Source Requantization: Successive Degradation and Bit Stealing

Aaron S. Cohen Stark C. Draper Emin Martinian Gregory W. Wornell
Massachusetts Institute of Technology, Cambridge, MA 02139
{acohen,scd,emin,gww}@allegro.mit.edu

Abstract

We consider source requantization in two forms — successive degradation (i.e., source fidelity reduction) and bit stealing (i.e., information embedding) — when no forward planning has been done to facilitate the requantization. We focus on finite-alphabet sources with arbitrary distortion measures as well as the Gaussian-quadratic scenario. For the successive degradation problem, we show an achievable distortion-rate trade-off for non-hierarchically structured rate-distortion achieving codes, and compare it to the distortion-rate trade-off of successively refinable codes.

We further consider source requantization in the form of bit stealing, whereby an information embedder acts on a quantized source, producing an output at the same rate. Building on the successive degradation results, we develop achievable distortion-rate trade-offs. Two cases are considered, corresponding to whether the source decoder is informed of any bit stealing or not. In the latter case, the embedder must produce outputs in the original source codebook.

For the Gaussian-quadratic scenario, all trade-offs are within 0.5 bits/sample of the distortion-rate bound. Furthermore, for bit stealing, the use of simple post-reconstruction processing that is only a function of the embedded rate can eliminate the loss experienced by uninformed decoders.

1 Introduction

We consider aspects of the problem of source requantization when no provisions have been made for requantization. First, we examine the successive degradation, or quantization rate reduction, problem. As a motivating example, consider transporting a source through a multi-hop network, where the source is originally encoded at rate R_{orig} with distortion d_0. If the capacity of some intermediate link along the path drops to a residual rate $R_{\text{res}} < R_{\text{orig}}$ due to fading, congestion, etc., a transcoder is required to reduce the rate of the quantized source to R_{res}, which in turn increases

This work has been supported in part by NSF Grant No. CCR-0073520 and a grant from Microsoft Research.

the distortion to $d > d_0$. Ideally, we would like transcoding to be efficient in the sense that R_{res} and d still lie on the rate-distortion curve.

If the source was originally encoded in a successively refinable manner [1], one could easily reduce the rate by erasing the least significant descriptions. We consider what can be achieved even without requiring such special source coding structure at the original encoder by considering arbitrary random rate-distortion codebooks.[1]

In the network congestion example, when a source packet and data packet arrive at a link that cannot support their combined rate, an alternative to dropping source bits to accommodate the data packet is to inject the data bits into the source bits via information embedding into the source reconstruction. If the embedded (stolen) rate is R_{emb}, then the residual rate that is effectively allocated to the source representation is $R_{\text{res}} = R_{\text{orig}} - R_{\text{emb}}$. We examine the degree to which this can be efficient in terms of the associated distortion-rate trade-offs.

The basic difference between the scenarios considered here and those in, e.g., [1, 2], is the sign of the change in total rate. Those works consider how best to make use of an additional source observation and an increase in the total rate (above the rate of the first codebook) to minimize the distortion of the final source estimate(s). In this work the total rate change is negative, and when that change must be made there is no source observation to work with, nor have prior provisions been made for the rate decrease. In contrast to the results of [1], these constraints lead to an unavoidable rate-distortion inefficiency, even in the Gaussian-Quadratic case. Furthermore, as opposed to, e.g., [3, 4, 5], the order of operations considered herein is reversed. In those works information embedding occurs first, followed by compression. We consider the reverse scenario where information embedding must be done in an already-compressed source.

Section 2 considers requantization in the form of successive degradation, while Section 3 in turn considers requantization in the form of bit stealing, both for informed and uninformed destination source decoders.

2 Successive Degradation

Fig. 1 depicts the successive degradation scenario. The source $\mathbf{x}$ is encoded at rate R_{orig} by a rate-distortion achieving source code using a random codebook, producing source reproduction $\hat{\mathbf{x}}$ at distortion $d_0 \approx E\left[D(\hat{\mathbf{x}}, \mathbf{x})\right]$. Throughout this paper, we assume that the source coder is generated in an i.i.d. manner; this structure is sufficient to achieve the rate distortion bound and we believe that the output of any good source coder will look approximately i.i.d. The reproduction $\hat{\mathbf{x}}$ is processed at the successive degrader to produce a second source reproduction $\bar{\mathbf{x}}$. We define $R_{\text{res}}(d)$ to be the successive degradation rate-distortion function. That is, $R_{\text{res}}(d)$ is the smallest residual rate such that we can keep the distortion level $E\left[D(\bar{\mathbf{x}}, \mathbf{x})\right] < d + \delta$. In Appendix A we show the following:

[1] By 'arbitrary random rate-distortion codebook' we mean that the codebook is generated according to an i.i.d. distribution $p(\hat{x}|x)$, without any special structure such as, e.g., the Markov structure of a successively-refinable codebook which, in the notation of [1], satisfies $p(\hat{x}_1, \hat{x}_2|x) = p(\hat{x}_1|\hat{x}_2)p(\hat{x}_2|x)$ where $E\left[D(\hat{x}_1, x)\right] \geq E\left[D(\hat{x}_2, x)\right]$.

