy required to reproduce k source samples with a given fidelity after transmission over a memoryless Gaussian channel. In particular, we analyze the reduction in transmitted energy that accrues thanks to the availability of noiseless feedback. Allowing a nonvanishing excess distortion probability ϵ boosts the asymptotic fundamental limit by a factor of $1-\epsilon$, with or without feedback. If feedback is available, achieving guaranteed distortion with finite average energy is possible.**

***Index Terms*—Joint source-channel coding, lossy compression, single-shot method, finite-blocklength regime, rate-distortion theory, feedback, memoryless channels, Gaussian channels, energy-distortion tradeoff, Shannon theory.**

I. Introduction

Disposing of the restriction on the number of channel uses per source sample, we limit the total available transmitter energy E and we study the tradeoff between the source dimension k, the total energy E and the fidelity of reproduction achievable in the transmission over an AWGN channel, with and without feedback.

The problem of transmitting a message with minimum energy was posed by Shannon [1], who showed that the minimum energy per information bit compatible with vanishing block error probability converges to $N_0 \log_e 2$ as the number of information bits goes to infinity, where $\frac{N_0}{2}$ is the noise power per degree of freedom. In other words, if the source produces equiprobable binary strings of length k, the minimum energy per information bit compatible with vanishing block error probability converges to [1]

$$\frac{E^\star(k,0,\epsilon)}{kN_0} \to \log_e 2 = -1.59 \text{ dB} \tag{1}$$

as $k \to \infty$, $\epsilon \to 0$. The fundamental limit in (1) holds regardless of whether feedback is available. Moreover, this fundamental limit is known to be the same regardless of whether the channel is subject to fading or whether the receiver is coherent or not [2]. Polyanskiy et al. refined (1) as [3, Theorem 3]

$$E^\star(k,0,\epsilon)\frac{\log e}{N_0} = k + \sqrt{2k\log e}\,Q^{-1}(\epsilon) - \frac{1}{2}\log k + O(1) \tag{2}$$

where Q is the standard normal complementary cumulative distribution function, for transmission without feedback, and as [3, Theorem 8]

$$E^\star_f(k,0,\epsilon)\frac{\log e}{N_0} = (1-\epsilon)k + O(1) \tag{3}$$

for transmission with feedback. Moreover, [3, Theorem 7] (see also (13)) shows that in fact

$$E^\star(k,0,0)\log e = kN_0, \tag{4}$$

i.e. in the presence of full noiseless feedback, Shannon's limit (1) can be achieved with equality already at $k=1$ and $\epsilon = 0$.

For the finite blocklength behavior of energy per bit in fading channels, see [4].

For the transmission of a memoryless source over the AWGN channel under an average distortion criterion, Jain et al. [5, Theorem 1] pointed out that as $k \to \infty$,

$$\frac{E^\star(k,d)}{k}\frac{\log e}{N_0} \to R(d). \tag{5}$$

where $R(d)$ is the source rate-distortion function. Note that (5) still holds even if noiseless feedback is available.

Unlike Polyanskiy et al. [3], we allow analog sources and arbitrary distortion criteria, and unlike Jain et al. [5], we are interested in a nonasymptotic analysis of the minimum energy per sample.

We treat several scenarios that differ in how the distortion is evaluated and in whether feedback is available.

The main results of this paper are the following:

- Under *average distortion* constraint, $\mathbb{E}\left[\mathsf{d}(S^k,\hat{S}^k)\right] \le d$, the total minimum energy required to transmit k source samples over an AWGN channel *with feedback* satisfies

$$E^\star_f(k,d)\cdot\frac{\log e}{N_0} = kR(d) + O(\log k)\,. \tag{6}$$

- Under *excess distortion* constraint, $\mathbb{P}[\mathsf{d}(S^k,\hat{S}^k) > d] \le \epsilon$, the total minimum energy required to transmit k memoryless source samples over an AWGN channel *without feedback* satisfies

$$E^\star(k,d,\epsilon)\cdot\frac{\log e}{N_0} = kR(d) + \sqrt{k\left(2R(d)\log e + \mathcal{V}(d)\right)}Q^{-1}(\epsilon) + O(\log k) \tag{7}$$

where $\mathcal{V}(d)$ is the rate-dispersion function of the source.

