On Uncoded Transmission and Blocklength

Yuval Kochman
School of CSE, HUJI
Jerusalem, Israel
Email: yuvalko@cs.huji.ac.il

Gregory W. Wornell
EECS Dept., MIT
Cambridge, MA 02139, USA
Email: gww@mit.edu

***Abstract*—This work considers the definition of the excess-distortion exponent, used to measure the asymptotic finite block-length behavior of joint source-channel coding. We arrive at the conclusion that it is not a meaningful measure for the operational tradeoffs of a scheme. We propose a new definition, which makes a distinction between the processing block of the coding scheme (which implies delay and may be connected to complexity), the fidelity blocklength (reflecting the quality of the reconstruction as required by the application), and the resource blocklength (depending on hardware or shared medium considerations). As an aside, the exponent of uncoded schemes is analyzed. This results in finding the joint source-channel coding excess-distortion exponent in some cases where it was not known previously.**

I. Introduction

Shannon's separation principle establishes that there is no loss in the conceptually and practically convenient decomposition of the transmission problem into digital compression and communication problems: in the infinite-blocklength limit, the expected distortion of a separation-based scheme is optimal among all schemes. This optimal expected distortion D^* satisfies $R(D^*) = C$.[1]

In spite of the optimality of separation in the asymptotic expected-distortion sense, joint source-channel coding (JSCC) is the focus of ever-growing research. This is largely due to the sub-optimality of separation in network settings and under channel uncertainty; however, even in a point-to-point scenario with full channel knowledge, JSCC is advantageous in terms of finite-blocklength performance.

It is long known that for some combinations of source statistics, channel statistics and distortion measure, *uncoded transmission* (i.e., the encoder and decoder process each sample independently) can also achieve expected distortion D^*; see e.g. [1] for the white quadratic-Gaussian case. Gastpar et al. [2] have found necessary and sufficient conditions for this fortunate situation where uncoded schemes are optimal, for memoryless finite-alphabet sources and channels.

An alternative to the expected distortion measure is given by the concept of *excess-distortion probability*, borrowed from lossy source coding [3]: the situation where the distortion exceeds some prescribed threshold is seen as an error event, and the relation between the threshold, the probability and the blocklength is studied. Formally, we can define the problem as follows.

Let $\mathcal{S}$, $\hat{\mathcal{S}}$, $\mathcal{X}$ and $\mathcal{Y}$ be the source, reconstruction, channel input and channel output alphabets, respectively, all assumed to be finite. The source is defined by some i.i.d. distribution $p(S)$ and a non-negative single-letter distortion measure $d(S, \hat{S})$. The channel is defined by a conditional distribution $w(Y|X)$ and a non-negative cost function $\gamma(X)$. All of these quantities are defined over the obvious alphabets. Let $h_m = (f_m, g_m)$ be a JSCC scheme of length m: the encoder f_m maps a source block $\boldsymbol{S} = S_1^m$ into channel inputs $\mathbf{X} = X_1^m$ while the decoder g_m maps a block of channel outputs $\mathbf{Y} = Y_1^m$ into the reconstruction $\hat{\boldsymbol{S}} = \hat{S}_1^m$.[2]

We can further consider the concatenation of many m-blocks and obtain infinite-duration processes $\{S\}$, $\{\hat{S}\}$, $\{X\}$, $\{Y\}$. The scheme parses the process $\{S\}$ into m-blocks that we denote as processing blocks. Regardless of this parsing, we can check whether the JSCC scheme satisfies the distortion and cost constraints with respect to the processes. Specifically, the distortion of the *fidelity block* (indexed by its length n and location n_0) is a time-average of the single-letter distortion:

$$d_{n,n_0}(\{S\}, \{\hat{S}\}) = \frac{1}{n} \sum_{i=n_0+1}^{n_0+n} d(S_i, \hat{S}_i) \quad . \tag{1}$$

Similarly, the cost of the *resource block* is a time-average of the single-letter cost:

$$\gamma_{k,k_0}(\{X\}) = \frac{1}{k} \sum_{i=k_0+1}^{k_0+k} \gamma(X_i) \ . \tag{2}$$

