

Entropy of Highly Correlated Quantized Data

Daniel Marco, *Member, IEEE*, and David L. Neuhoff, *Fellow, IEEE*

Abstract—This paper considers the entropy of highly correlated quantized samples. Two results are shown. The first concerns sampling and identically scalar quantizing a stationary continuous-time 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 rate goes to infinity. The second result provides an upper bound to the rate at which the joint entropy tends to infinity, in the case of an infinite-level uniform threshold scalar quantizer and a stationary Gaussian random process. Specifically, an asymptotic formula for the conditional entropy of one quantized sample conditioned on the previous quantized sample is derived. At high sampling rates, these results indicate a sharp contrast between the large encoding rate (in bits/sec) required by a lossy source code consisting of a fixed scalar quantizer and an ideal, sampling-rate-adapted lossless code, and the bounded encoding rate required by an ideal lossy source code operating at the same distortion.

Index Terms—Entropy, entropy-rate, Gaussian random process, oversampling, quantization threshold crossing, quantized random process, rate-distortion, sampling.

I. INTRODUCTION

CONSIDER the sequence of integer cell indices I_1, \dots, I_N produced by a scalar quantizer q operating on N samples $X(1/N), X(2/N), \dots, X(1)$ taken uniformly in the unit time interval $[0, 1]$ from a continuous-time stationary random process $X(t)$. We are interested in the number of bits per second it takes to losslessly encode $I_1, \dots, I_N = q(X(1/N)), \dots, q(X(1))$, as N increases. With ideal lossless coding, this rate is the joint entropy $H(I_1, \dots, I_N)$ of the quantized samples, which one may view as the product of the sampling rate N times the entropy per sample

$$H(I_1, \dots, I_N) = N \frac{H(I_1, \dots, I_N)}{N}.$$

As the sampling rate increases, the entropy per sample decreases to zero as the samples become increasingly correlated. The question is: Does it decrease fast enough that the joint entropy remains bounded? In this paper, we answer this question in the negative. Specifically, we show that under the mild condition that $X(t)$ crosses at least one quantizer threshold

Manuscript received August 30, 2006; revised August 21, 2008. Current version published April 21, 2010. This work was supported in part by NSF Grants ANI-0112801 and CCF 0329715 and in part by the Center for the Mathematics of Information at California Institute of Technology. Portions of this work were presented at the IEEE International Symposium on Information Theory, Adelaide, Australia, July 2005.

D. Marco was with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA. He is now with Final Israel Ltd., Herzliya, Israel (e-mail: danielmarco@gmail.com).

D. L. Neuhoff is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: neuhoff@eecs.umich.edu).

Digital Object Identifier 10.1109/TIT.2010.2044073

with positive probability, $H(I_1, \dots, I_N) \rightarrow \infty$ as $N \rightarrow \infty$. Therefore, the rate, in bits per second, required to encode the output of the quantizer increases without bound. The main idea of the proof is to show that as $N \rightarrow \infty$, it is possible to obtain from the quantized samples an increasingly accurate approximation of the first time that the process crosses some specified quantization threshold in some finite interval. Since it is shown that such a first crossing occurs with positive probability and that when it occurs it is a random variable with a continuous component, it follows that it has infinite entropy. As a result, the entropy of the quantized samples can be shown to approach infinity as $N \rightarrow \infty$. Some technical issues require handling. For example, it needs to be shown that a first crossing does indeed exist with positive probability. Also, a suitable definition is needed for an approximate first crossing time that can be determined from the quantized samples and that converges in an appropriate sense to the actual first crossing time.

As a concrete indication of the rate of entropy growth, we also show that when a uniform threshold scalar quantizer with infinitely many levels is applied to a stationary Gaussian random process with autocorrelation function $\rho(\tau)$, the conditional entropy of one quantized sample given the previous is, asymptotically

$$H(I_2|I_1) = \left| M \sqrt{1 - \rho(1/N)} \log(1 - \rho(1/N)) \right| (1 + o_N) \quad (1)$$

where M is a constant depending only on the position of the mean in its quantization cell and the ratio of the quantizer step-size to the standard deviation of X , o_N denotes a quantity that goes to zero as N increases, and all logarithms in this paper have base 2. From this, it follows that

$$\begin{aligned} H(I_1, \dots, I_N) &\leq (N-1)H(I_2|I_1) + H(I_1) \\ &= N \left| M \sqrt{1 - \rho(1/N)} \log(1 - \rho(1/N)) \right| \\ &\quad \times (1 + o_N). \end{aligned} \quad (2)$$

For example, if the autocorrelation function of the random process is twice differentiable at the origin (as when the spectrum is Gaussian or bandlimited), then (1) simplifies to

$$H(I_2|I_1) = Mc \frac{\log N}{N} (1 + o_N) \quad (3)$$

and consequently (2) becomes

$$H(I_1, \dots, I_N) \leq (Mc \log N) (1 + o_N) \quad (4)$$

where $c = \sqrt{-\rho''(0)}/2$. If, however, the autocorrelation function approaches the origin with slope having nonzero absolute value s (e.g., $\rho(\tau) = e^{-|\tau|^s}$ with $s = 1$), then

$$H(I_2|I_1) = \frac{M\sqrt{s} \log N}{2\sqrt{N}} (1 + o_N)$$

and in this case, (2) becomes

$$H(I_1, \dots, I_N) \leq \left(\frac{M}{2} \sqrt{Ns} \log N \right) (1 + o_N). \quad (5)$$

