Entropy of Quantized Data at High Sampling Rates

Daniel Marco
EE Department
California Institute of Technology
Pasadena CA 91101, USA
Email: idaniel@caltech.edu

David L. Neuhoff
EECS Department
University of Michigan
Ann Arbor MI 48109, USA
Email: neuhoff@umich.edu

***Abstract*—This paper considers the entropy of the highly correlated quantized samples resulting from sampling at high rate. Two results are shown. The first concerns sampling and identically scalar quantizing a stationary random process over a finite interval. It is shown that if the process crosses a quantization threshold with positive probability, then the joint entropy of the quantized samples tends to infinity as the sampling interval goes to zero. The second result provides an upper bound to the rate at which the joint entropy tends to infinity, in the case of infinite-level uniform threshold scalar quantizers and a stationary Gaussian random process whose mean lies at a midpoint of some quantization cell. Specifically, an asymptotic formula for the conditional entropy of one quantized sample conditioned on another quantized sample is derived.**

I. Introduction

The problem of analyzing the effectiveness of high rate sampling and scalar quantization has been widely addressed in the literature, for example [1]-[6]. Roughly speaking, the problem is the following: Suppose a continuous-time signal is sampled, scalar quantized, encoded into bits and then reconstructed using some type of interpolation. Is it possible to tradeoff quantizer resolution for higher sampling rate, while maintaining the same distortion between the original and reconstructed signals? Does the rate-vs.-distortion performance of such systems become better or worse as the sampling rate increases? How does one optimize such systems? And how specifically does the rate-vs.-distortion performance vary with the sampling rate?

In this paper, we do not consider the problem in its most general form, but rather restrict discussion to the following: Let a continuous-time stationary random process X, bandlimited or not, be sampled in the interval $[0,1]$ every τ seconds, quantized with a fixed scalar quantizer, binary encoded with an ideal entropy coder, and reconstructed with linear interpolation, producing

$$\widehat{X}_t = \sum_{n=0}^{\lceil \frac{1}{\tau} \rceil - 1} Q(X_{n\tau}) p_\tau(t - n\tau) \ ,$$

where Q denotes the quantization rule of the fixed scalar quantizer, and p_τ is an interpolation pulse, which may change with τ. Let $D_{Q,p,\tau}$ denote the mean squared-error (MSE) in $\widehat{X}$ as a reproduction of X (averaged over the time interval [0,1]). Let $H_\tau(Q(X))$ denote the joint entropy of the quantized samples divided by the number of samples ($\lceil \frac{1}{\tau} \rceil$). H is subscripted by τ to reflect the dependence of the statistics of the quantized samples on τ. (Note that the entropy coder must adapt to τ.) Let $R_\tau(Q) \triangleq \frac{1}{\tau} H_\tau(Q(X))$ denote the rate in bits/second produced by the ideal entropy encoder when quantizing and encoding the samples in the interval $[0,1]$. The first question to address is: For a fixed scalar quantizer Q, what happens to $R_\tau(Q)$ and $D_{Q,p_\tau,\tau}$ as $\tau \to 0$?

We first argue that $\liminf_{\tau \to 0} \inf_p D_{Q,p,\tau} > 0$. This is because for any choice of p, $D_{Q,p,\tau}$, can be no smaller than the MSE of the best linear filter $D_{Q,LF}$, i.e. the Wiener filter, for estimating X_t from the quantized, but not sampled, version of X_t over all time. That is, $D_{Q,LF}$ is the least MSE when linearly estimating X_t from $X_{Q,s} \triangleq Q(X_s)$, $-\infty \leq s \leq \infty$. Thus, by the well-known MSE formula for Wiener filters

$$\begin{aligned} \liminf_{\tau \to 0} \inf_p D_{Q,p,\tau} &\geq \inf_{LF} D_{Q,LF} \\ &= \int_{-\infty}^{\infty} \frac{P_X(\omega) P_{X_Q}(\omega) - |P_{XX_Q}(\omega)|^2}{P_{X_Q}(\omega)} \, d\omega \end{aligned}$$

where P_X and P_{X_Q} are the power spectral densities of X and X_Q, respectively, and P_{XX_Q} is the cross power spectral density. Since the MSE of the Wiener filter is ordinarily positive, we see that as claimed above, linear interpolation yields distortion bounded away from zero.