As quantifying the exact tradeoff between the blocks (parametrized by m, n, k, n_0, k_0), the distortion and cost thresholds and the probability of failure is a very complicated task, it is plausible to consider the asymptotic behavior in the limit of long blocks. The JSCC excess-distortion exponent, defined for discrete memoryless source-channel pairs by Csiszár [4] and extended to Gaussian source-channel pairs by Zhong et al. [5], takes the processing, fidelity and resource blocks to be the same: $n = k = m$ and $n_0 = k_0 = 0$. Then, having a single blocklength, the exponent gives the rate of decay of the excess-distortion probability as the blocklength increases, for a fixed distortion threshold.[3]

This work was supported in part by NSF under Grant No. CCF-1017772, AFOSR under Grant No. FA9550-11-1-0183, and by Hewlett-Packard Laboratories.

[1] Sometimes this is referred to as the "OPTA".

[2] We exclude the case where the number of channel uses differs from the source blocklength, since it is not needed for the main message of this work.

[3] An alternative approach is the JSCC dispersion [6] where the failure probability is held fixed, while the threshold is allowed to change. Due to space constraints, we do not consider dispersion in this paper, but most of the conclusions regarding the exponent are valid for both.

Uncoded transmission is of special practical importance. As mentioned above, it can achieve D^*, the optimum in the expected sense. But what is the JSCC exponent of uncoded transmission? In terms of the definitions above, such a scheme has processing blocklength $m = 1$, which is in contradiction to setting $n = k = m$ and considering asymptotics. The natural workaround is to define an m-dimensional scheme as one operating independently over each symbol, and then apply the exponent framework. As this procedure assigns an increasing processing blocklength m to a scheme that operationally has processing blocklength one, it clearly cannot reflect the true tradeoff between that blocklength and performance. In particular, it obscures the advantage of using a short processing blocklength. Indeed we demonstrate, that while in some cases following this procedure shows that uncoded transmission achieves the optimal exponent (including cases where the exponent was not previously known), in others it is inferior to digital schemes.

Since the unnecessary coupling between the processing, fidelity and resource blocklengths prohibits one from deriving meaningful operational tradeoffs, we conclude the paper by suggesting an alternative definition.

II. Notation and Background

A. JSCC Exponents

Recall that for the sake of the JSCC exponent, the processing, fidelity and resource blocks are the same. Consequently, we can treat distortion and cost violations in a unified manner: for some thresholds (D, Γ), we consider the case where a scheme h_n has either excess distortion or a cost violation as an "error". The error probability is thus:

$$\epsilon(h_n, D, \Gamma) = \Pr\{d_{n,0}(\{S\}, \{\hat{S}\}) > D \text{ or } \gamma_{n,0}(\{X\}) > \Gamma\}, \quad (3)$$

where d_{n,n_0} and γ_{k,k_0} were defined in (1), (2).

Let $\mathcal{H} = \{h_n : n = 1, 2, \ldots\}$ be some sequence of schemes. The upper and lower excess-distortion exponents are defined as [4]:[4]

$$\begin{aligned} \overline{E}(D, \Gamma) &= \sup_{\mathcal{H}} \limsup_{n\to\infty} -\frac{1}{n} \log \epsilon(h_n, D, \Gamma) \\ \underline{E}(D, \Gamma) &= \sup_{\mathcal{H}} \liminf_{n\to\infty} -\frac{1}{n} \log \epsilon(h_n, D, \Gamma) \quad . \end{aligned} \quad (4)$$

When these two are equal, then the following limit exists and the JSCC excess-distortion exponent is well-defined:

$$E(D, \Gamma) = \lim_{n\to\infty} \sup_{\mathcal{H}} -\frac{1}{n} \log \epsilon(h_n, D, \Gamma) \ . \quad (5)$$

These JSCC exponents can be bounded in terms of source and channel exponents as follows. The source excess-distortion exponent [3] is

$$F(R, D) = \inf_{q: R_q(D) \geq R} D(q\|p) \ , \quad (6)$$

[4]Csiszár's original definition assumes no channel input constraints. Zhong et al. [5] needed to impose a constraint for dealing with continuous alphabets; they chose, similar to the definition of the channel-coding exponent, to require each block output of f_n to satisfy the constraint. Our definition leads to the same results, since a post-encoder device may convert violations into errors.