Proceedings of the DATA COMPRESSION CONFERENCE (DCC'02)
1068-0314/02 $17.00 © 2002 IEEE

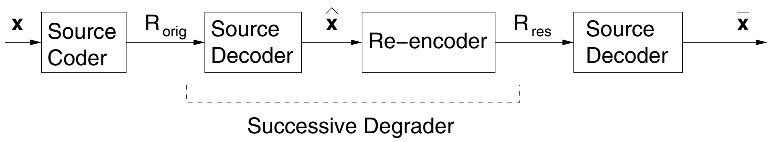

Figure 1: Requantization for successive degradation: reducing the rate of an already-quantized source $\hat{\mathsf{x}}$; $R_{\text{res}} < R_{\text{orig}}$.

Theorem 1 *Let* x *be drawn i.i.d.* $\sim p(x)$ *and let* $p(\hat{x}|x)$ *determine the design of a random source-codebook with rate* $R_{\text{orig}} = I(\hat{\mathsf{x}}; \mathsf{x})$ *and distortion* $d_0 < E[D(\hat{\mathsf{x}}, \mathsf{x})] + \delta$*. An achievable upper bound on the successive degradation rate-distortion function* $R_{\text{res}}(d)$ *is given by* $R_{\text{res}}(d) \leq \inf_{p(\bar{x}|\hat{x})} I(\bar{\mathsf{x}}; \hat{\mathsf{x}})$ *where* $p(\bar{x}, \hat{x}, x) = p(\bar{x}|\hat{x})p(\hat{x}|x)p(x)$ *and* $E[D(\bar{\mathsf{x}}, \mathsf{x})] \leq d$*.*

2.1 Gaussian-Quadratic Results

In Appendix B we show that an achievable upper-bound on the successive degradation rate-distortion function $R_{\text{res}}(d)$ for i.i.d. Gaussian sources under a mean-squared distortion (MSD) measure is, for $d > d_0$,

$$R_{\text{res}}(d) \leq \min\left\{\frac{1}{2}\log\left[\frac{\sigma_{\mathsf{x}}^2 - d_0}{d - d_0}\right], \frac{1}{2}\log\left[\frac{\sigma_{\mathsf{x}}^2}{d_0}\right]\right\}. \tag{1}$$

As $d \to \sigma_{\mathsf{x}}^2$, $R_{\text{res}} \to 0$. As $d \to d_0$, $R_{\text{res}} \to R_{\text{orig}}$. Because the approach in Appendix B requantizes the source as if it is an i.i.d. Gaussian source, while in fact it is a member of a finite set of $2^{nR_{\text{orig}}}$ vectors, there is some finite range of distortions above d_0 for which requantization will not save us any rate, it is better to leave the original source code as is. To find the point where successive degradation (without time-sharing) becomes useful, set the first term in the minimization of (1) equal to R_{orig} to get

$$d = d_0 + (\sigma_{\mathsf{x}}^2 - d_0)2^{-2R_{\text{orig}}} = d_0\left(2 - 2^{-2R_{\text{orig}}}\right) \triangleq d_*. \tag{2}$$

This says that each time we independently requantize the source with an independently generated source code of the same rate, the overall distortion increases by roughly d_0. We now upper bound the rate-loss in successive degradation (which can be further decreased with time-sharing). Since the biggest rate loss occurs at $d = d_*$ in (2), the maximum loss is

$$R_{\text{res}}(d_*) - R(d_*) \leq \frac{1}{2}\log\left[\frac{\sigma_{\mathsf{x}}^2}{d_0}\right] - \frac{1}{2}\log\left[\frac{\sigma_{\mathsf{x}}^2}{d_0(2 - 2^{-2R_{\text{orig}}})}\right] = \frac{1}{2}\log[2 - 2^{-2R_{\text{orig}}}] < \frac{1}{2}. \tag{3}$$

Our results indicate there is a positive rate-loss: unlike in the successive refinement problem [1], in successive degradation we do not have access to the original source signal when degrading the source, only to a quantized version of it. In recent work [6] we have shown that random codebooks can be constructed so that $R_{\text{res}}(d)$ is also lower-bounded by the right-hand-side of (1). In other words, in the worst case (1) holds with equality.

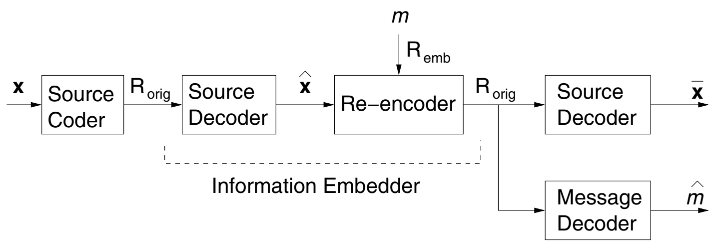

Figure 2: Requantization for bit stealing: embedding a message into an already-quantized source; $R_{\text{emb}} < R_{\text{orig}}$.