We now consider what happens to $R_\tau(Q)$ as the sampling interval decreases to zero. Does it go to zero? Remain finite? Tend to infinity? The answer is not obvious in that $R_\tau(Q)$ is the product of $1/\tau$ samples/sec., which increases to infinity, and $H_\tau(Q(X))$ bits/sample, which decreases to zero, due to the fact that the samples (and consequently, the quantized samples) become increasingly correlated as τ decreases. Essentially, the question asks if the increasing correlation can be sufficiently exploited to counteract the increasingly large number of samples.

The first substantial result of this paper shows that under a very mild condition, the answer is no. That is, for any quantizer Q satisfying this mild condition, $R_\tau(Q)$ tends to infinity. In other words, although $H_\tau(Q(X))$ approaches zero, it becomes large relative to τ.

The fact that $R_\tau(Q)$ tends to infinity is not what the authors anticipated. Instead, it seemed plausible that $R_\tau(Q)$ would follow the trend of ideal rate-distortion coding and approach a finite value. Specifically, suppose that instead of

[0]This work was supported by NSF Grant ANI-0112801.

scalar quantizing and ideal entropy coding over the finite interval $[0,1]$, the process were sampled over the interval $(-\infty,\infty)$, and the samples were lossily encoded with an ideal rate-distortion encoder with distortion equal to the limiting distortion of the scalar quantizer based system, i.e., distortion $D = \liminf_{\tau\to 0} \inf_p D_{Q,p,\tau}$. Such an encoder would produce $\frac{1}{\tau}\mathcal{R}_\tau(D)$ bits/sec, where $\mathcal{R}_\tau(\cdot)$ is the rate-distortion function (in bits/sample) of the sampled random process (subscripted by τ to reflect the dependence of the discrete-time process on τ). In the Gaussian case, for example, it is well-known that as τ decreases to zero, $\frac{1}{\tau}\mathcal{R}_\tau(D)$ approaches $\mathcal{R}(D)$, the rate-distortion function of the continuous-time random process X [7]. Since $\mathcal{R}(D)$ is ordinarily finite, we see that restricting the lossy encoder to be a combination of a fixed scalar quantizer and an ideal (sampling-rate adapted) entropy coder over the finite interval [0,1] increases the limiting rate from finite to infinite. This may be surprising because one generally expects scalar quantization with entropy coding to have rate that exceeds the rate-distortion function by at most a constant (c.f. [8], [9]). However, this expectation applies to discrete-time processes, and if the rate $H_\tau(Q(X))$ of scalar quantization with entropy coding exceeds $\mathcal{R}_\tau(D)$ by, say, c bits/sample, then when τ is small, $R_\tau(Q) = \frac{1}{\tau}H_\tau(Q(X))$ exceeds $\mathcal{R}(D) \approx \frac{1}{\tau}\mathcal{R}_\tau(D)$ by approximately c/τ, which is large.

The question then arises as to how fast $R_\tau(Q)$ approaches infinity. In general, this is a difficult question. To obtain a partial answer, we make the problem tractable by considering Gaussian sources, uniform threshold quantization, and conditional entropy coding that attains $H\big(Q(X_\tau)\,|\,Q(X_0)\big)$ instead of $H_\tau(Q(X))$. The second substantial result of this paper (Theorem 7) can be applied to show that when τ is small

$$R_\tau(Q) \;\overset{\sim}{\leq}\; \bar{R}_\tau(Q) \;\overset{\Delta}{=}\; \frac{1}{\tau}H\big(Q(X_\tau)\,|\,Q(X_0)\big) \approx \; -\frac{1}{\tau}M\sqrt{1-\rho(\tau)}\,\log_2\sqrt{1-\rho(\tau)}\;, \quad (1)$$

where

$$M \;=\; \frac{2\sqrt{2}}{\pi}\sum_{k=0}^{\infty} e^{-\frac{(k+\frac{1}{2})^2\Delta^2}{2\sigma^2}}\;,$$

Δ is the step size of the uniform quantizer, σ^2 is the variance of X, and $\rho(\cdot)$ is its normalized covariance function, i.e. $\rho(\tau)$ is the correlation coefficient between adjacent samples of X. Note that the above expression for $\bar{R}_\tau$ depends only on the behavior of $\rho(\tau)$ for τ near zero, where $\rho(\tau)\approx 1$. This means it does not depend on spectral characteristics of X, such as whether the process is bandlimited or not.