where q is a distribution over $\mathcal{S}$, $R_q(D)$ is its rate-distortion function and $D(\cdot\|\cdot)$ is the divergence. The channel sphere-packing exponent [7] is

$$E^{sp}(R, \Gamma) = \max_{\phi: E_\phi(\gamma) \leq \Gamma} \min_{v: I(\Phi, v) \leq R} D(v\|w|\phi) \ , \quad (7)$$

where ϕ is a distribution over $\mathcal{X}$ and $D(\cdot\|\cdot|\cdot)$ is the conditional divergence. The channel random-coding exponent is given by

$$E^r(R, \Gamma) = \max_{\phi: E_\phi(\gamma) \leq \Gamma} \min_v \Big[D(v\|w|\phi) + |I(\phi, v) - R|^+ \Big] \ , \quad (8)$$

where $|\cdot|^+$ means limiting to non-negative values. The critical rate $R^{cr} \leq C$ is the rate threshold above which $E^r(R) = E^{sp}(R)$. The achievable channel exponent is given by:

$$E^a(R, \Gamma) = \max\Big(E^r(R, \Gamma), E^x(R, \Gamma)\Big) \quad (9)$$

where $E^x(R, \Gamma)$ is the expurgated error exponent [8]. The expurgated exponent is higher than the random-coding exponent for rates below R^x (which, in turn, is at most R^{cr}). We do not give the expressions for E^x and R^x due to space limitations.

The bounds on the exponents (4) are summarized in the following.

Theorem 1: Bounds on the JSCC excess-distortion exponents [4].

$$\underline{\underline{E}}(D, \Gamma) \leq \underline{E}(D, \Gamma) \leq \overline{E}(D, \Gamma) \leq \overline{\overline{E}}(D, \Gamma)$$

where

$$\underline{\underline{E}}(D, \Gamma) = \inf_R \Big[F(R, D) + E^a(R, \Gamma) \Big] \quad (10)$$

$$\overline{\overline{E}}(D, \Gamma) = \inf_R \Big[F(R, D) + E^{sp}(R, \Gamma) \Big] \quad (11)$$

If (10) is minimized by some R above the channel critical rate R^{cr}, then the limits coincide and the exponent (5) is well-defined. This happens when the distortion threshold D is not too far above D^*, i.e. $D \leq D^{cr}$ for some *critical distortion* D^{cr}.

The bound (11) may be re-written as minimization over source types rather than over rates, as follows: (a similar result holds for (10) as well)

$$\overline{\overline{E}}(D, \Gamma) = \inf_q \left[D(q\|p) + E^{sp}\left(R_q(D), \Gamma\right) \right]. \quad (12)$$

We define a separation-based scheme as one that is comprised of source and channel parts, separated by a a digital interface between the source and channel parts which has fixed rate and is *arbitrary*. The first requirement forbids an interface where the rate depends on the source sequence, while the second means that we cannot count on "close" quantized source points being mapped to "close" channel inputs; for a formal definition see e.g. [6]. We note that although the achievability proof of the lower exponent in [4] is based on an unequal error protection (UEP) scheme which is fully digital, it is not separation-based, as the source-channel interface has source-dependent rate. In fact, except in some degenerate cases, the lower exponent is strictly higher than the exponent achieved by any sequence of separation-based schemes.

B. Symmetric Source-Channel Pairs

Symmetric sources and channels allow the exponent expressions to be simplified, and will be used in the sequel. There are several definitions for a symmetric channel, out of which we choose Gallager's definition [8, p. 94]. We also define a symmetric source in a dual way.

To that end, a matrix is said to be *symmetric* if it can be divided column-wise into sub-matrices, where in each sub-matrix all columns are permutations of each other and all rows are permutations of each other. A source is symmetric if the distortion-measure matrix (where rows are indexed by the source value and columns by the reconstruction) is symmetric and the source distribution is uniform. A channel is symmetric if the transition matrix (where rows are indexed by the input value and columns by the output) is symmetric and the cost is uniform.

Lemma 1: For a symmetric source, $R_q(D)$ is maximized by the uniform distribution, thus

$$F_p(R, D) = \begin{cases} 0 & \text{if } R_p(D) \leq R \\ \infty & \text{otherwise} \end{cases} . \tag{13}$$

For a symmetric channel, the conditional divergence $D(v\|w|\phi)$ is independent of ϕ and the optimal input distribution for the channel capacity as well as for the exponents $E_w^{sp}(R)$ (7), $E_w^r(R)$ (8) and $E_w^x(R)$ is uniform.