3 Embedding in a Quantized Source

Fig. 2 depicts the scenario where requantization takes the form of embedding a message into a quantized source. The source $\mathbf{x}$ is encoded at rate R_{orig} by a rate-distortion achieving source code using a random codebook, producing source reproduction $\hat{\mathbf{x}}$. A message m of rate R_{emb} is embedded into host $\hat{\mathbf{x}}$. Because we want to embed a message into the quantized source, the rates into and out of the embedder are identical. The decoder produces two estimates: of the source, $\bar{\mathbf{x}}$, and of the message, $\hat{m}$.

We consider two embedding subclasses of such problems. In the first class, the source destination is informed of any bit stealing and thus the embedder is free to create an entirely new source codebook, requantize the source to that code, and transmit that description; the remaining bits are used for the message. In the second class, the source destination is not informed of any bit stealing, and thus the embedder must produce outputs in the original source codebook. In the Gaussian-quadratic case we show that, even with this additional constraint, if the decoder becomes informed of the rate of the embedding R_{emb}, then through the use of a simple post-reconstruction scaling, the same distortion can be achieved as if a totally new codebook were generated — as in the first class.

3.1 Embedding for Informed Decoders: Rate-Splitting

In this case, it suffices to apply the results of Section 2 to the embedding problem. In effect, we split the rate into two streams. One, of rate R_{emb}, encodes the message. The other, of rate $R_{\text{orig}} - R_{\text{emb}}$ is the residual rate left to describe the source. Theorem 1 gives us a strategy to achieve a particular rate-distortion trade-off between R_{emb} and d for any finite-alphabet source.

To explore this trade-off for Gaussian sources and MSD, substitute $R_{\text{orig}} - R_{\text{emb}}$ into (1) for R_{res}, and solve for d to see that this method achieves distortion equal to

$$d(R_{\text{orig}} - R_{\text{emb}}) = d_0 + (\sigma_x^2 - d_0)2^{-2(R_{\text{orig}} - R_{\text{emb}})} = d_0 + d_0\left(1 - 2^{-2R_{\text{orig}}}\right)2^{2R_{\text{emb}}}. \quad (4)$$

3.2 Embedding for Uninformed Decoders

In this case, the source decoders in Fig. 2 are identical. We define $R_{\text{emb}}(d)$ to be the largest embedding rate such that we can keep the distortion level $E\left[D(\bar{\mathbf{x}}, \mathbf{x})\right] < d + \delta$.

Proceedings of the DATA COMPRESSION CONFERENCE (DCC'02)
1068-0314/02 $17.00 © 2002 IEEE

In Appendix C we show the following:

Theorem 2 *Let x be drawn i.i.d $\sim p(x)$ and let $p(\hat{x}|x)$ determine the design of a random source-codebook with rate $R_{\text{orig}} = I(\hat{\mathsf{x}};\mathsf{x})$ and distortion $d_0 < E\left[D(\hat{\mathsf{x}},\mathsf{x})\right] + \delta$. An achievable lower-bound on the rate-distortion embedding function is given by $R_{\text{emb}}(d)$ that satisfy $R_{\text{orig}} - R_{\text{emb}}(d) \leq \inf_{p(\bar{x}|\hat{x})} I(\bar{\mathsf{x}};\hat{\mathsf{x}})$ where $p(\bar{x},\hat{x},x) = p(\bar{x}|\hat{x})p(\hat{x}|x)p(x)$, $p_{\hat{\mathsf{x}}}(x) = p_{\bar{\mathsf{x}}}(x)$, and $E\left[D(\bar{\mathbf{x}},\mathbf{x})\right] < d + \delta$.*

3.2.1 Gaussian-Quadratic Results

Let x be an i.i.d. zero-mean Gaussian sequence of variance σ_{x}^2. Assume we are given $\hat{\mathbf{x}}$, the output of a rate R_{emb} source coder with MSD $\frac{1}{n}E\left[\sum_{i=1}^{n}(\hat{\mathsf{x}}_i - \mathsf{x}_i)^2\right] \leq d_0 + \delta$. If we wish to embed in this quantized source a message of rate $R_{\text{emb}} \leq R_{\text{orig}}$, then Appendix D shows that the following MSD is achievable by an uninformed decoder:

$$\begin{aligned} d_{\text{U}}(R_{\text{emb}}) &= \frac{1}{n}E\left[\sum_{i=1}^{n}(\bar{\mathsf{x}}_i - \mathsf{x}_i)^2\right] \\ &= \sigma_{\mathsf{x}}^2\left(2 - 2^{-2R_{\text{orig}}}\right) - 2\sigma_{\mathsf{x}}^2\left(1 - 2^{-2R_{\text{orig}}}\right)\sqrt{1 - 2^{-2(R_{\text{orig}} - R_{\text{emb}})}}. \end{aligned} \tag{5}$$

In the case of a decoder that is informed of the rate of embedding, Appendix D further shows we can improve the distortion by multiplying $\bar{\mathbf{x}}$ by β where

$$\beta = \sqrt{1 - 2^{-2(R_{\text{orig}} - R_{\text{emb}})}}. \tag{6}$$

In particular, this scaling is invertible, so it does not effect the embedding rate, but improves distortion to

$$d_{\text{I}}(R_{\text{emb}}) = d_0 + d_0\left(1 - 2^{-2R_{\text{orig}}}\right)2^{2R_{\text{emb}}} \tag{7}$$

where $d_0 = D(R_{\text{orig}}) = \sigma_{\mathsf{x}}^2 2^{-2R_{\text{orig}}}$. Comparing (7) with (4) we see that the informed decoder does as well as one given a completely new codebook.