We now describe what can be deduced about higher-order conditional entropies. For a Gaussian random process with autocorrelation function that is twice differentiable at the origin, (3) and the fact that $H(I_1, \dots, I_N) \rightarrow \infty$ as $N \rightarrow \infty$ imply that for $L \geq 2$

$$\frac{A_{N,L}}{N} \leq H(I_L | I_{L-1}, \dots, I_1) \leq Mc \frac{\log N}{N} (1 + o_N)$$

where $A_{N,L}$ is a quantity going to infinity as N goes to infinity. That is, for $L \geq 2$, $H(I_L | I_{L-1}, \dots, I_1) \rightarrow \infty$ at a rate greater than $1/N$, but no greater than $(\log N)/N$.

On the one hand, the result that joint entropy increases to infinity is surprising in that it contrasts sharply with the finite rate attainable by ideal lossy source coding. To see this, consider the fact that linear reconstruction of a continuous-time approximation of $X(t)$ from the quantized samples creates a reproduction $\tilde{X}(t)$ with mean-squared error (MSE) distortion, denoted D_N , bounded below by the ordinarily strictly positive distortion, denoted D_w , of a Wiener filter applied to the output of the quantizer operating directly on $X(t)$ (with no sampling). Therefore, a system consisting of a sampler with sampling rate N , scalar quantizer, ideal (sampling-rate-adapted) lossless encoder, corresponding decoder, and linear reconstruction has MSE $D_N \geq D_w > 0$ and rate going to infinity as $N \rightarrow \infty$. In contrast, ideal lossy coding of the samples with MSE D_w requires rate $NR_N(D_w)$ bits/second, where $R_N(\cdot)$ denotes the Shannon rate-distortion function (in bits/sample) of the stationary sampled process $X_n = X(n/N)$. Since as $N \rightarrow \infty$, $NR_N(D_w)$ converges to $\mathcal{R}(D_w)$, the Shannon rate-distortion function of the continuous-time process $X(t)$, and since linear reconstruction of $X(t)$ from the decoded samples can be done with MSE approaching D_w , ideal lossy coding can attain MSE no larger than that of the scalar quantizer based system, with rate remaining bounded rather than tending to infinity as N increases. The infinite gap between ideal lossy source coding and scalar quantization based coding contrasts with the well known fact that when coding discrete-time processes, scalar quantization with entropy coding (i.e., lossless coding at the entropy-rate of the quantizer output) suffers only a small penalty relative to ideal lossy source coding that is bounded above by 0.255 bits/sample at high rates and that does not substantially exceed this at any rate [1]–[3]. Evidently, at high sampling rates, this penalty, multiplied by the sampling rate, approaches infinity.

On the other hand, the increase of entropy with N is consistent with the previous results of which we are aware. For example, Shamai [4] showed that when a binary quantizer with threshold at zero is applied to a random process bandlimited to W Hz, N times the entropy-rate¹ $H_\infty(I)$ of the sampled and quantized process can be as large as,

$${}^1H_\infty(I) = \lim_{N \rightarrow \infty} \frac{1}{N} H(I_1, \dots, I_N)$$

but no larger than, $(2W \log N)(1 + o_N)$ bits/second. Since $NH_\infty(I) \leq H(I_1, \dots, I_N)$, this shows that the latter might grow as $\log N$ for a bandlimited process,² which is the same rate of growth shown in (4) when the quantizer is uniform threshold and the process is Gaussian with an autocorrelation function that is twice differentiable at the origin. Cvetkovic and Vetterli [5], [6] describe a simple scheme for losslessly encoding the output of a sampler and scalar quantizer applied to a bounded and bandlimited deterministic signal. The coding rate of their scheme increases as $\log N$. While this is not a result about the entropy of the quantized samples of a random process, it nonetheless shows a comparable behavior. For the same type of signals, Cvetkovic and Daubechies [7] and Ishwar *et al.* [8] describe schemes for encoding the output of a sampler and dithered scalar quantizer, with rate again increasing as $\log N$ (see also [9] for an earlier discussion of the dithering method, without the low rate encoding of [7] and [8]). Finally, for bounded, nonbandlimited deterministic signals, Kumar *et al.* [10] describe a scheme for coding the output of a sampler and dithered scalar quantizer with rate depending on the tail of the spectrum. For example, the rate grows as $(\log N)^2$ for a signal with an exponentially decaying spectrum. This and (5) suggest that for nonbandlimited processes, entropy can increase more rapidly than $\log N$.

The significance of our result that entropy increases to infinity depends on the fact stated earlier that ordinarily the MSE, D_N , cannot decrease to zero as sampling rate N increases. For if to the contrary some special reconstruction method could make D_N go to zero, then rate increasing to infinity would not be so significant, because it might increase at the pace dictated by the rate-distortion function $\mathcal{R}(D_N)$. Alternatively, the increase of rate would not be so significant if some special reconstruction method made it possible to increase the coarseness of the quantizer with N in such a way that both rate and MSE remain bounded. Accordingly, we now summarize the results of which we are aware on the limiting MSE of linear and nonlinear reconstructions made from quantized samples.