This work was supported in part by the National Science Foundation (NSF) under Grant CCF-1016625, by the NSF CAREER award under Grant CCF-1253205, and by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under Grant CCF-0939370.

- Under *excess distortion* constraint, the total minimum average energy[1] required to transmit k memoryless source samples over an AWGN channel *with feedback* satisfies

$$E_f^\star(k,d,\epsilon)\cdot\frac{\log e}{N_0}=kR(d)(1-\epsilon)-\sqrt{\frac{k\mathcal{V}(d)}{2\pi}}e^{-\frac{(Q^{-1}(\epsilon))^2}{2}}+O(\log k) \tag{8}$$

Particularizing (8) to $\epsilon = 0$ also covers the case of guaranteed distortion. The first term in the expansion (8) can be achieved even without feedback, as long as $\epsilon > 0$ and the power constraint is understood as averaged over the codebook.

Section II focuses on energy-probability of error tradeoff in the transmission of a single random variable over an AWGN channel. Section III introduces the distortion measure and leverages the results of Section II to study finite blocklength energy-distortion tradeoff.

II. Energy-limited feedback codes for non-equiprobable messages

In this section, we study the transmission of a discrete random variable over an AWGN channel under an energy constraint. In particular, we would like to know how much information can be pushed through the channel, if a total of E units of energy is available to accomplish the task. Formally, the codes studied in this section are defined as follows.

Definition 1. *An energy-limited code without feedback for the transmission of a discrete random variable W taking values in $\mathcal{W}$ over an AWGN channel is defined by:*

1) *A sequence of encoders $\mathsf{f}_n\colon \mathcal{W}\mapsto\mathcal{A}$, defining the channel inputs*

$$X_n=\mathsf{f}_n(W) \tag{9}$$

satisfying

$$\mathbb{P}\left[\sum_{j=1}^{\infty}X_j^2\le E\right]=1 \tag{10}$$

2) *A decoder $\mathsf{g}\colon\mathcal{B}^\infty\mapsto\mathcal{W}$.*

Definition 2. *An energy-limited feedback code for the transmission of a random variable W taking values in $\mathcal{W}$ over an AWGN channel is defined by:*

1) *A sequence of encoders $\mathsf{f}_n\colon\mathcal{W}\times\mathcal{B}^{n-1}\mapsto\mathcal{A}$, defining the channel inputs*

$$X_n=\mathsf{f}_n\left(W,Y^{n-1}\right) \tag{11}$$

satisfying

$$\sum_{j=1}^{\infty}\mathbb{E}\left[X_j^2\right]\le E \tag{12}$$

2) *A decoder $\mathsf{g}\colon\mathcal{B}^\infty\mapsto\mathcal{W}$.*

[1]The energy constraint in (7) is understood on a per-codeword basis. The energy constraint in (8) is understood as average over the codebook.

Bit number	Sequence of time slots				
1	1	2	4	7	$\cdots$
2	3	5	8	$\cdots$	
3	6	9	$\cdots$		
$\vdots$					
$\ell(W)$					

Fig. 1. Illustration of the diagonal numbering of channel uses in Theorem 1.

An (E,ϵ) code for the transmission of random variable W over the Gaussian channel is a code with energy bounded by E and $\mathbb{P}\left[W\neq\widehat{W}\right]\le\epsilon$.

As in the setting of Shannon [1], definitions 1–2 do not impose any restrictions on the number of degrees of freedom n, restricting instead the total available energy.

Recently, Polyanskiy et al. [3, Theorem 7] showed a dynamic programming algorithm for the error-free transmission of a single bit over an AWGN channel with feedback that attains *exactly* Shannon's optimal energy per bit tradeoff

$$E=N_0\log_e 2. \tag{13}$$

Our first non-asymptotic achievability result leverages that algorithm to transmit error-free a binary representation of a random variable over the AWGN channel by means of a variable-length separate compression/transmission scheme.