As one limiting case, we substitute $R_{\text{emb}} = 0$ into (7) to get [c.f. (2)] $d_{\text{I}}(0) = d_0(2 - 2^{-2R_{\text{orig}}}) \simeq 2d_0$. This can be interpreted as follows. If $R_{\text{emb}} = 0$ the embedder has nothing to do and no additional distortion should be incurred, i.e $E\left[(\bar{\mathsf{x}} - \mathsf{x})^2\right] = d_0$. But if we wish to embed at a slightly positive rate, $R_{\text{emb}} > 0$, because the embedder is restricted to mappings between quantization vectors, the codeword will be moved from the original quantization vector to a different one in the source codebook to encode the message. In a good vector quantizer, quantization vectors have a minimal spacing of approximately $2d_0$. For positive rates then, we expect to see a minimal distortion on the order of $2d_0$.

As a second limiting case, we substitute $R_{\text{emb}} = R_{\text{orig}}$ into (6) and (7) to get $\beta = 0$ and $d_{\text{I}}(R_{\text{orig}}) = \sigma_{\mathsf{x}}^2$. In contrast, in the case of uninformed decoders ($\beta = 1$), since $\beta|_{R_{\text{emb}}=0} \neq 1$, then $d_{\text{U}}(0) > d_{\text{I}}(0) = d_0(2 - 2^{-2R_{\text{orig}}})$. In particular, substituting $R_{\text{emb}} = R_{\text{orig}}$ into (5) yields $d_{\text{U}}(R_{\text{emb}}) = 2\sigma_{\mathsf{x}}^2 - d_0$. This can be interpreted as follows. When $R_{\text{emb}} = R_{\text{orig}}$ the source is erased, and each quantization vector is used to encode a different message. Erasing all information relevant to the source incurs distortion σ_x^2. Mapping a message independent of the source to one of the codebook codewords incurs an additional distortion $(\sigma_x^2 - d_0)$, the variance of the quantization vectors. This results in an overall distortion of $2\sigma_x^2 - d_0$. Clearly though, if all bits are used to send the embedded message, there is no data about the original source left, so the decoder (if informed) should set $\bar{x} = 0$, yielding a MSD of σ_x^2, which is reflected by $\beta|_{R_{\mathrm{emb}}=R_{\mathrm{orig}}} = 0$.

A normalized distortion measure for the informed decoder is, via (7),

$$d_{\mathrm{I}}^{\mathrm{norm}}(R_{\mathrm{emb}}) = d_{\mathrm{I}}(R_{\mathrm{emb}})/\sigma_x^2 = 2^{-2R_{\mathrm{orig}}} + 2^{-2(R_{\mathrm{orig}}-R_{\mathrm{emb}})}\left(1 - 2^{-2R_{\mathrm{orig}}}\right). \tag{8}$$

A normalized measure for uninformed decoders, $d_{\mathrm{U}}^{\mathrm{norm}}(R_{\mathrm{emb}})$ can be derived by dividing (5) by σ_x^2.

In Fig. 3 normalized distortion is plotted versus residual source coding rate, $R_{\mathrm{res}} = R_{\mathrm{orig}} - R_{\mathrm{emb}}$, for $R_{\mathrm{orig}} = 2$. The normalized distortions of the uninformed decoder, $d_{\mathrm{U}}^{\mathrm{norm}}(R_{\mathrm{emb}})$, and of the informed decoder, $d_{\mathrm{I}}^{\mathrm{norm}}(R_{\mathrm{emb}})$, are plotted as dashed and dash-dotted curves, respectively.[2] Time-sharing, when possible, makes the points that lie along the dotted line achievable. At a particular residual source-coding rate, the minimum achievable distortion is lower-bounded by the distortion-rate function $D(R_{\mathrm{orig}} - R_{\mathrm{emb}})$, which is plotted with the solid curve.

In Fig. 3, note the "requantization" gap between the bit stealing curves and the distortion-rate bound at $R_{\mathrm{orig}} - R_{\mathrm{emb}} = R_{\mathrm{orig}}$, i.e., $R_{\mathrm{emb}} = 0$, which occurs because the embedder is requantizing the quantized source. Unless the original quantization point is mapped back to itself, a non-zero minimal distortion is incurred. Since, using the approach detailed in Appendix C, for any $R_{\mathrm{emb}} > 0$ the probability of mapping back to the same point goes to zero, we observe a gap. Unfortunately, one cannot always avoid this discontinuity via time-sharing because one cannot time-share within a single vector codebook. Only if there is a sequence of quantized sources transmitted can time-sharing be implemented. In this special case one can time-share between the achievable distortions given by (8) and doing nothing ($R_{\mathrm{emb}} = 0$) to get the points along the dotted line.