For example, if $\rho(\tau) = e^{-|\tau|}$, which implies X is Markov, and if τ is small, then $\sqrt{1-\rho(\tau)/\sigma^2} \approx \sqrt{\tau}$, and

$$\bar{R}_\tau(Q) \;\approx\; -\frac{M}{2}\frac{\log_2\tau}{\sqrt{\tau}}\;. \quad (2)$$

Or if $\rho(\tau) = e^{-\tau^2}$ and τ is small, then $\sqrt{1-\rho(\tau)/\sigma^2} \approx \tau$, and

$$\bar{R}_\tau(Q) \;\approx\; -M\log_2\tau\;. \quad (3)$$

The results so far show, in effect, that for any target distortion d that is less than the variance of X,

$$\inf_{Q:\liminf_{\tau\to 0}\inf_p D_{Q,p,\tau}\leq d}\;\lim_{\tau\to 0} R_\tau(Q) \;=\; \infty \quad (4)$$

That is, if Q is chosen so that the limiting distortion will be at most d, then rate in bits/sec must necessarily tend to infinity as τ decreases. Notice that in (4), we permit the interpolation pulse, as well as the entropy coder, to change with τ, but we do not permit the quantizer to change. One is naturally lead to ask if the result would change if the quantizer were also allowed to change with τ, subject only to a constraint that the distortion be at most d. Under some additional conditions, such as that the process be ergodic and have piecewise continuous sample functions, one can show

$$\lim_{\tau\to 0}\;\inf_{Q,p:D_{Q,p,\tau}\leq d} R_\tau(Q) = \infty \quad (5)$$

(The formal statement and proof are omitted as space does not permit adequate discussion of the technicalities involved.) This shows that even when the scalar quantizer is optimized for the specific sampling rate, the encoded rate in bits/sec. still tends to infinity as the sampling rate increases.

We conclude this section by noting that this work was originally motivated by the desire to understand the potential efficiency of dense sensor networks when given the task of measuring and encoding a two-dimensional field [10] to within some target MSE. To simplify discussion, consider a one-dimensional field/signal indexed by t, $0 \leq t \leq 1$. The sensor measurements can be viewed as samples of the signal. Because these samples are at disparate locations, their encoding must be of the type called "distributed". A natural, though suboptimal, approach to distributed lossy encoding is scalar quantization followed by Slepian-Wolf distributed lossless coding. (Transform coding, DPCM and vector quantization across samples are not permitted.) In this case, the average number of encoded bits per unit length will be $R_\tau(Q)$, when samples spaced τ apart are scalar quantized with Q. The question then arises as to the behavior of $R_\tau(Q)$ as the sensor density increases. On the one hand, as τ decreases, the data from more and more sensors must be encoded. On the other hand, the data produced by these sensors becomes increasingly correlated, and this correlation can be exploited by the Slepian-Wolf lossless coder. It was natural to conjecture that the Slepian-Wolf coder would sufficiently exploit the correlation to counterbalance the increasing number of sensors. Unfortunately, the results described earlier indicate that $R_\tau(Q)$ tends to infinity, which means that the efficiency of a dense sensor network must necessarily degrade, when each sample is quantized with a common scalar quantizer Q. It is not known if this negative result is due to the type of lossy coding considered here (identical scalar quantization of each sample) or is characteristic of lossy distributed coding.