As mentioned earlier, the limiting MSE of linear reconstruction is bounded below by D_w , the MSE of a Wiener filter applied to the output of the quantizer operating directly on $X(t)$, without sampling. While we are not aware of any general result that characterizes those processes for which $D_w > 0$, it seems evident that this holds except in highly specialized situations, for example, when $X(t)$ is a random telegraph process (which switches at random times between $+1$ and -1), and q is a quantizer with a threshold at the origin.

Similarly, for nonlinear reconstructions, MSE can sometimes be made to go to zero and sometimes not, and it is not understood how to characterize the boundary between the two cases. However, there are a handful of examples for which it is known whether MSE tending to zero is possible or not, which we now list. First, if the sample functions of a stationary random process

²Actually, Shamai was interested in the information-rate (in bits per second) between a bandlimited continuous-time process $X(t)$ and the output of a sampler and binary quantizer, which reduces to $NH_\infty(I)$. In this context, our result shows that the information rate in bits/second provided by a sampler and binary quantizer tends to infinity; the only condition being that $X(t)$ crosses the quantizer threshold with positive probability.

are sinusoidal with random phase and known mean, amplitude, and frequency, and if the quantizer has a threshold that is crossed by the sample functions of this process, then it is easy to see that MSE can be made to go to zero with a nonlinear reconstruction that at time t uses the past quantized samples to estimate the times of the two most recent threshold crossings, from which an estimate of the unknown phase is produced. More generally, for bandlimited signals and processes, it is known that MSE can be made to go to zero as N increases provided there are sufficiently many threshold crossings [11]–[13]. However, bandlimitedness is insufficient in and of itself, because Bar-David [13] has described a stationary, bandlimited, zero-mean Gaussian random process for which MSE cannot go to zero when the quantizer is binary with threshold at zero, even with nonlinear reconstruction. The following are two instances of nonbandlimited random processes for which it is known that MSE cannot go to zero with nonlinear reconstruction. Slepian [15] has shown that MSE cannot go to zero for a Gauss-Markov (i.e., Ornstein-Uhlenbeck) process and a binary quantizer with threshold at the mean, and Marco [16] has shown that MSE cannot go to zero for any quantizer and any stationary Markov process that is continuous almost surely and has absolutely continuous second-order distribution.

In conclusion, based on the results of which we are aware, it is our belief that MSE can be made to go to zero only in highly structured cases, such as the random telegraph and sinusoidal examples. Consequently, in most cases, it is significant that entropy increases to infinity as sampling rate grows.

The study of the entropy of scalar quantization plus entropy coding at high sampling rates is relevant to the problem of field gathering with wireless sensor networks [17]–[20], where scalar quantization plus Slepian-Wolf lossless coding is an interesting approach [18] to distributed lossy coding of the samples taken by an array of sensors measuring a field such as the temperature in some geographic region. With Slepian-Wolf coding [21] the quantized samples are encoded using a number of bits equal to their joint entropy. In this context, the fact that entropy tends to infinity indicates that for efficient operation of such a system, the sensors should not be spaced too closely [18]. Evidently, with scalar quantization, the increasing correlation between neighboring sensors is not sufficient to mitigate the increasing number of sensors. In making this assertion, we assume that, as elementary arguments suggest, the joint entropy of N quantized samples taken from a stationary 2-D field will also tend to infinity.

An outline of the paper follows. Section II states and proves the result that entropy tends to infinity as $N \rightarrow \infty$, using the fact that threshold crossing times are asymptotically well determined from the sampled and quantized process. To make this argument precise, such crossing times must be carefully defined. Some needed technical characteristics of random processes (separability, measurability and continuity) are briefly reviewed in Appendix A. Section III states the result on the asymptotic formula for the conditional entropy $H(I_2|I_1)$. This is proved in Section IV, with the proofs of certain lemmas left to Section V, and some details left to Appendices B, C and D. This is a delicate argument that uses the light tails of the Gaussian distribution to show that when samples are closely spaced, the conditional entropy of one quantized sample given the previous

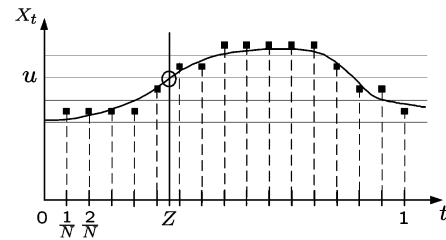


Fig. 1. Sample path of the random process X_t on the interval $[0, 1]$, which is sampled and quantized. u is the quantization threshold considered, Z is the first crossing time of u , and $1/N$ is the sampling interval.

is dominated by just two terms of the defining sum. In addition, this section also provides an asymptotic formula for $H(I_2|I_1)$ when the uniform threshold scalar quantizer cells shrink as N increases. Concluding remarks are offered in Section VI.

II. JOINT ENTROPY OF QUANTIZED SAMPLES AT HIGH SAMPLING RATES

Consider a continuous-time, stationary random process that is sampled every $T = 1/N$ seconds. Each sample is quantized using an arbitrary, yet unchanging, scalar quantizer. In this section, we show that under very mild conditions the joint entropy of the quantized samples in any fixed finite time interval tends to infinity as N goes to infinity. Specifically, we assume that the process is continuous in probability and that it crosses some quantizer threshold with positive probability. The process need not be bandlimited.