Theorem 1. *There exists a zero-error feedback code for the transmission of a random variable W over the AWGN channel with energy*

$$E<N_0\log_e 2\,(H(W)+1) \tag{14}$$

Proof. The encoder converts the source into a variable-length string using a Huffman code, so that the codebook is prefix-free and the expectation of the encoded length $\ell(W)$ is bounded as

$$\mathbb{E}\left[\ell(W)\right]<H(W)+1\,. \tag{15}$$

Next each bit (out of $\ell(W)$) is transmitted at the optimal energy per bit tradeoff $N_0\log_e 2$ using the zero-error feedback scheme in [3, Theorem 7]. Transmissions corresponding to different bits are interleaved diagonally (see Fig. 1): the first bit is transmitted in time slots $1,2,4,7,11,\ldots$, the second one in $3,5,8,12,\ldots$, and so on. The channel encoder is silent at those indices allocated to source bits $\ell(W)+1,\ell(W)+2,\ldots$ For example, if the codeword has length 2 nothing is transmitted in time slots $6,9,13,\ldots$. The receiver decodes the first transmitted bit focusing on the time slots $1,2,4,7,11,\ldots$ It proceeds successively with the second bit, etc., until it forms a codeword of the Huffman code, at which point it halts. Thus, it does not need to examine the outputs of the time slots corresponding to information bits that were not transmitted, and in which the encoder was silent.

Since the scheme spends $N_0\log_e 2$ energy per bit, The total energy to transmit the codeword representing W is

$$\ell(W)N_0\log_e 2. \tag{16}$$

Taking the expectation of (16) over W and applying (15), (14) follows. $\square$

For concreteness, in the proof of Theorem 1 we have referred to the degrees of freedom as time slots. In practice, decoding in finite time is feasible if the channel is not bandlimited.

In the converse direction, achieving zero error probability requires that

$$H(W) \leq \frac{E}{N_0} \log e. \tag{17}$$

Indeed, due to the zero-error requirement and data processing, $H(W) = I(W; \mathsf{g}(Y^\infty)) \leq I(X^\infty; Y^\infty)$, which in turn is bounded by the right side of (17) [1].

Our next achievability result studies the performance of a variable-length separated scheme.

Theorem 2. *Denote*

$$\varepsilon(E, m) \triangleq 1 - \frac{1}{\sqrt{\pi N_0}} \int_{-\infty}^{\infty} \left(1 - Q\left(\frac{x + \sqrt{E}}{\sqrt{\frac{N_0}{2}}}\right)\right)^{m-1} e^{-\frac{x^2}{N_0}} dx. \tag{18}$$

Assume that W takes values in $\{1, 2, \ldots, M\}$. There exists an (E, ϵ) non-feedback code for the transmission of random variable W over an AWGN channel without feedback such that

$$\epsilon \leq \min_{\substack{E_1 > 0, E_2 > 0: \\ E_1 + E_2 \leq E}} \left\{\mathbb{E}\left[\varepsilon\left(E_1, W\right)\right] + \varepsilon(E_2, \lfloor \log_2 M \rfloor + 1)\right\} \tag{19}$$

where $\ell(W)$ denotes the length of W.

Proof. Assume that the outcomes of W are ordered in decreasing probabilities. Consider the following variable-length separated achievability scheme: the source outcome m is first losslessly represented as a binary string of length $\lfloor \log_2 m \rfloor$ by assigning it to m-th binary string in $\{\emptyset, 0, 1, 00, 01, \ldots\}$ (the most likely outcome is represented by the empty string). Then, all binary strings are grouped according to their encoded lengths. A channel codebook is generated for each group of sequences. The encoded length is sent over the channel with high reliability, so the decoder almost never makes an error in determining that length. Then the encoder makes an ML decision only between sequences of that length. A formal description and an error analysis follow.

Codebook: the collection of $M + \lfloor \log M \rfloor + 1$ codewords

$$\mathbf{c}_j = \sqrt{E_1}\, \mathbf{e}_j, \; j = 1, 2, \ldots, M \tag{20}$$

$$\mathbf{c}_j = \sqrt{E_2}\, \mathbf{e}_j, \; j = M + 1, M + \lfloor \log_2 M \rfloor + 1 \tag{21}$$

where $\{\mathbf{e}_j, \; j = 1, 2, \ldots\}$ is an orthonormal basis of $L_2(\mathbb{R}^\infty)$.

Encoder: The encoder sends the pair $(m, \lfloor \log_2 m \rfloor)$ by transmitting $\mathbf{c}_m + \mathbf{c}_{m + \lfloor \log_2 m \rfloor + 1}$.