We define a symmetric source-channel pair as the combination of a symmetric source with a symmetric channel. Combining the lemma above with (11) we have that for a symmetric pair:

$$\overline{\overline{E}}_{p,w}(D) = E_w^{sp}(R_p(D)) = \min_{v: I(\phi,v) \leq R_p(D)} D(v\|w|\phi) \tag{14}$$

where ϕ is uniform.

III. The JSCC Exponent of Uncoded Transmission

Any uncoded scheme $h = h_1 = (f_1, g_1)$ induces a sequence of schemes $\mathcal{H}$, formed by independently applying h on each source sample and channel output. We define the exponents of an uncoded scheme as in (4), with respect to the sequence $\mathcal{H}$. In order to evaluate these quantities, let $\phi_{q,h}(X)$ and $\psi_{v,h}(\hat{S}|S)$ be the channel input distribution and the reproduction conditional distribution, induced by the source distribution $q(X)$, the channel distribution $v(Y|X)$ and the scheme h.

Proposition 1: The excess-distortion exponent of an uncoded scheme h is given by:

$$E_{p,w,h}(D, \Gamma) = \inf_{q,v \notin \mathcal{L}} [D(q\|p) + D(v\|w|\phi_{q,h})] \tag{15}$$

where the admissible distribution region is given by:

$$\mathcal{L} = \mathcal{L}(D, \Gamma) = \{q, v : E_{q,\psi_{v,h}} d(S, \hat{S}) \leq D, E_{\phi_{q,h}} \gamma(X) \leq \Gamma\} .$$

The proof is a direct consequence of Sanov's Theorem. Of special interest is the case where the source and channel are *compatible*, i.e., $\mathcal{X} = \mathcal{S}$ and $\mathcal{Y} = \hat{\mathcal{S}}$, and the scheme is *direct mapping*, i.e., the encoder and decoder are given by $f(S) = S$ and $g(Y) = Y$, respectively. In this case, $\phi_{q,h}(X) = q(X)$ and $\psi_{v,h}(\hat{S}|S) = v(\hat{S}|S)$. Recall that an optimal scheme is one that achieves the optimum expected distortion D^*. Since an optimal uncoded scheme must be "information lossless" [2], i.e., $I(S, \hat{S}) = I(X, Y)$, direct mappings capture the essence of such schemes.[5] In the next sections, we show when direct mappings that achieve D^*, are also optimal in the sense of exponent.

A. Optimality of Uncoded Schemes for Symmetric Pairs

For symmetric pairs, the uncoded scheme exponent (15) specializes to (assuming the use of direct mapping):

$$E_{p,w,h}(D) = \inf_{v: E(d(S,\hat{S})) > D} D(v\|w|p)]. \tag{16}$$

The following result connects optimality of an uncoded scheme in the expected distortion sense, to optimality in the excess-distortion sense.

Theorem 2: For a symmetric source-channel pair, assume that there exists an uncoded scheme h that achieves D^*. Then the JSCC excess-distortion exponent exists, is given by (14), and is achieved by the scheme.

Proof outline: For simplicity, consider the compatible case where h is a direct mapping, then we need to show that (14) and (16) are equivalent. In this case, necessarily w is the RDF-achieving channel. The necessary condition [2] reduces in the symmetric case to:

$$\log \frac{w(Y|X)}{[p(X)w(Y|X)]_y} = -c \cdot d(X, Y) + k \, \forall X \in \mathcal{X}, Y \in \mathcal{Y},$$

where $c \geq 0$ and k are normalization constants. However, these "local" properties do not suffice, as the exponent may be governed by an empirical channel v far from w. To that end, we first notice that the optimal v must lead to the same marginal on Y as w, and then that for any such v,

$$D(v\|w|p) = I(p, v) + c \cdot E_{p,v} d(x, y) + k.$$

In light of this, the minimizations (14) and (16) are indeed equivalent.

Comparing to the result of Theorem 1, we see that for $D > D^{cr}$ optimal direct mapping has a better exponent than $\underline{\underline{E}}$ achieved by Csiszàr's UEP scheme, and it closes the gap between the lower and upper exponents. For lower D, however, the two approaches yield the same exponent. It is interesting to note, that in the symmetric case the digital scheme does not require any unequal error protection, but reduces to a separation-based scheme.