As mentioned in the introduction, the key idea³ is to approximate, using the quantized samples (ever more accurately, as $N \rightarrow \infty$), the first time the process crosses the specified quantization threshold in the given time interval. Since this first time is a random variable with a continuous component, it has infinite entropy, thus implying the entropy of the quantized samples tends to infinity as $N \rightarrow \infty$. Fig. 1 illustrates a sample path, its samples, a threshold, and the time of first threshold crossing. While the basic idea is fairly simple, it does involve certain technical hurdles that need to be overcome. Specifically, the following are needed.

1. Some random process regularity conditions, such as separability, that insure the validity of elementary steps of the derivation.
2. A useful definition of a quantizer-threshold first-crossing time that can be approximated from quantized samples.
3. A proof that for any specified finite time interval, the event, denoted F , that a *first* threshold crossing occurs in that interval has positive probability.
4. The identification of an event $B \subset F$ with positive probability, with the property that when B occurs, the first crossing time is a well behaved continuous random variable, and which facilitates the next item.
5. The definition of an approximate first crossing time that can be determined from the quantized samples, and a proof that when B occurs, it converges in probability to the first crossing time.

³The authors thank Bruce Hajek for this idea.

6. A proof that 3–5 above imply that the entropy of the approximate first crossing time, and consequently, the entropy of the quantized samples, converge to infinity.

We now state the random process assumptions and regularity conditions we require, definitions of threshold crossing times, and the main result of this section along with its proof. Appendix A provides definitions and a brief review of the regularity conditions.

The continuous-time random process is represented as the collection of random variables $\{X_t(\omega) : t \in (-\infty, \infty), \omega \in \Omega\}$ defined on the complete probability space (Ω, \mathcal{F}, P) . Let X_t denote the random variable at time t , let $X(\omega)$ denote the sample path corresponding to ω , and let X be a shorthand for the entire random process.

We assume throughout that X is stationary and continuous in probability, which are conditions determined by the finite-dimensional distributions of X . We also assume that the specific functional representation $X_t(\omega)$ of X is separable and measurable, which continuity in probability permits one to assume. These conditions insure that a set of the form $\{\omega : X_t(\omega) \leq r, t_1 < t < t_2\}$ is an event in \mathcal{F} with a probability depending only on r and $t_2 - t_1$, and that expected values and time integrals can be exchanged in appropriate circumstances.

A quantizer is characterized by a finite or countably infinite, strictly increasing sequence of thresholds $\{\tau_i\}$, a corresponding partition $\{S_i\}$ of \mathbb{R} into intervals, called cells, such that the left and right boundaries of the cell S_i are τ_i and τ_{i+1} , respectively, and a quantization rule $q : \mathbb{R} \rightarrow \mathbb{Z}$ defined by $q(x) = i$, when $x \in S_i$. The quantizer will be denoted q .

Definition 1: (inspired by [22, p. 146ff]): Let q be a quantizer, and let $h : \mathbb{R} \rightarrow \mathbb{R}$ be some function. Then h has a q -threshold-crossing if there exists $t, \delta > 0$, and quantization cells S_i, S_j such that $h(s) \in S_i, t - \delta < s < t$, and $h(s) \in S_j, t < s < t + \delta$. When δ is specified, it is said to be a (q, δ) -threshold-crossing. When t is specified, the threshold crossing is said to be at t .

It can happen that a function h has a number of q -threshold crossings in a finite interval such as $[0, 1]$, but no *first* crossing. For example, suppose q has just one threshold at 0, and $h(t) = \sin(1/t)$.

The main result of this section is the following.

Theorem 1: Let X be a continuous-time, stationary and continuous in probability random process. Let q be a quantizer such that X has a q -threshold-crossing with positive probability. Given a positive integer N , let $I_k^N(\omega) = q(X_{k/N}(\omega))$, $k = 0, \dots, N - 1$, denote the N quantized samples in the time interval $[0, 1]$. Then

$$\lim_{N \rightarrow \infty} H(I_0^N, I_1^N, \dots, I_{N-1}^N) = \infty.$$

Though for simplicity, the theorem is stated for the time interval $[0, 1]$, it holds equally for any interval. As suggested earlier, the proof proceeds by showing that from the N quantized samples, one can approximately compute the time of the first quantizer threshold crossing in the time interval $(0, 1)$, and then showing that the approximation, and consequently the quan-

tized samples from which they are computed, have entropy increasing to infinity. To show that the approximation has entropy approaching infinity, we will use the following.

Lemma 2: Let Z and \hat{Z}^N , $N \in \mathbb{Z}_+$, be real-valued random variables defined on (Ω, \mathcal{F}, P) such that each \hat{Z}^N is discrete, and there is an event $B \in \mathcal{F}$ such that (a) $P(B) > 0$, (b) for every $s > 0$, $\Pr(|Z - \hat{Z}^N| < s|B) \rightarrow 1$ as $N \rightarrow \infty$, and (c) for every $\varepsilon > 0$, there exists $s > 0$ such that $\Pr(|Z - z| < s|B) < \varepsilon$ for all z . Then $H(\hat{Z}^N) \rightarrow \infty$ as $N \rightarrow \infty$.