The remainder of this paper is organized as follows. Section II shows that $R_\tau(Q)$ tends to infinity as τ goes to zero. Section III develops the asymptotic (as correlation goes to one) formula for the conditional entropy of one quantized sample

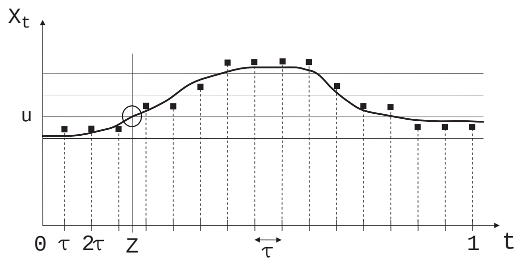

Fig. 1. A sample path of the random process X_t on the interval $[0,1]$, which is sampled and quantized. u is the quantization threshold considered, Z is the first crossing time of u, and τ is the sampling interval.

conditioned on another quantized sample, which is used in (1). Section IV offers concluding remarks. We comment that in the sequel we will outline the proofs, but due to space limitations, we will not actually provide the proofs in this extended abstract.

II. Joint Entropy of Quantized Samples at High Sampling Rates

Consider a continuous-time, stationary and continuous in probability[1] random process that is sampled every τ seconds. Each sample is quantized using an arbitrary, yet unchanging, scalar quantizer. This section addresses the question of what happens to the joint entropy of the scalar quantized samples from some finite time interval, as the sampling interval τ tends to zero? It will be shown that under a very mild condition the joint entropy tends to infinity as $\tau \to 0$. Specifically, we assume that with positive probability the random process crosses (see Definition 1 below) some quantization threshold. This will be our only additional assumption about the random process, aside from stationarity.

The key idea[2] is showing that as $\tau \to 0$, one can obtain from the quantized samples an increasingly accurate description of a quantity that has infinite entropy. This in turn will imply that the joint entropy of the quantized samples tends to infinity as $\tau \to 0$. More specifically, we consider the finite time interval $[0,1]$, and let the approximated quantity be the time of the first crossing in $[0,1]$ of some quantization threshold, which can be increasingly well approximated from the quantized samples of the random process. It will follow that the joint entropy of the outputs of the quantizers tends to infinity. This is illustrated in Figure 1. While the idea is fairly simple, it does involve certain technical hurdles that need to be overcome. Specifically, the following steps are needed:

1. A useful definition of a crossing, whose time can be approximated.
2. A proof that with positive probability a first crossing does indeed occur.
3. The definition of a random variable that is a function of the quantization indices that approximates the time of the first crossing, and a proof that it converges conditionally in probability to the time of the first crossing.
4. A proof that the time of the first crossing is absolutely continuous. Continuity is sufficient to insure that it has infinite entropy, and absolute continuity will be used in the next step.
5. A proof that 2, 3, and 4 above imply that the entropy of the approximating random variable tends to infinity as the sampling interval goes to zero.

Let $\{X_t,\ t \in T\}$, $T = (-\infty, \infty)$, be a continuous-time random process, denoted for short by X, which is defined over the probability space $(\Omega, \mathcal{F}, P)$. Let $\omega \in \Omega$ denote a point in Ω, which corresponds to a sample path denoted $X(\omega)$. Let $X_t(\omega)$ denote the value of the sample path ω at time t.

Next, let $\tau > 0$ denote a sampling interval, and let $I_k^\tau(\omega)$, $k \in \mathbb{Z}$, denote the index of the quantization cell containing $X_{k\tau}(\omega)$, when quantized by some scalar quantizer Q.

Definition 1: ([11], p. 147) Let $u \in \mathbb{R}$, let $\delta > 0$, and let $h : \mathbb{R} \to \mathbb{R}$ be some function. Then h has a u*-upcrossing* of size δ at t if $h(t-s) < u$ and $h(t+s) \geq u$ for all $0 < s < \delta$. h has a u*-upcrossing* at t if there exists $\delta > 0$ such that h has a u-upcrossing of size δ at t.

Note that in [11] a definition very similar to the above is referred to as *strict* upcrossings. Note further that without loss of generality it suffices to focus on upcrossings only, as opposed to both up and downcrossings.

We are now ready to provide a formal definition of first crossing (this corresponds to Step 1).

$$Z_\delta(\omega) \;=\; \begin{cases} t, & X(\omega) \text{ has a first } u\text{-upcrossing in [0,1], it} \\ & \text{is at } t\text{, and it is of size } \delta \\ \infty, & \text{else} \end{cases} .$$

Observe that Z_δ is an extended-value random variable. Note further that Z_δ might equal ∞ not only due to not having any u-upcrossing in $[0,1]$, but also due to having an infinite number of upcrossings with no first upcrossing, e.g. $\sin \frac{1}{t}$, with $u = 0$.