Examples for symmetric pairs where an optimal uncoded scheme exists include uniform source with difference distortion measure, sent over an unconstrained modulo-additive channel, where the distortion measure and noise distribution "agree", e.g., binary-symmetric source and channel with Hamming distortion. We choose to present a different case, as follows.

[5] In general, optimal uncoded schemes may include operations such as randomly mapping the same source symbol to some two channel inputs x_1 and x_2, if $w(Y|X = x_1) = w(Y|X = x_2)$.

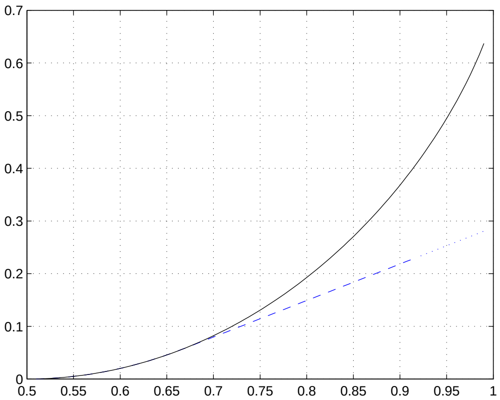

Fig. 1: Erasure exponents as a function of D for $\epsilon = 0.5$. Solid black curve is the JSCC exponent $E(R)$ according to Theorem 2. Blue dashed curve is the lower exponent according to Theorem 1. It becomes dotted in the expurgated region, where the lower exponent was not evaluated.

Example 1: **Symmetric binary erasure source-channel pair**. Let the source be uniform over $\{0, 1\}$, the reconstruction alphabet be $\{0, 1, B\}$ and the distortion measure be erasure:

$$d(S, \hat{S}) = \begin{cases} 0, & S = \hat{S} \\ 1, & \hat{S} = B \\ \infty, & \text{otherwise.} \end{cases} \tag{17}$$

Let the channel be a binary erasure channel (BEC) with the same alphabets (where the erasure symbol is B) and with uniform cost. It is easy to see that the optimal expected distortion D^* equals the erasure probability ϵ, and that direct mapping is optimal. By Theorem 2, the JSCC excess-distortion exponent is $E(D) = D_b(D\|\epsilon)$ where D_b is the binary divergence:

$$D_b(a\|b) = a \log \frac{a}{b} + (1-a) \log \frac{1-a}{1-b} \ . \tag{18}$$

For comparison, the random-coding exponent gives:

$$\underline{\underline{E}}(D) = \min_{\epsilon \le \epsilon' \le D} \left\{ D_b(\epsilon'\|\epsilon) + |D - \epsilon'|^+ \log(2) \right\},$$

for any D below the expurgation regime: $D \le 1 - {R^x(\epsilon)}/{\log(2)}$. The gap between the optimal exponent and the lower exponent is demonstrated in Figure 1.

B. Asymmetric Pairs

For channels that are not symmetric, the optimal (exponent-achieving) input distribution may be rate-dependent, and the same also applies for the test channel that the source code materializes. For such settings the UEP scheme has the inherent advantage that the digital codes may be adjusted according to the rate required by the source type. However, these schemes may still be inferior above the critical distortion, when the sphere-packing bound is not known to be tight. An uncoded scheme may therefore perform better or worse than the bound (10), depending on the distortion threshold.

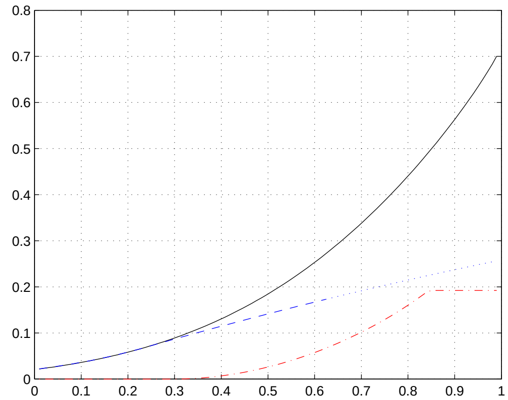

Fig. 2: Constrained erasure exponents as a function of D for $\epsilon = 0.3$, $p = 0.3$ and $\Gamma = 0.6$. Dash-dotted red curve is direct mapping exponent $E^{sp}(R)$ according to Proposition 1. Blue dashed curve is the lower exponent according to Theorem 1. It becomes dotted in the expurgated region, where the lower exponent was not evaluated. Solid black curve is the upper exponent according to the same theorem.