Decoder: Having received the infinite string corrupted by i.i.d. Gaussian noise $\mathbf{z}$, the decoder first (reliably) decides between $\lfloor \log_2 M \rfloor + 1$ possible values of $\lfloor \log_2 m \rfloor$ based on the minimum distance:

$$\hat{\ell} \triangleq \operatorname{argmin} \|\mathbf{z} - \mathbf{c}_j\|, \; j = M+1, \ldots, M + \lfloor \log_2 M \rfloor + 1 \tag{22}$$

As shown in [6, p. 258]), [7], [3, Theorem 3], the probability of error of such a decision is given by $\varepsilon(E, \lfloor \log_2 M \rfloor + 1)$. This accounts for the second term in (19). The decoder then decides between $2^{\hat{\ell}}$ messages[2] j with $\lfloor \log j \rfloor = \hat{\ell}$:

$$\hat{\mathbf{c}} \triangleq \operatorname{argmin} \|\mathbf{z} - \mathbf{c}_j\|, \; j = 2^{\hat{\ell}}, \ldots, \min\{2^{\hat{\ell}+1} - 1, M\} \tag{23}$$

The probability of error of this decision rule is similarly upper bounded by $\varepsilon\left(E, m\right)$, provided that the value of $\lfloor \log_2 m \rfloor$ was decoded correctly: $\hat{\ell} = \lfloor \log_2 m \rfloor$. Since $2^{\lfloor \log_2 m \rfloor} \leq m$, this accounts for the first term in (19). $\square$

Normally, one would choose $1 \ll E_2 \ll E_1$ so that the second term in (19), which corresponds to the probability of decoding the length incorrectly, is negligible compared to the first term, and the total energy $E \approx E_1$. Moreover, if W takes values in a countably infinite alphabet, one can truncate it so that the tail is negligible with respect to the first term in (19). To ease the evaluation of the first term in (19), one might use $i \leq \frac{1}{P_W(i)}$. In the equiprobable case, this weakening leads to $\mathbb{E}\left[\varepsilon\left(E_1, W\right)\right] \leq \varepsilon\left(E_1, M\right)$.

If the power constraint is average rather than maximal, a straightforward extension of Theorem 2 ensures the existence of an (E, ϵ) code (average power constraint) for the AWGN channel with

$$\epsilon \leq \mathbb{E}\left[\varepsilon\left(E_1(\lfloor \log_2 W \rfloor), W\right)\right] + \varepsilon(E_2, \lfloor \log M \rfloor + 1), \tag{24}$$

where $E_1 \colon \{0, 1, \ldots, \lfloor \log_2 M \rfloor\} \mapsto \mathbb{R}_+$ and $E_2 \in \mathbb{R}_+$ are such that

$$\mathbb{E}\left[E_1(\lfloor \log_2 W \rfloor)\right] + E_2 \leq E. \tag{25}$$

III. Asymptotic expansions of the energy-distortion tradeoff

A. Problem setup

This section focuses on the energy-distortion tradeoff in the JSCC problem. Like in Section II, we limit the total available transmitter energy E without any restriction on the (average) number of channel uses per source sample. Unlike Section II, we allow general (not neccesarily discrete) sources, and we study the tradeoff between the source dimension k, the total energy E and the fidelity of reproduction. Thus, we identify the minimum energy compatible with target distortion without any restriction on the time-bandwidth product (number of degrees of freedom). We consider both the average and the excess distortion criteria.

Formally, we let the source be a k-dimensional vector S^k in the alphabet $\mathcal{S}^k$. A (k, E, d, ϵ) energy-limited code is an

[2]More precisely, $2^{\hat{\ell}}$ messages if $\hat{\ell} \leq \lfloor \log_2 M \rfloor - 1$ and $M - 2^{\lfloor \log_2 M \rfloor} + 1 \leq 2^{\lfloor \log_2 M \rfloor}$ messages if $\hat{\ell} = \lfloor \log_2 M \rfloor$.