The following lemma establishes Step 2 and more.

Lemma 2: There exists $\delta > 0$ such that

$$\Pr\left(Z_\delta \in [\frac{\delta}{4}, \frac{\delta}{2}]\right) \;>\; 0 \;.$$

Next, in order to define a random variable that approximates Z_δ (i.e. the first part of Step 3), let $\tilde{I}$ denote the index of the quantization cell containing u. Without loss of generality we assume this cell has the form $[u, v)$. Given $0 < \tau < 1$ and $N_\tau = \lfloor \frac{1}{\tau} \rfloor$, we define

$$K^\tau(\omega) \;=\; \text{smallest } k \in \{0, 1, \ldots, N_\tau - 1\} \text{ such that } \left(I_k^\tau(\omega) < \tilde{I} \text{ and } I_{k+1}^\tau(\omega) \geq \tilde{I}\right), \text{ or } 2N_\tau \text{ if no such } k \text{ exists.}$$

Now let

$$\widehat{Z}^\tau(\omega) \;=\; K^\tau(\omega)\,\tau \;,$$

[1]Continuity in probability is a very mild technical condition that is needed to allow stationarity to imply that the events $\{\omega : X_t(\omega) \leq r,\ \ a < t < b\}$ and $\{\omega : X_t(\omega) \leq r,\ \ a+s < t < b+s\}$ have the same probability, and to ensure that when taking the expected value of the time integral of a function of the random process, the expected value can be brought inside the integral.

[2]The authors would like to thank Bruce Hajek for providing this idea.

which is the approximation of $Z_\delta(\omega)$ that can be determined from the quantized samples. The second part of Step 3 is provided by the following lemma:

Lemma 3: Let $\delta > 0$ be given. For any $\varepsilon > 0$,

$$\lim_{\tau\to 0} \Pr\left(|Z_\delta - \widehat{Z}^\tau| < \varepsilon \,|\, Z_\delta \in \left[\frac{\delta}{4}, \frac{\delta}{2}\right]\right) \;=\; 1 \;.$$

The next lemma establishes Step 4.

Lemma 4: Z_δ is absolutely continuous on $[0, 1]$.

We comment that continuity alone of Z_δ implies $H(Z_\delta) = \infty$. However, absolute continuity is needed for Theorem 6 below.

Finally, the following lemma shows that the entropy of the approximating random variable tends to infinity as the sampling interval goes to zero (Step 5), when the conditions of Steps 2,3 and 4 are met.

Lemma 5: Let B be an interval, let Z be a random variable that is absolutely continuous on B, and let $\Pr(Z \in B) > 0$. Let $\{Z_n\}$ be a sequence of discrete random variables such that for any $\varepsilon > 0$, $\lim_{n\to\infty} \Pr(|Z - Z_n| < \varepsilon \,|\, Z \in B) = 1$. Then,

$$\lim_{n\to\infty} H(Z_n) = \infty \;.$$

The main result of this section is now obtained in the following way. Lemma 2 is used to choose $0 < \delta < 1$ for which $\Pr\left(Z_\delta \in \left[\frac{\delta}{4}, \frac{\delta}{2}\right]\right) > 0$. For this choice of δ we apply Lemmas 3, 4 and 5, where $B = \left[\frac{\delta}{4}, \frac{\delta}{2}\right]$ in the last lemma, and observe that for any τ the joint entropy of the quantized samples is at least as large as the entropy of $\widehat{Z}^\tau$. Formally, the result is stated as follows:

Theorem 6: Let X be a continuous-time stationary and continuous in probability random process defined over the probability space $(\Omega, \mathcal{F}, P)$. Let Q be a quantizer having a quantization threshold u such that $\Pr(X \text{ has a } u\text{-upcrossing in } [0,1]) > 0$. Let also $0 < \tau < 1$, and $N_\tau = \left\lfloor \frac{1}{\tau} \right\rfloor$. Then

$$\lim_{\tau\to 0} H(I_0^\tau, I_1^\tau, \ldots, I_{N_\tau}^\tau) \;=\; \infty \;.$$

Note that the condition of Theorem 6 that $\Pr(X \text{ has a } u\text{-upcrossing in } [0,1]) > 0$, is benign and holds for a large class of processes, for example, for any u it holds for any stationary Gaussian process with nonzero variance [11] (p. 153).