A Successively Degrading a Finite-Alphabet Source

Let $x_1, x_2, \ldots, x_n$ be drawn i.i.d. $\sim p(x)$ and let $D(\mathbf{x}, \hat{\mathbf{x}})$ be a distortion measure for this source. Given a randomly-generated source code that achieves distortion d_0, we can re-encode this codebook into a second codebook of rate R that achieves distortion d as long as $R > R_{\mathrm{res}}(d)$, where R_{res} is the rate-distortion function for this source/first codebook pair. We show that an achievable upper-bound on $R_{\mathrm{res}}(d)$ is

$$R_{\mathrm{res}}(d) \leq \inf_{p(\bar{x}|\hat{x})} I(\bar{x}; \hat{x}) \tag{9}$$

where $p(\bar{x}, \hat{x}, x) = p(\bar{x}|\hat{x})p(\hat{x}|x)p(x)$ and $E[D(x, \bar{x})] \leq d$.

[2] As shown in (7), the distortion-rate performance of the informed decoder for Gaussian sources is equivalent to the performance of the rate-splitting embedder (4). This means that the dash-dot line in Fig. 3 also shows us the distortion-rate performance of successive degradation source coding for the quadratic Gaussian case discussed in Section 2.1.

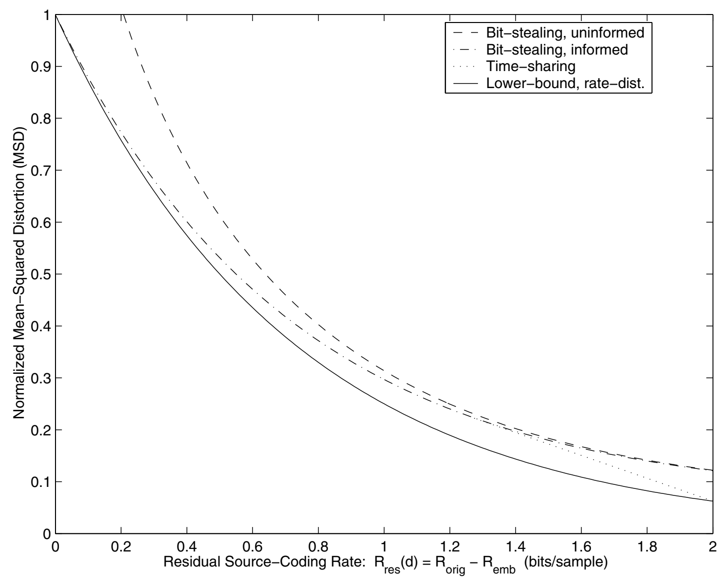

Figure 3: Residual-Rate-Distortion behavior; $R_{\text{orig}} = 2$. The dashed and dash-dot curves plot $d_{\text{U}}^{\text{norm}}(R_{\text{emb}})$ and $d_{\text{I}}^{\text{norm}}(R_{\text{emb}})$, respectively; the latter is equivalent to successive degradation performance. Time-sharing (when possible) makes the points along the dotted line achievable. The solid curve is the lower-bound corresponding to the Gaussian distortion-rate function $D(R_{\text{orig}} - R_{\text{emb}})$. Note $d_{\text{U}}^{\text{norm}}(R_{\text{emb}} = R_{\text{orig}}) = 2 - 2^{-2R_{\text{orig}}} \approx 2$.

Quantizer: Assume the source is quantized using a randomly generated codebook according to $p(\hat{x}|x)$ where $R_{\text{orig}} > I(\hat{x}; x)$ and $d_0 \leq E\left[D(\hat{x}, x)\right] + \delta$. In particular, the source codebook $\mathcal{C}$ consists of $2^{nR_{\text{orig}}}$ n-length sequences generated according to $\prod_{i=1}^{n} p(\hat{x}_i)$. These sequence are labeled $\hat{\mathbf{x}}(1), \ldots, \hat{\mathbf{x}}(2^{nR_{\text{orig}}})$. For source encoding, find the indices i such that $(\mathbf{x}, \hat{\mathbf{x}}(i)) \in A_{\epsilon}^{*(n)}$, where $A_{\epsilon}^{*(n)}$ denotes the set of jointly strongly typical sequences of length n. If there is more than one such index, choose any one of them. Transmit that i.

Source Degrader: Given that we want to increase the source distortion to $d > d_0$, at the source degrader search over joint distributions $p(\bar{x}, \hat{x}, x) = p(\bar{x}|\hat{x})p(\hat{x}|x)p(x)$ such that (1) $p(\hat{x}|x)$ and $p(x)$ are as defined as above, and (2) $E\left[D(\bar{x}, x)\right] < d + \delta$. Pick the distribution that satisfies these constraints while minimizing $I(\bar{x}; \hat{x})$.

Generate a second codebook $\mathcal{C}_2$ which consists of 2^{nR} n-length sequences generated according to $\prod_{i=1}^{n} p(\bar{x}_i)$. These sequence are labeled $\bar{\mathbf{x}}(1), \ldots, \bar{\mathbf{x}}(2^{nR})$. For encoding, find the indices j such that $(\hat{\mathbf{x}}, \bar{\mathbf{x}}(j)) \in A_{\epsilon}^{*(n)}$. If there is more than one such index, choose any one of them. Transmit that j.