Proof: To show that entropy goes to infinity, we will show that when N is large, all outcomes of \hat{Z}^N have uniformly small conditional probabilities. Specifically, we now show that for any $\varepsilon > 0$, $\Pr(\hat{Z}^N = z|B) < \varepsilon$ for all sufficiently large N and all z . Accordingly, given $\varepsilon > 0$, choose $s_\varepsilon > 0$ such that $\Pr(|Z - z| < s_\varepsilon|B) < \varepsilon/2$ for all z , and choose N_o such that $\Pr(|Z - \hat{Z}^N| \geq s_\varepsilon|B) < \varepsilon/2$ for all $N \geq N_o$. Then for $N \geq N_o$ and any z

$$\begin{aligned} \Pr(\hat{Z}^N = z|B) &= \Pr(\hat{Z}^N = z, |Z - \hat{Z}^N| < s_\varepsilon|B) \\ &\quad + \Pr(\hat{Z}^N = z, |Z - \hat{Z}^N| \geq s_\varepsilon|B) \\ &\leq \Pr(|Z - z| < s_\varepsilon|B) + \Pr(|Z - \hat{Z}^N| \geq s_\varepsilon|B) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

We now bound the entropy of \hat{Z}^N

$$\begin{aligned} H(\hat{Z}^N) &\geq H(\hat{Z}^N|B)P(B) \\ &= P(B) \sum_z \Pr(\hat{Z}^N = z|B) \log \frac{1}{\Pr(\hat{Z}^N = z|B)} \\ &\geq P(B) \sum_z \Pr(\hat{Z}^N = z|B) \log \frac{1}{\varepsilon} \\ &= P(B) \log \frac{1}{\varepsilon} \end{aligned}$$

where the second inequality holds whenever N is sufficiently large. Taking the limit as $N \rightarrow \infty$ and then letting $\varepsilon \rightarrow 0$ shows that $\lim_{N \rightarrow \infty} H(\hat{Z}^N) = \infty$, and completes the proof. \square

Proof of Theorem 1: We will apply Lemma 2. The fact that there is a q -threshold-crossing with positive probability permits us to fix $\delta > 0$ such that the probability of a (q, δ) -threshold-crossing is positive. (If the probability of a (q, δ) -threshold-crossing were zero for all $\delta > 0$, then the probability of a q -threshold-crossing would be zero.) Moreover, since the process is defined on the whole real line, for any finite interval (t_1, t_2) , the probability of a (q, δ) -threshold-crossing occurring in that interval is positive (if not, then by stationarity the probability of a (q, δ) -threshold-crossing in each interval of the form $(\frac{m}{2}(t_2 - t_1), (\frac{m}{2} + 1)(t_2 - t_1))$, with m an integer, would be zero, and the union bound would show that the probability of a (q, δ) -threshold-crossing anywhere in $(-\infty, \infty)$ is zero).

We are now ready to define Z and \widehat{Z}^N to be, respectively, the first threshold-crossing time and its approximation mentioned earlier. Specifically

$$Z(\omega) = \begin{cases} t, & X(\omega) \text{ has a first } q\text{-threshold-crossing in } (0, 1), \\ & \text{it is at } t, \text{ and it is a } (q, \delta) \text{ crossing.} \\ 2, & \text{otherwise} \end{cases}$$

and

$$\widehat{Z}^N(\omega) = \frac{K^N(\omega) + \frac{1}{2}}{N}$$

where $K^N(\omega)$ is the smallest integer $k \in \{0, 1, \dots, N - 1\}$ such that $I_k^N(\omega) \neq I_{k+1}^N(\omega)$, or $3N$ if no such k exists.

Let B in Lemma 2 be the event that a (q, δ) -threshold-crossing occurs in the time interval $(\delta/4, \delta/2)$. The above argument establishes $P(B) > 0$, which is condition (a) of the lemma. To check condition (b), we observe that when B occurs, there can be no earlier threshold-crossings in $(0, 1)$; so this is the first threshold-crossing in $(0, 1)$. Hence, Z equals the time of this threshold-crossing. Moreover, when $1/N < \delta/4$, this threshold-crossing will necessarily be reflected in the quantized samples. That is, the sample just to the left of the threshold crossing ($1/N < \delta/4$ insures there is such a sample) and the sample just to the right will be the first pair of consecutive samples to lie in different quantization cells.⁴ It follows that when B occurs and $1/N < \delta/4$, $|Z - \widehat{Z}^N| < 1/N$. Hence, for any ε , when $1/N < \min\{\varepsilon, \delta/4\}$, $\Pr(|Z - \widehat{Z}^N| < \varepsilon|B) = 1$, so condition (b) holds.