Theorem 6 shows in fact that $R_\tau = \frac{1}{\tau} H_\tau(Q(X)) = H(I_0^\tau, I_1^\tau, \ldots, I_{N_\tau - 1}^\tau) \to \infty$ as $\tau \to \infty$. Namely, the rate in bits per second of the sampled and quantized process tends to infinity as the sampling interval goes to zero.

III. Asymptotic formula for conditional entropy

The following is the second main result of this paper.

Theorem 7: Let X_1 and X_2 be jointly Gaussian random variables with zero mean, variance σ^2, and correlation coefficient ρ that are quantized with an infinite-level uniform threshold quantizer with step size Δ, whose k^{th} cell is $[(k - \frac{1}{2})\Delta, (k + \frac{1}{2})\Delta)$. Let $\lambda = \frac{\Delta}{\sigma}$. If I_1 and I_2 denote the integers representing the quantization indices associated with X_1 and X_2, respectively, then

$$\lim_{\rho\to 1} \frac{H(I_2|I_1)}{-M_\lambda \sqrt{1-\rho}\, \log \sqrt{1-\rho}} \;=\; 1 \;,$$

where

$$M_\lambda \;=\; \frac{2\sqrt{2}}{\pi} \sum_{k=0}^{\infty} e^{-\frac{(k+\frac{1}{2})^2 \lambda^2}{2}} \;.$$

We observe that Theorem 7 provides an accurate estimate to the rate of an entropy coder that encodes each index conditioned on the previous, when the source is Gaussian and the quantizer is uniform. As indicated below, it also provides an asymptotically accurate upper bound to the rate R_τ of an entropy coder that jointly encodes all indices.

$$\begin{aligned} R_\tau(Q) = \frac{1}{\tau} H_\tau\big(Q(X)\big) &= H(I_0^\tau) + \sum_{k=1}^{N_\tau - 1} H(I_k^\tau | I_{k-1}^\tau, \ldots, I_0^\tau) \\ &\le H(I_0^\tau) + (N_\tau - 1) H(I_1^\tau | I_0^\tau) \approx \frac{1}{\tau} H(I_1^\tau | I_0^\tau) \;. \end{aligned} \quad (6)$$

Substituting the result of Theorem 7 into the right-hand side of the above yields (1), which as mentioned earlier, provides an upper bound to the rate at which R_τ goes to infinity.

Equations (2) and (3) provide explicit expressions for the upper bound in the case of exponential and Gaussian autocorrelation functions. Next is a corollary to Theorem 7.

Corollary 8:

$$\lim_{\lambda\to 0} \lim_{\rho\to 1} \frac{H(I_2|I_1)}{-\frac{2}{\sqrt{\pi}} \frac{1}{\lambda} \sqrt{1-\rho}\, \log \sqrt{1-\rho}} \;=\; 1 \;.$$

Letting $\lambda \to 0$ means the quantizers become high resolution. Intuitively this should increase the conditional entropy. Indeed, the corollary shows that asymptotically, as $\lambda \to 0$, the conditional entropy increases as $\frac{1}{\lambda}$.

The complete proof of Theorem 7 is quite long and complicated and so we only provide some intuition and discussion of what is involved.

We observe that finding a conditional entropy of the form $H(I_2|X_1)$ is not trivial. However, since for such a case the conditional distribution of X_2 given X_1 is Gaussian, this problem has in fact been solved in [12], where the output entropy of a uniform quantizer with a Gaussian source has been evaluated for asymptotically small rate. For the case examined here, namely, finding the conditional entropy $H(I_2|I_1)$, the derivation becomes significantly more difficult. The reason for this stems from the fact that I_1 is the quantized version of X_1, which in turn makes the conditional distribution of X_2 given $I_1 = k$ no longer Gaussian.