Decoding: At the decoder the source estimate is $\bar{\mathbf{x}}(j)$.

Probability of error: The probability that the source coder cannot find at least one $\hat{\mathbf{x}}(i)$ jointly strongly typical with $\mathbf{x}$ goes to zero for n large. This follows from standard joint strong typicality reasoning because $R_{\text{orig}} > I(\hat{x}; x)$. The probability that the source degrader cannot find a $\bar{\mathbf{x}}(j)$ jointly typical with $\hat{\mathbf{x}}(i)$ goes to zero for n large. This follows from standard joint strong typicality reasoning because $R > I(\bar{x}; \hat{x})$.

Distortion: We now analyze the expected distortion between $\bar{\mathbf{x}}$ and $\mathbf{x}$. Because $\mathbf{x}$ and $\hat{\mathbf{x}}$ are jointly typical, $\hat{\mathbf{x}}$ and $\bar{\mathbf{x}}$ are jointly typical, and $x \multimap \hat{x} \multimap \bar{x}$ form a Markov chain, as n gets large $(\mathbf{x}, \bar{\mathbf{x}}) \in A_\epsilon^{*(n)}$ with probability approaching one by the Markov lemma. Because they are jointly strongly typical, $E[D(\mathbf{x}, \bar{\mathbf{x}})] = \sum_{x,\bar{x}} p(x, \bar{x}) D(x, \bar{x})$.

Since $R > I(\bar{x}, \hat{x})$ is achievable, and $R > R_{\text{res}}(d)$, then (9) provides an upper bound on the rate-distortion function $R_{\text{res}}(d)$.

B Successively Degrading a Gaussian Source

To extend the results of Appendix A to continuous alphabets, we must partition $x \to x_p$, $\hat{x} \to \hat{x}_p$ and $\bar{x} \to \bar{x}_p$ carefully, so as to preserve the Markov relationship $x_p \multimap \hat{x}_p \multimap \bar{x}_p$ for each partition. See [7] for more details.

Define the following test channel: $\hat{x} = \alpha x + e_{\text{orig}}$ where $\alpha = (1 - d_0/\sigma_x^2)$ and $e_{\text{orig}} \sim N(0, \alpha d_0)$. This is the standard test-channel for rate-distortion source coders. Next define a second standard test channel $\bar{x} = \gamma \hat{x} + e_{\text{res}}$ where $\gamma = (1 - \Delta/\sigma_{\hat{x}}^2)$ and $e_{\text{res}} \sim N(0, \gamma\Delta)$. This compound test channel gives us overall MSD $d = E\left[(\bar{x} - x)^2\right] = d = d_0 + \Delta$, and

$$I(\bar{x}; \hat{x}) = \frac{1}{2} \log \left[\frac{\sigma_x^2 - d_0}{\Delta}\right]. \tag{10}$$

Substitute $\Delta = d - d_0$ into (10) to get an upper bound on $R_{\text{res}}(d)$,

$$R_{\text{res}}(d) \leq \frac{1}{2} \log \left[\frac{\sigma_x^2 - d_0}{d - d_0}\right], \tag{11}$$

where $d - d_0 > 0$.

C Embedding in a Quantized Finite-Alphabet Source

Let $x_1, x_2, \ldots, x_n$ be drawn i.i.d. $\sim p(x)$ and let $D(\mathbf{x}, \hat{\mathbf{x}})$ be a distortion measure for this source. Given a randomly-generated source code that achieves distortion d_0, we can embed a message of rate R into this codebook causing overall distortion d as long as $R < R_{\text{emb}}(d)$, where R_{emb} is the rate-distortion function for this source/first codebook pair. We show that an achievable lower-bound on $R_{\text{emb}}(d)$ is given by the $R_{\text{emb}}(d)$ that satisfy

$$R_{\text{orig}} - R_{\text{emb}}(d) \leq \inf_{p(\bar{x}|\hat{x})} I(\bar{x}; \hat{x}) \tag{12}$$

where $p(\bar{x}, \hat{x}, x) = p(\bar{x}|\hat{x}) p(\hat{x}|x) p(x)$, $p_{\bar{x}}(x) = p_{\hat{x}}(x)$, and $E[D(x, \bar{x})] \leq d$.

Quantizer: Assume the source is quantized using a randomly generated codebook according to $p(\hat{x}|x)$ where $R_{\text{orig}} > I(\hat{x}; x)$ and $d_0 \leq E[D(\hat{x}, x)] + \delta$. In particular, the source codebook $\mathcal{C}$ consists of $2^{nR_{\text{orig}}}$ n-length sequences generated according to $\prod_{i=1}^n p(\hat{x}_i)$. These sequence are labeled $\hat{\mathbf{x}}(1), \ldots, \hat{\mathbf{x}}(2^{nR_{\text{orig}}})$. For source encoding, find the indices i such that $(\mathbf{x}, \hat{\mathbf{x}}(i)) \in A_\epsilon^{*(n)}$, where $A_\epsilon^{*(n)}$ denotes the set of jointly strongly typical sequences of length n. If there is more than one such index, choose any one of them. Transmit that i.