energy-limited code for $(S^k, \mathcal{S}^k)$ with total energy E and probability $\leq \epsilon$ of exceeding d:

$$\mathbb{P}[\mathsf{d}(S^k, \hat{S}^k) \leq d] \leq \epsilon, \tag{26}$$

where $\mathsf{d}\colon \mathcal{S}^k \times \hat{\mathcal{S}}^k \mapsto [0, +\infty]$ is the distortion measure. Similarly, a (k, E, d) energy-limited code is an energy-limited code for $(S^k, \mathcal{S}^k)$ with total energy E and average distortion not exceeding d:

$$\mathbb{E}\left[\mathsf{d}(S^k, \hat{S}^k)\right] \leq d. \tag{27}$$

The goal of Section III is to characterize the minimum energy required to transmit k source samples at a given fidelity, i.e. to characterize the following fundamental limits:

$$E_f^\star(k, d) \triangleq \left\{\inf E \colon \exists \text{ a } (k, E, d) \text{ feedback code}\right\}, \tag{28}$$

$$E_f^\star(k, d, \epsilon) \triangleq \left\{\inf E \colon \exists \text{ a } (k, E, d, \epsilon) \text{ feedback code}\right\} \tag{29}$$

as well as the corresponding limits $E^\star(k, d)$ and $E^\star(k, d, \epsilon)$ of the energy-limited non-feedback codes.

B. Regularity assumptions on the source

We assume that the source, together with its distortion measure, satisfies the following assumptions:

A1 The source $\{S_i\}$ is stationary and memoryless, $P_{S^k} = P_{\mathsf{S}} \times \ldots \times P_{\mathsf{S}}$.

A2 The distortion measure is separable, $\mathsf{d}(s^k, z^k) = \frac{1}{k}\sum_{i=1}^k \mathsf{d}(s_i, z_i)$.

A3 The distortion level satisfies $d_{\min} < d < d_{\max}$, where $d_{\min}$ is the infimum of values at which the minimal mutual information quantity $\mathbb{R}_S(d)$ is finite, and $d_{\max} = \inf_{\mathsf{z} \in \widehat{\mathcal{M}}} \mathbb{E}\left[\mathsf{d}(\mathsf{S}, \mathsf{z})\right]$, where the expectation is with respect to the unconditional distribution of S.

A4 The rate-distotortion function is achieved by a unique $P_{\mathsf{Z}^\star|\mathsf{S}}$: $R(d) = I(\mathsf{S}; \mathsf{Z}^\star)$.

A5 $\mathbb{E}\left[\mathsf{d}^{12}(\mathsf{S}, \mathsf{Z}^\star)\right] < \infty$ where the expectation is with respect to $P_{\mathsf{S}} \times P_{\mathsf{Z}^\star}$.

We showed in [8] that under assumptions A1–A5 for all $0 \leq \epsilon \leq 1$

$$\left.\begin{array}{l}\mathbb{R}_{S^k}(d, \epsilon) \\ H_{d,\epsilon}(S^k)\end{array}\right\} = (1-\epsilon)kR(d) - \sqrt{\frac{k\mathcal{V}(d)}{2\pi}}e^{-\frac{(Q^{-1}(\epsilon))^2}{2}} + O\left(\log k\right). \tag{30}$$

where

$$\mathbb{R}_{S^k}(d, \epsilon) \triangleq \min_{\substack{P_{Z^k|S^k}\colon \mathcal{S}^k \mapsto \hat{\mathcal{S}}^k \colon \\ \mathbb{P}\left[\mathsf{d}(S^k, Z^k) > d\right] \leq \epsilon}} I(S^k; Z^k), \tag{31}$$

$$H_{d,\epsilon}(S^k) \triangleq \min_{\substack{\mathsf{c}\colon \mathcal{S}^k \mapsto \hat{\mathcal{S}}^k \colon \\ \mathbb{P}\left[\mathsf{d}(S^k, \mathsf{c}(S^k)) > d\right] \leq \epsilon}} H(\mathsf{c}(S^k)). \tag{32}$$

C. Energy-limited feedback codes

Our first result in this section is a refinement of (5).