The key to finding this conditional entropy lies in understanding the behavior of the conditional distribution of X_2 given that $I_1 = k$. The expression for $f_{X_2|I_1}(x|k)$ is

$$\begin{aligned} f_{X_2|I_1}(x|k) \;&=\; \frac{\Pr(I_1 = k \,|\, X_2 = x) f_{X_2}(x)}{\Pr(I_1 = k)} \\ &=\; \frac{1}{P_k} f_{X_2}(x) \left[Q\big(\frac{t_k - \rho x}{\sigma_\rho}\big) - Q\big(\frac{t_{k+1} - \rho x}{\sigma_\rho}\big) \right] \;, \end{aligned}$$

where f_{X_1} is the zero mean, variance σ^2 Gaussian density, P_k is the probability of X_1 lying in cell k, and $t_k = (k - \frac{1}{2})\Delta$ is the left threshold of the k^{th} quantization cell.

Consider the influence of k, i.e. the quantization cell in which X_1 lies, on $f_{X_2|I_1}(x|k)$. Suppose that ρ is very close to one, thus, X_2 and X_1 are very correlated, which means that if X_1 lies in quantization cell k, we expect that with high probability X_2 would lie in the same cell. Indeed, upon plotting of $f_{X_2|I_1}(x|k)$ with $\rho = 0.99$ and $k = 1$, we observed that X_2 lies in cell 1 with high probability. However, we further noticed that if k is sufficiently large, then a seemingly strange thing happens, namely, X_2 lies with high probability in a different cell than the one in which X_1 lies. Specifically, for the case that $\rho = 0.99$ and $k = 17$, X_2 lies with high probability in cell 16 rather than 17. The reason for this shifting towards the origin phenomenon is quite simple. Specifically, if $X_1 = x_1$, then $EX_2 = \rho x_1$. Thus, no matter how close to one ρ may be, the distance between the value of X_1, i.e. x_1, and the mean of X_2, i.e. ρx_1, equals $(1-\rho)x_1$, which can be made arbitrarily large, (i.e. it can equal the length of many quantization cells) by letting x_1 be sufficiently large. The take away message from this discussion is that for any value of ρ, if k is sufficiently small, then X_2 lies in the k^{th} cell with high probability, and if k is too large, then this is no longer true.

The proof now follows in four main steps, each showing that one term on the right-hand side of the equation below converges to one as $\rho \to 1$.

$$\frac{H(I_2|I_1)}{M_\lambda \mathcal{H}(\sqrt{1-\rho})} = \frac{H(I_2|I_1)}{\sum_{|k|\le N(\rho)} H(I_2|I_1=k)P_k} \times \frac{\sum_{|k|\le N(\rho)} H(I_2|I_1=k)P_k}{\sum_{|k|\le N(\rho)} \left(H_{k-1|k}+H_{k|k}+H_{k+1|k}\right)P_k} \times \frac{\sum_{|k|\le N(\rho)} \left(H_{k-1|k}+H_{k|k}+H_{k+1|k}\right)P_k}{\sum_{|k|\le N(\rho)} \left(H_{k-1|k}+H_{k+1|k}\right)P_k} \times \frac{\sum_{|k|\le N(\rho)} \left(H_{k-1|k}+H_{k+1|k}\right)P_k}{M_\lambda \mathcal{H}(\sqrt{1-\rho})}$$

where $\mathcal{H}(p) = -p\log p$, $H_{l|k} = \mathcal{H}(P_{l|k})$, $P_{l|k} = \Pr(I_2 = l|I_1 = k)$, and $N(\rho)$ is a carefully chosen integer function of ρ that goes to infinity as $\rho \to 1$. Specifically, the idea is to choose $N(\rho)$ so that on the one hand it grows slowly enough so that for $|k| \le N(\rho)$, $f_{X_2|I_1}(x|k)$ will be so concentrated on the k^{th} quantization cell that $P_{k|k} \approx 1$ and $H(I_2|I_1 = k) \approx H_{k-1|k} + H_{k|k} + H_{k+1|k}$ (so the shifting towards the origin phenomenon described above does not occur), while on the other hand it grows rapidly enough that the first term on the right hand side above tends to one as $\rho \to 1$. A choice of $N(\rho)$ that satisfies these two competing requirements is

$$N(\rho) = \left\lfloor \left(\ln\frac{1}{1-\rho}\right)^{\frac{3}{4}} - \frac{1}{2} \right\rfloor .$$

To obtain the third term approximation we recall that $P_{k|k}$ is a point on the right side of the curve $-p\log p$, while $P_{k-1|k} + P_{k+1|k}$ is close to the origin. The approximation then follows from the fact that the slope of $-p\log p$ at $p = 0$ is infinite while the slope at $p = 1$ is finite. Lastly, the fourth term approximation is obtained by evaluating the required probabilities and identifying and discarding negligible terms.