Finally, to verify condition (c), we observe that for any z and any s

$$\Pr(|Z - z| < s|B) \leq \frac{\Pr(|Z - z| < s)}{P(B)}.$$

The numerator of the right-hand expression is bounded from above by $\Pr(\text{any } (q, \delta)\text{-threshold-crossing (not necessarily a first crossing) occurs in } (z - s, z + s))$. By stationarity, this is the same for all values of z . As $s \rightarrow 0$, it converges to $\Pr((q, \delta)\text{-threshold-crossing occurs at } t = 0)$, which in turn is zero because it equals the expected fraction of time at which a (q, δ) -threshold-crossing occurs, which is zero, because (q, δ) -threshold-crossings can occur at most a countable number of times in any sample path (One can formalize this argument [23] using the indicator function of the event that a (q, δ) -threshold-crossing occurs at time t . Such an argument uses the measurability of the process to permit an expectation and time integral to be interchanged). In conclusion, we have that for every $\varepsilon > 0$, there exists $s > 0$ such that $\Pr(|Z - z| < s) < \varepsilon P(B)$ for all z , and, consequently, $\Pr(|Z - z| < s|B) < \varepsilon$ for all z , which establishes condition (c).

Lemma 2 now implies that $H(\widehat{Z}^N) \rightarrow \infty$. Since \widehat{Z}^N is entirely determined by I_0^N, \dots, I_{N-1}^N

$$\lim_{N \rightarrow \infty} H(I_0^N, \dots, I_{N-1}^N) \geq \lim_{N \rightarrow \infty} H(\widehat{Z}^N) = \infty$$

⁴If there is a sample directly at the threshold-crossing time, then it will be this sample plus the one either to the right or left that forms the first consecutive pair to lie in different cells.

which completes the proof of Theorem 1. □

We note that the process in Theorem 1 is required to be continuous in probability so as to guarantee (along with stationarity) the preservation of probability under shifting and to permit exchange of integration and expectation.

III. ASYMPTOTIC FORMULA FOR CONDITIONAL ENTROPY FOR A GAUSSIAN SOURCE

Let X_1 and X_2 denote jointly Gaussian random variables with mean μ , variance σ^2 , and correlation coefficient $\rho \leq 1$. Let q denote the quantization rule of an infinite-level uniform threshold scalar quantizer with step size Δ , offset $\theta \in [0, 1]$,⁵ and cells $S_k = [t_k, t_{k+1})$, $k \in \mathbb{Z}$, where $t_k \triangleq (k - \theta)\Delta$, is the left threshold of the k th cell. Let $I_1 = q(X_1)$ and $I_2 = q(X_2)$, where $q(x) = k$ if $x \in S_k$. Let $\lambda = \frac{\Delta}{\sigma}$.

The following is the main result of this section.

Theorem 3: Suppose $\mu \in S_j$, $j \in \mathbb{Z}$. Let $\theta^* = \frac{\mu - (j - \theta)\Delta}{\Delta}$. Then

$$\lim_{\rho \rightarrow 1} \frac{H(I_2|I_1)}{-M_\lambda^{\theta^*} \sqrt{1 - \rho} \log \sqrt{1 - \rho}} = 1$$

where

$$M_\lambda^{\theta^*} = \frac{\sqrt{2}}{\pi} \sum_{k=-\infty}^{\infty} e^{-\frac{(k - \theta^*)^2 \lambda^2}{2}}$$

is a positive constant that depends on the ratio $\frac{\Delta}{\sigma}$ and on the position of the mean μ in its cell.

This theorem shows that when ρ is close to one, $H(I_2|I_1)$ factors into a term depending on ρ and another depending on λ and θ . We show in Appendix B that $M_\lambda^\theta \rightarrow \frac{2}{\sqrt{\pi}\lambda}$ uniformly in θ as $\lambda \rightarrow 0$. Numerical calculations show that when $\lambda \leq 1$ (i.e., $\frac{\Delta}{\sigma} \leq 1$), the approximation $M_\lambda^\theta \approx \frac{2}{\sqrt{\pi}\lambda}$ is accurate to within fraction 10^{-8} of the true value for all $\theta \in [0, 1]$. Numerical calculations also indicate that as $\rho \rightarrow 1$, the convergence of conditional entropy to the asymptotic expression is fairly slow.

Next, consider a generalization of the asymptotic formula for conditional entropy in Theorem 3 for the situation that the quantizer cells shrink as the correlation between the input random variables increases. This is relevant to situations where it is desired to attain smaller MSE as correlation increases. For example, the scheme described in the introduction for field gathering with a wireless sensor network (involving fixed scalar quantizers and linear reconstruction) has rate in (bits per unit area) increasing to infinity and distortion bounded away from zero, as sensor density increases. However, it might be desired to have distortion go to zero as sensor density increases, which can be attained using quantizers whose cells shrink to zero as sensor density increases. Clearly, such a scheme has larger rate than its nonshrinking cells counterpart.