Theorem 3. *Let the source and its distortion measure satisfy assumptions A1–A5. The minimum energy required to transmit k source symbols with average distortion $\leq d$ over an AWGN channel with feedback satisfies*

$$E_f^\star(k, d) \cdot \frac{\log e}{N_0} = kR(d) + O(\log k) \tag{33}$$

Proof. Achievability. The expansion in (33) is achieved by the following separated source/channel scheme. For the source code, we use the code of Yang and Zhang [9] (abstract alphabet) that compresses the source down to M representation points with average distortion d such that

$$\log M = kR(d) + O(\log k). \tag{34}$$

For the channel code, we transmit the binary representation of M error-free using the optimal scheme of Polyanskiy et al. [3, Theorem 7], so that

$$\log M = \frac{E}{N_0} \log e. \tag{35}$$

Converse. By data processing, similar to (17),

$$kR(d) \leq \frac{E}{N_0} \log e. \tag{36}$$

□

For the transmission of a Gaussian sample over the feedback AWGN channel, the Schalkwijk-Bluestein scheme [10], [11] attains exactly the fundamental limit

$$E_f^\star(k, d) \cdot \frac{\log e}{N_0} = kR(d) \tag{37}$$

for $k = 1$. For k Gaussian samples, transmitting the Schalkwijk-Bluestein codewords corresponding to i-th source sample in time slots $i, k+i, 2k+i, \ldots$ attains (37) exactly for all $k = 1, 2, \ldots$.

Theorem 4. *In the transmission of a source satisfying the assumptions A1–A5 over an AWGN channel with feedback, the minimum average energy required for the transmission of k source samples under the requirement that the probability of exceeding distortion d is no greater than $0 \leq \epsilon < 1$ satisfies, as $k \to \infty$,*

$$E_f^\star\left(k, d, \epsilon\right) \frac{\log e}{N_0} = (1-\epsilon)kR(d) - \sqrt{\frac{k\mathcal{V}(d)}{2\pi}}e^{-\frac{(Q^{-1}(\epsilon))^2}{2}} + O\left(\log k\right) \tag{38}$$

Proof. Achievability. Pair a lossy compressor $S^k \to W$ with excess-distortion probability ϵ and $H(W) = H_{d,\epsilon}(S^k)$ with the achievability scheme in Theorem 1 and apply (14) and (30).

Converse. Again, the converse result follows proceeding as in (17), invoking (30). □

Theorem 4 demonstrates that allowing feedback and average power constraint reduces the asymptotically achievable minimum energy per sample by a factor of $1 - \epsilon$. That limit is approached from below rather than from above, i.e. finite blocklength helps.

D. Energy-limited non-feedback codes

Our next result generalizes [3, Theorem 3]. Loosely speaking, it shows that the energy E, probability of error ϵ and distortion d of the best non-feedback code satisfy

$$\frac{E}{N_0}\log e - kR(d) \approx \sqrt{k\mathcal{V}(d) + \frac{2E}{N_0}\log e}\cdot Q^{-1}(\epsilon)$$

i.e. source and channel dispersions add up, as in the usual (non-feedback) joint source-channel coding problem [12], [13]. More precisely, we have the following:

Theorem 5 (Asymptotics, no feedback)**.** *In the transmission of a stationary memoryless source (satisfying the assumptions A1–A5) over the AWGN channel, the minimum energy necessary for achieving probability* $0 < \epsilon < 1$ *of exceeding distortion* d *satisfies, as* $k \to \infty$*,*

$$E^\star(k,d,\epsilon)\frac{\log e}{N_0} = kR(d) + \sqrt{k\left(2R(d)\log e + \mathcal{V}(d)\right)}Q^{-1}(\epsilon) + O(\log k) \quad (39)$$

Proof. ArXiv preprint [14]. □

If the maximal power constraint in (10) is relaxed to (12), then $E_a^\star(k,d,\epsilon)$, the minimum average power required for transmitting k source samples over an AWGN channel with the probability of exceeding distortion d smaller than or equal to $0<\epsilon<1$ satisfies, under assumptions A1–A5:

$$E_a^\star(k,d,\epsilon)\frac{\log e}{N_0} = R(d)(1-\epsilon)k + O\left(\sqrt{k\log k}\right), \quad (40)$$

i.e. the asymptotically achievable minimum energy per sample is reduced a factor of $1-\epsilon$ if a maximal power constraint is relaxed to an average one.