C. The Effect of an Input Constraint

Introducing a channel input constraint (via a non-uniform cost function) reveals another limitation of uncoded schemes. In a digital scheme, a random codebook which satisfies the constraint in average, may be always replaced by one in which each individual codeword satisfies the constraint (by expurgation arguments); in fact, the random-coding exponent is achieved by a single-decomposition codebook, see e.g. [9]. On the other hand, if an i.i.d. source is mapped directly to the channel inputs, then an input sequence of any length has some probability of breaking the constraint. Furthermore, an uncoded scheme might not use the channel well, i.e., might work with an input cost that is *lower* than the constraint. We demonstrate this in the following.

Example 2: **Constrained symmetric binary erasure source-channe pair.** We change Example 1 in the following way. We add a third input symbol $X = B$, with $w(B|B) = \epsilon$ and $w(1|B) = w(0|B) = (1-\epsilon)/2$. Since $w(Y|X = B)$ equals the distribution of Y under the optimal (symmetric) input distribution on $(0, 1)$, this additional input is "useless". Define the cost function $\gamma(0) = \gamma(1) = 1$, $\gamma(B) = 0$, then we have that $C(\Gamma) = \Gamma(1-\epsilon)$. We also add to the source a third letter $S = B$, with $d(B, Y) = 0$ for all Y, thus this source letter is "insignificant". Let B have probability $1 - p$, while 0 and 1 have probability $p/2$. The RDF of this source is given by $R(D) = p(1 - D)$. It is easy to see that for $p = \Gamma$, direct mapping achieves the optimum $D^* = \epsilon$.

Considering the exponent, note that for any finite block length, the probability of direct mapping to break the input cost constraint $\Gamma = p$ is about $1/2$, thus this scheme has a zero exponent! A non-zero exponent can be achieved when direct mapping is not optimal, i.e., when $\Gamma > p$. In that case, direct mapping will achieve for $D \ge \epsilon$:

$$E^{sl}_{p,\epsilon}(D, \Gamma) = \min \left\{ D_b(\Gamma\|p), \min_{p \le q \le \Gamma} \left[D_b(q\|p) + q \cdot D_b(D\|\epsilon) \right] \right\}.$$

This reflects a loss compared to the lower exponent:

$$\overline{\overline{E}}_{p,\epsilon}(D,\Gamma) = \min_{q \geq p} \left[D_b(q\|p) + \Gamma \cdot D_b \left(1 - \frac{q}{\Gamma}(1-D) \| \epsilon \right) \right],$$

and indeed may be even worse than the lower exponent guaranteed by Theorem 1 for some or for all D, as demonstrated in Figure 2. Direct mapping achieves $\overline{D} = \epsilon$, thus it has zero exponent for $D < \epsilon$. The digital UEP scheme, in contrast, can use the channel input with cost Γ, resulting in $\overline{D} = 0$. For high distortions, the exponent of direct mapping is limited by that of the source weight being higher than Γ.

IV. Optimal Scheme for the Cost-Constrained Erasure Pair

Although in the example above an uncoded scheme does not achieve the upper exponent, we show that a modification of it does. This approach may be extended to apply to other cases of logarithmic distortion measures (see [10]), but due to space constraints we do not define these here.

Theorem 3: For the cost-constrained erasure pair of Example 2, the JSCC exponent exists and equals the upper exponent.

Proof outline: The encoder can measure the portion of source letters different from B. Denote that binary type by q. It then uses direct mapping of the source to the channel input, except that if $q > \Gamma$ then the last $(q - \Gamma)n$ source samples which are not B are still mapped to B. Now first make the (unrealistic) assumption that the decoder also knows how many samples at the end of block were mapped to B. In that case, it can ignore these outputs and produce reconstruction B. One may verify, that this leads to the upper exponent, It remains to show how the decoder can be made aware of this number. This can be achieved using a prefix, e.g. one using a concatenation of the Alias coding of the integers with some channel code; this small amount of information can be sent with an error probability decaying faster than the exponent, at the cost of vanishing excess distortion.