Information Embedder: Given that we want to embed a message at distortion d, at the information embedder search over joint distributions $p(\bar{x}, \hat{x}, x) = p(\bar{x}|\hat{x})p(\hat{x}|x)p(x)$ such that (1) $p(\hat{x}|x)$ and $p(x)$ are as defined as above, (2) $p_{\bar{x}}(x) = p_{\hat{x}}(x)$, and (3) $E[D(\bar{x}, x)] < d + \delta$. Pick the distribution that satisfies these constraints while minimizing $I(\bar{x}; \hat{x})$.

Randomly bin $\mathcal{C}$ into 2^{nR} subcodes $\mathcal{C}_j$ where $R = R_{\text{orig}} - I(\bar{x}; \hat{x}) - \epsilon$, $\epsilon > 0$.. That is, for each $\hat{\mathbf{x}}(i)$ pick an index j uniformly distributed over $1, \ldots, 2^{nR}$ and assign $\hat{\mathbf{x}}(i)$ to subcode $\mathcal{C}_j$. On average there are $2^{n(R_{\text{orig}}-R)}$ $\hat{\mathbf{x}}$ sequences in each $\mathcal{C}_j$. Relabel the $\hat{\mathbf{x}}$ sequences in $\mathcal{C}_j$ as $\bar{\mathbf{x}}(j,k)$ where $j \in 1, \ldots, 2^{nR}$ and $k \in 1, \ldots, 2^{n(R_{\text{orig}}-R)}$. Finally, given a source codeword $\hat{\mathbf{x}}(i)$, and a message $m = m$, find the indices k such that $(\hat{\mathbf{x}}(i), \bar{\mathbf{x}}(m,k)) \in A_\epsilon^{*(n)}$. If there is more than one such index pick any one. Transmit the index l such that $\hat{\mathbf{x}}(l) = \bar{\mathbf{x}}(m,k)$.

Decoding: The source estimate is $\hat{\mathbf{x}}(l)$, and the message estimate is $\hat{m} = m$ such that $\bar{\mathbf{x}}(m,k) = \hat{\mathbf{x}}(l)$.

Probability of error: The probability that the source coder cannot find at least one $\hat{\mathbf{x}}(i)$ jointly strongly typical with $\mathbf{x}$ goes to zero for n large. This follows from standard joint strong typicality reasoning because $R_{\text{orig}} > I(\hat{x}; x)$.

The probability that the information embedder cannot find a $\bar{\mathbf{x}}(m,k)$ jointly typical with $\hat{\mathbf{x}}(i)$ goes to zero for n large. To see this, note that the probability that the original source-quantization vector $\hat{\mathbf{x}}(i)$ falls into the selected bin m is 2^{-nR}, which goes to zero for n large. Then, conditioned on the event that $\hat{\mathbf{x}}(i)$ is not in bin m, the rest of the codewords in bin m look like i.i.d. sequences generated independent of $\hat{\mathbf{x}}(i)$ according to $\prod_{i=1}^n p_{\bar{x}}(x_i)$. (Since the elements of $\mathcal{C}_m$ are also in $\mathcal{C}$, and those in $\mathcal{C}$ were generated independently according to $\prod_{i=1}^n p_{\hat{x}}(x_i)$, because $p_{\hat{x}}(x) = p_{\bar{x}}(x)$ the codewords in bin m look like they were generated according to $\prod_{i=1}^n p_{\bar{x}}(x_i)$.) The probability that at least one of these sequences, $\bar{\mathbf{x}}(m,k)$, is jointly strongly typical with $\hat{\mathbf{x}}(i)$ goes to one for n large because there are $2^{n(R_{\text{orig}}-R)}$ codewords in bin m and $R_{\text{orig}} - R > I(\bar{x}; \hat{x})$. The probability that $\hat{m} \neq m$ is zero because there is no noise in the system.

Distortion: We now analyze the expected distortion between $\bar{\mathbf{x}}$ and $\mathbf{x}$. Because (1) $\mathbf{x}$ and $\hat{\mathbf{x}}$ are jointly typical, (2) $\hat{\mathbf{x}}$ and $\bar{\mathbf{x}}$ are jointly typical, and (3) $x \leftrightarrow \hat{x} \leftrightarrow \bar{x}$ form a Markov chain, as n gets large $(\mathbf{x}, \bar{\mathbf{x}}) \in A_\epsilon^{*(n)}$ with probability approaching one by the Markov lemma. Because they are jointly strongly typical, $E[D(\mathbf{x}, \bar{\mathbf{x}})] = \sum_{x,\bar{x}} p(x,\bar{x})D(x,\bar{x})$.

Since $R = R_{\text{orig}} - I(\bar{x}; \hat{x}) - \epsilon$ is achievable for all $\epsilon > 0$, and $R \leq R_{\text{emb}}(d)$, then (12) provides a lower bound on the rate-distortion function $R_{\text{emb}}(d)$.