Proof of (40)*.* Observe that Theorem 4 ensures that a smaller average energy than that in (40) is not attainable even with full noiseless feedback. In the achievability direction, let $(\mathsf{f}^\star, \mathsf{g}^\star)$ be the optimal variable-length source code achieving the probability of exceeding d equal to ϵ' (see [8, Section III.B]). Denote by $\ell(\mathsf{f}^\star(s))$ the length of $\mathsf{f}^\star(s)$. Let M be the size of that code. Set the energy to transmit the codeword of length $\ell(\mathsf{f}^\star(S^k))$ to

$$\ell(\mathsf{f}^\star(S^k))N_0\log_e 2 + \sqrt{k\log k} \quad (41)$$

As shown in [8],

$$\mathbb{E}\left[\ell(\mathsf{f}^\star(S))\right] = (1-\epsilon')kR(d) - \sqrt{\frac{k\mathcal{V}(d)}{2\pi}}e^{-\frac{(Q^{-1}(\epsilon'))^2}{2}} + O(\log k) \quad (42)$$

Choosing $\epsilon' = \epsilon - \frac{a}{\sqrt{k}}$ for some a, we conclude that indeed the average energy satisfies (40). Moreover, [3, Theorem 3] implies that the expression inside the expectation in (24) is $O\left(\frac{1}{\sqrt{k}}\right)$. It follows that for a large enough a, the excess distortion probability is bounded by ϵ, and the proof is complete. □

IV. Conclusions

1) In the absence of feedback, source and channel dispersions add up.
2) With feedback, zero-error transmission of a discrete random variable is possible.
3) With feedback and average distortion, the asymptotic fundamental limit is approached at a fast speed $O\left(\frac{\log k}{k}\right)$, where k is the number of source samples.
4) With feedback, the asymptotically achievable minimum energy per sample is reduced by a factor of $1-\epsilon$, where ϵ is the excess distortion probability. This asymptotic fundamental limit is approached from below, i.e., counterintuitively, smaller source blocklengths may lead to better energy-distortion tradeoffs.
5) With feedback, the average energy vs. distortion tradeoff is governed by channel noise power, and the source rate-distortion and rate-dispersion functions. In particular, the channel dispersion plays no role.

References

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

[2] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Information Theory, Special Issue on Shannon Theory: Perspective, Trends and Applications*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[3] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Minimum energy to send k bits through the Gaussian channel with and without feedback," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4880–4902, 2011.

[4] W. Yang, G. Durisi, and Y. Polyanskiy, "Minimum energy to send k bits over Rayleigh-fading channels," in *IEEE International Symposium on Information Theory, submitted*, 2015.

[5] A. Jain, D. Gunduz, S. R. Kulkarni, H. V. Poor, and S. Verdú, "Energy-distortion tradeoffs in Gaussian joint source-channel coding problems," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3153–3168, 2012.

[6] J. M. Wozencraft and I. M. Jacobs, *Principles of communication engineering*. John Wiley & Sons, 1965.

[7] Y. P. Pyatoshin, "Some properties of m-ary communication systems with coding," *Problemy Peredachi Informatsii*, vol. 4, no. 1, pp. 45–51, 1968.

[8] V. Kostina, Y. Polyanskiy, and S. Verdú, "Variable-length compression allowing errors (extended)," *ArXiv preprint, arXiv:1402.0608*, Jan. 2014.

[9] E. Yang and Z. Zhang, "On the redundancy of lossy source coding with abstract alphabets," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1092–1110, May 1999.

[10] J. Schalkwijk and L. Bluestein, "Transmission of analog waveforms through channels with feedback," *IEEE Transactions on Information Theory*, vol. 13, no. 4, pp. 617–619, 1967.

[11] T. J. Cruise, "Achievement of rate-distortion bound over additive white noise channel utilizing a noiseless feedback channel," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 583–584, Apr. 1967.

[12] D. Wang, A. Ingber, and Y. Kochman, "The dispersion of joint source-channel coding," in *Proceedings 49th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, Sep. 2011.

[13] V. Kostina and S. Verdú, "Lossy joint source-channel coding in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2545–2575, May 2013.

[14] V. Kostina, Y. Polyanskiy, and S. Verdú, "Joint source-channel coding with feedback (extended)," *ArXiv preprint*, Jan. 2015.