A similar approach can be applied to the *quadratic-Gaussian* case. This continuous-alphabet setting is not formally defined here due to space constraints. Zhong et al. [5] have shown that an exponent similar to Gallager's lower exponent can be achieved in this setting. We can show that the JSCC exponent indeed exists and equals the equivalent of the *upper* exponent. This exponent can be achieved by applying a scalar factor to the source block that normalizes it to a fixed radius that satisfies the power constraint. Given the normalization factor, the decoder can undo this normalization and the exponent is achieved. The factor, in turn, can be revealed with sufficient accuracy and vanishing cost, just as in the erasure case.

V. Discussion: Where is the Advantage of Uncoded Transmission?

In this work we have evaluated the excess-distortion exponent achievable by single-letter schemes, and compared it to the lower exponent of Csiszár [4], proven via UEP coding. We have shown that in some cases the single-letter scheme outperforms the UEP exponent, and even closes the gap to the upper exponent. Furthermore, we have shown an example where, while the single-letter scheme fails, a variant of it does achieve the upper exponent.

Considering the cases where the exponent of the optimal uncoded scheme is worse than that of the UEP scheme, we argue that this happens only because of the definition of the exponent. After all, the uncoded scheme always operates at processing blocklength $m = 1$, while the UEP scheme needs a long processing blocklength. It is true, that the uncoded scheme may require larger fidelity blocklength n or resource blocklength k, in order to average over the randomness of the source and channel. These blocklengths have a very different significance, and there is no operational reason to equate them as the exponent definition [4] does. This equating of blocklength may be legacy from digital (separate source and channel) problems, where such an operation does not incur any loss; for example, expurgation of channel codebooks avoids input constraint violations altogether, while keeping the rate and error probability the same as for an average constraint.[6]

In order to solve this problem, we propose to define an exponent triplet (or pair in the absence of an input constraint), where the different blocklengths de-coupled. For any finite-blocklength (e.g. uncoded) scheme, the processing exponent is always infinite. Specifically, when uncoded transmission achieves the "traditional" JSCC exponent $E(D)$ (as in Theorem 2), that would result in infinite processing exponent, together with a fidelity exponent which equals $E(D)$, thus demonstrating the advantage of uncoded transmission. In other cases, the infinite processing exponent may come at the cost of a lower fidelity or resource one. Our new definition, not included in this version, will appear in the full paper.

References

[1] T. J. Goblick. Theoretical limitations on the transmission of data from analog sources. *IEEE Trans. Info. Theory*, IT-11:558–567, 1965.

[2] M. Gastpar, B. Rimoldi, and M. Vetterli. To code or not to code: Lossy source-channel communication revisited. *IEEE Trans. Info. Theory*, IT-49:1147–1158, May 2003.

[3] K. Marton. Error exponent for source coding with a fidelity criterion. *IEEE Trans. Info. Theory*, IT-20:197–199, Mar. 1974.

[4] I. Csiszár. On the error exponent of source-channel transmission with a distortion threshold. *IEEE Trans. Info. Theory*, IT-28:823–838, Nov. 1982.

[5] Y. Zhong, F. Alajaji, and L.L. Campbell. Joint source-channel coding excess distortion exponent for some memoryless continuous-alphabet systems. *IEEE Trans. Info. Theory*, IT-55:1296–1319, Mar. 2009.

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

[7] E. A. Haroutunian. Estimates of the error exponent for the semi-continuous memoryless channel. *Problemy Pered. Inform. (Problems of Inform. Trans.)*, 4, No. 4:37–48, 1968.

[8] R. G. Gallager. *Information Theory and Reliable Communication*. Wiley, New York, N.Y., 1968.

[9] I. Csiszár and J. Korner. *Information Theory - Coding Theorems for Discrete Memoryless Systems*. Academic Press, New York, 1981.

[10] T. Courtade and T. Weissman. Multiterminal source coding under logarithmic loss. *IEEE Trans. Info. Theory*, submitted.

[6] This is true as long as the processing and resource blocklengths are taken to infinity together with a fixed ratio. Otherwise, an input constraint is a hard problem, see e.g. the extensive literature on transmission under a power peak constraint.