D Embedding in a Quantized Gaussian Source

To extend the results of Appendix C to continuous alphabets, we must partition $x \to x_p$, $\hat{x} \to \hat{x}_p$ and $\bar{x} \to \bar{x}_p$ carefully, so as to preserve the Markov relationship $x_p \leftrightarrow \hat{x}_p \leftrightarrow \bar{x}_p$ for each partition. See [7] for more details. In this section we derive a compound test channel for Gaussian source.

First, define the following test channel: $\hat{x} = \alpha x + e_{\text{orig}}$ where $\alpha = (1 - d_0/\sigma_x^2)$ and $e_{\text{orig}} \sim N(0, \alpha d_0)$. This is the standard test-channel for rate-distortion source coders. Next define an information embedding test channel [8] $\bar{x} = \gamma \hat{x} + e_{\text{emb}}$ where $e_{\text{emb}} \sim N(0, (1 - \gamma^2)(\sigma_x^2 - d_0))$. Note that $E[\bar{x}^2] = E[\hat{x}^2]$ which satisfies the constraint $p_{\hat{x}}(x) = p_{\bar{x}}(x)$ since both are zero-mean Gaussian random variables. Finally, define $\hat{\hat{x}} = \beta\bar{x}$. In Appendix C we did not allow post-processing of the estimate. That was the case of an 'uninformed' decoder, when the decoder does not know embedding has occurred. In that case $\beta = 1$. We analyze the case $\beta = 1$ below, but also analyze the case of 'informed' decoders when β can be chosen as a function of the embedding rate R. The multiplication by β does not effect the achievable transmission rates, since the multiplication can be inverted, hence the achievability arguments from Appendix C still hold. The effect of the β scaling is thus restricted to distortion calculations. We optimize over all possible scalings γ and β.

From rate-distortion results for white Gaussian sources and mean-squared distortion measures, we know that to achieve distortion d_0, $R_{\text{orig}} > \frac{1}{2}\log\left(\frac{\sigma_x^2}{d_0}\right) = I(x;\hat{x})$. From Theorem 2 we know that we can embed at a rate R such that $R_{\text{orig}} - R = I(\hat{x};\bar{x}) = \frac{1}{2}\log\left(\frac{1}{1-\gamma^2}\right)$. Solve for γ as a function of R to get $\gamma = \sqrt{1-2^{-2(R_{\text{orig}}-R)}} \geq 0$, which holds with equality when $R = R_{\text{orig}}$. Next, solve for the expected distortion and substitute in for α, γ, and the variances of e_{orig} and e_{emb} to get

$$
\begin{aligned}
d = E\left[(\hat{\hat{x}} - x)^2\right] &= (\alpha\beta\gamma - 1)^2\sigma_x^2 + \alpha\beta^2\gamma^2 d_0 + \beta^2(1-\gamma^2)(\sigma_x^2 - d_0) \\
&= \sigma_x^2\left[\beta^2\left(1 - \frac{d_0}{\sigma_x^2}\right) - 2\beta\left(1 - \frac{d_0}{\sigma_x^2}\right)\sqrt{1 - \frac{d_0}{\sigma_x^2}2^{2R_{\text{emb}}}} + 1\right]. \qquad (13)
\end{aligned}
$$

For uninformed decoders, substitute $\beta = 1$ and $d_0 = \sigma_x^2 2^{-2R_{\text{orig}}}$ into (13) to get (5). For informed decoders, differentiate (13) with respect to β and set equal to zero to get $\beta = \sqrt{1-2^{-2(R_{\text{orig}}-R)}}$, which upon substitution in to (13) with $d_0 = \sigma_x^2 2^{-2R_{\text{orig}}}$ yields (7). Solving (7) for R yields a lower-bound on $R_{\text{emb}}(d)$ because $R_{\text{emb}}(d) \geq R$.

References

[1] W. H. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Trans. Info. Theory*, vol. 37, pp. 269–275, Mar. 1991.

[2] H. Viswanathan and T. Berger, "Sequential coding of correlated sources," *IEEE Trans. Info. Theory*, vol. 46, pp. 236–246, Jan. 2000.

[3] D. Karakos and A. Papamarcou, "Relationship between quantization and distribution rates of digitally watermarked data," in *Proc. IEEE Inter. Symp. Info. Theory*, 2000.

[4] D. Kundur, "Implications for high capacity data hiding in the presence of lossy compression," in *Proc. IEEE Int. Conf. On Info. Tech.: Coding & Comp.*, pp. 16–21, 2000.

[5] M. Ramkumar and A. N. Akansu, "Information theoretic bounds for data hiding in compressed images," in *Proc. IEEE Workshop on Multimedia Sig. Proc.*, 1998.

[6] A. Cohen, S. C. Draper, E. Martinian, and G. W. Wornell, "Stealing bits from a quantized source," *Submitted to Int. Symp. Info. Theory*, 2002.

[7] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder–II: General sources," *Information and Control*, vol. 38, pp. 60–80, 1978.

[8] M. H. Costa, "Writing on dirty paper," *IEEE Trans. Info. Theory*, vol. 29, pp. 439–441, May 1983.