IV. Conclusions

This paper considered the behavior of the entropy of highly correlated quantized data. First we examined the case that a stationary random process that crosses a quantization threshold with positive probability is sampled over some finite interval, and each sample is separately quantized with arbitrary, yet identical, scalar quantizer Q. It was shown that rate in bits per second $R_\tau(Q)$ of these quantized samples (i.e. their joint entropy) tends to infinity as the sampling interval goes to zero, while the induced MSE remains bounded away from zero.

Next, we established an upper bound to the rate at which $R_\tau(Q)$ above tends to infinity. This upper bound was obtained by upper bounding high order conditional entropies by first-order conditional entropy (as in (6)), namely, $H(I_k^\tau|I_{k-1}^\tau,\ldots,I_1^\tau,I_0^\tau) \le H(I_1^\tau|I_0^\tau)$. A simple asymptotic formula was provided for the first-order conditional entropy in the case of infinite-level uniform threshold quantizers and a stationary Gaussian random process whose mean lies at a midpoint of some quantization cell. This formula holds for bandlimited and non-bandlimited processes alike. Indeed, the convergence rate of the first-order conditional entropy depends only on the behavior of the autocorrelation function near the origin, thus it is of no consequence whether it is bandlimited or not. As examples, we considered two autocorrelation functions — an exponential and a Gaussian. The upper bound to the joint entropy was shown to go infinity at rates $\frac{\log\tau}{\tau}$ and $\log\tau$ for the former case and latter case, respectively.

References

[1] N. T. Thao and M. Vetterli, "Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates," *IEEE Trans. Signal Processing*, vol. 42, pp. 519–531, Mar. 1994.

[2] ——, "Lower bound on the mean-squared error in oversampled quantization of periodic signals using vector quantization analysis," *IEEE Trans. Info. Theory*, vol. 42, pp. 469–479, Mar. 1996.

[3] K. Benhenni and S. Cambanis, "The effect of quantization on the performance of sampling designs," *IEEE Trans. Info. Theory*, vol. 44, pp. 1981–1992, Sep. 1998.

[4] Z. Cvetkovic and I. Daubechies, "Single-bit oversampled A/D conversion with exponential accuracy in the bit-rate," *Data Compression Conference, DCC, Snowbird, UT*, pp. 343–352, Mar. 2000.

[5] Z. Cvetkovic and M. Vetterli, "On simple oversampled A/D conversion in $L^2(\mathbb{R})$," *IEEE Trans. Info. Theory*, vol. 47, pp. 146–154, Jan. 2001.

[6] A. Kumar, P. Ishwar, and K. Ramchandran, "On distributed sampling of smooth non-bandlimited fields," *IPSN, Berkeley*, pp. 89–98, Apr. 2004.

[7] T. Berger, *Rate Distortion Theory*, Prentice-Hall, Englewood Cliffs, 1971.

[8] H. Gish and J. N. Pierce, "Asymptotically efficient quantization," *IEEE Trans. Info. Theory*, vol. 14, pp. 676–683, Sep. 1968.

[9] J. Ziv, "On universal quantization," *IEEE Trans. Info. Theory*, vol. 31, pp. 344–347, May 1985.

[10] D. Marco, E. Duarte-Melo, M. Liu, and D. L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data," *IPSN, Palo Alto, CA.*, pp. 1–16, Apr. 2003.

[11] M. R. Leadbetter, G. Lindgren, and H. Rootzen, *Extremes and Related Properties of Random Sequences and Processes*, Springer Verlag, New York, 1983.

[12] D. Marco and D. L. Neuhoff, "Performance of low rate entropy-constrained quantizers," *ISIT, Chicago, IL.*, p. 495, Jul. 2004.