Theorem 4: Let X_1 and X_2 be quantized with an infinite-level uniform threshold quantizer with step size to standard deviation ratio $\lambda(\rho)$ that is any positively valued function that tends to

⁵ θ need only assume values in $[0, 1)$ rather than $[0, 1]$. However, allowing $\theta = 1$ facilitates symmetry arguments.

zero as $\rho \rightarrow 1$ and satisfies (a) $\lim_{\rho \rightarrow 1} \frac{\lambda^2(\rho)}{\mathcal{H}(1-\rho)} = \infty$, and (b) $\liminf_{\rho \rightarrow 1} \frac{(\log \frac{1}{1-\rho})^3}{\log \lambda(\rho)} > 2$. Then

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

An example of a possible $\lambda(\rho)$ is $\lambda(\rho) = \frac{1}{\log \frac{1}{1-\rho}}$. Notice that the factor that multiplies $\sqrt{1-\rho} \log \sqrt{1-\rho}$ in the denominator above is similar to the limiting value of M_λ^θ divided by $\lambda(\rho)$. This gives rise to the observation that the quantizer cells can be shrunk in two different ways—either, (a) by fixing λ , letting ρ go to 1, then letting λ go to 0, as in the discussion after Theorem 3, or (b) by letting λ be a function of ρ that decreases to zero as ρ goes to 1, as in Theorem 4.

IV. PROOFS OF THEOREM 3 AND THEOREM 4

Notation: We begin by introducing some notation to be used throughout. Let $P_k^\theta \triangleq \Pr(I_i = k)$, $i \in \{1, 2\}$, and $P_{l|k}^\theta \triangleq \Pr(I_2 = l | I_1 = k)$. For brevity, when it is clear from context, we drop the superscript θ from all functions that depend on it, such as $P_{l|k}$, P_k , t_k , etc. Note that $P_{l|k}$ depends on ρ , but we do not show this explicitly.

Let \mathcal{H} denote the entropy function, i.e., $\mathcal{H}(\dots, z_{-1}, z_0, z_1, \dots) = \sum_{k=-\infty}^{\infty} -z_k \log z_k$, where the z_k 's are a finite or countably infinite set of nonnegative numbers that need not sum to one, and where $-0 \log 0$ is taken to be 0. All logarithms are base 2. Let also $H_{l|k} = \mathcal{H}(P_{l|k})$, and for a set L let $H_{L|k} = \sum_{l \in L} H_{l|k}$. Thus, for example, with this notation

$$\begin{aligned} H(I_2|I_1 = k) &= \sum_{l=-\infty}^{\infty} H_{l|k} = H_{\{\dots, -1, 0, 1, \dots\}|k} \\ &= \sum_{l=-\infty}^{\infty} \mathcal{H}(P_{l|k}). \end{aligned}$$

When the set L has few elements (two or three), we abuse notation and omit the curly brackets. For instance, we write $H_{k-1,k,k+1|k}$ instead of $H_{\{k-1,k,k+1\}|k}$. Let also $H_q(f) = \mathcal{H}(\dots, P_{-1}(f), P_0(f), P_1(f), \dots)$, where $P_i(f) = \int_{S_i} f(x) dx$. Note that f need not be a probability density function (pdf), but it will be nonnegative.

Let $\mathcal{N}_{\mu, \sigma^2}(x)$ denote a Gaussian pdf with mean μ and variance σ^2 . Let

$$\sigma_\rho^2 \triangleq \sigma^2 (1 - \rho^2)$$

be the conditional variance of X_2 given X_1 . Let $Q(\alpha) \triangleq \int_\alpha^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ denote the standard ‘‘Q function’’. Elementary facts about the Q and \mathcal{H} functions that will be useful throughout are provided in Appendix C. We refer to them as Facts C1–C9. Let

$$g(\alpha) \triangleq \frac{Q(\alpha)}{\frac{1}{2} e^{-\frac{\alpha^2}{2}}}, \alpha \geq 0$$

and refer to it as the ‘‘correction factor’’ to the upper bound $\frac{1}{2} e^{-\frac{\alpha^2}{2}}$ of the Q function (see Fact C1 in Appendix C). A useful elementary fact about g is also provided in Appendix C.

Proof of Theorem 3: Without loss of generality let the means μ of X_1 and X_2 be zero, so that $\theta^* = \theta$. The key to finding a simple asymptotic expression for conditional entropy

$$\begin{aligned} H(I_2|I_1) &= \sum_{k=-\infty}^{\infty} H(I_2|I_1 = k) P_k \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \mathcal{H}(P_{l|k}) P_k \end{aligned}$$

is to truncate the summations, eliminating terms that are asymptotically negligible when ρ approaches one. As one step, we will specify an integer $K(\rho)$, and show that truncating the outer summation to $k \in \{-K(\rho), \dots, K(\rho)\}$ has, asymptotically, no effect.

Now consider truncating the inner summation, $H(I_2|I_1 = k) = \sum_{l=-\infty}^{\infty} \mathcal{H}(P_{l|k})$, $|k| \leq K(\rho)$. We begin by noting the similarity to the problem considered in [25], where it was shown that when quantizing a Gaussian random variable X with an infinite uniform scalar quantizer q with quantizer step size that grows large relative to the variance of the Gaussian variable, the entropy $H(q(X)) = \sum_{k=-\infty}^{\infty} \mathcal{H}(P_k)$ is dominated by just the two terms $\mathcal{H}(P_{\text{left}})$ and $\mathcal{H}(P_{\text{right}})$, corresponding to the cells flanking the cell containing the mean. The present problem is analogous in that the variance of the quantized variable, i.e., of X_2 given X_1 , goes to zero as $\rho \rightarrow 1$; so, it becomes small relative to the quantizer step size. Nonetheless, there is a significant difference. Here, we are concerned with the conditional entropy of $I_2 = q(X_2)$ given $I_1 = q(X_1)$, the quantized version of X_1 , rather than X_1 itself. Because of this, the conditional distribution of X_2 is not Gaussian, so the analysis of [25] does not apply.

To see how to proceed, consider an expression for the conditional pdf of X_2 given $I_1 = k$

$$\begin{aligned} f_{X_2|I_1}(x|k) &= \int_{t_k}^{t_{k+1}} f_{X_2|X_1}(x|y) \frac{f_{X_1}(y)}{P_k} dy \\ &= \frac{1}{P_k} \int_{t_k}^{t_{k+1}} \frac{1}{\sqrt{2\pi}\sigma_\rho} e^{-\frac{(x-\rho y)^2}{2\sigma_\rho^2}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy. \end{aligned} \quad (9)$$

One can see that $f_{X_2|I_1}(x|k)$ is a weighted average of conditional densities of X_2 given that X_1 equals particular values in S_k , where the weighting function is Gaussian. These conditional densities are Gaussian with mean ρ times the value of X_1 , and variance $\sigma_\rho^2 = \sigma^2(1 - \rho^2)$. Therefore, when ρ is close to one, these conditional densities are very narrow, with mean ρ times the given value of X_1 . Thus, when ρ is close to one, $f_{X_2|I_1}(x|k)$ concentrates on the interval $(\rho t_k, \rho t_{k+1})$, as illustrated in Fig. 2. Now, continuing to assume ρ is close to one, consider the effect of k , also illustrated in Fig. 2. On the one hand, when $|k|$ is not too large (specifically, when $(1 - \rho)|k| \ll 1$ so that $|\rho t_k - t_k| \ll \Delta$), $f_{X_2|I_1}(x|k)$ concentrates on the k th cell, i.e., the cell containing X_1 . On the other hand, when $|k|$ is much larger, the interval $(\rho t_k, \rho t_{k+1})$ on which $f_{X_2|I_1}(x|k)$ concentrates is disjoint from the k th cell. (We also note, though we do not exploit this, that when $|k|$ is large, $f_{X_2|I_1}(x|k)$ becomes very narrow, due to the fact the Gaussian weighting function concentrates at the endpoint closest to the origin of the k th cell.)

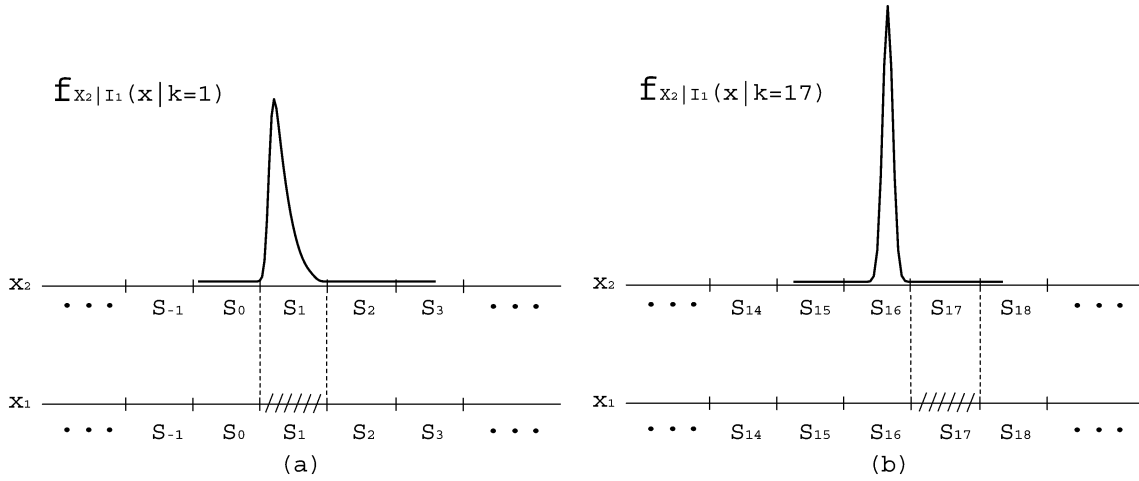


Fig. 2. Conditional pdf of X_2 given that X_1 lies in the k th quantization cell. The parameters are $\theta = \frac{1}{2}$, $\mu = 0$, $\Delta = 2$, $\sigma = 1$, $\rho = 0.99$, and (a) $k = 1$, (b) $k = 17$.

The fact that $f_{X_2|I_1}(x|k)$ does not approximately concentrate on one quantization cell when $|k|$ is large significantly complicates the approximation of $H(I_2|I_1 = k)$ for such k . Fortunately, in addition to choosing $K(\rho)$ large enough that the outer sum can be truncated to the terms $k \in \{-K(\rho), \dots, K(\rho)\}$, we are also able to choose $K(\rho)$ small enough that when $|k| \leq K(\rho)$, $(1 - \rho)|k|$ is sufficiently small that $f_{X_2|I_1}(x|k)$ concentrates so much on the k th cell that $H(I_2|I_1 = k)$ is dominated by just the two terms $\mathcal{H}(P_{k-1|k})$ and $\mathcal{H}(P_{k+1|k})$. More precisely, letting $x(\rho) \rightarrow a$ means that the convergence happens as $\rho \rightarrow 1$, the competing requirements on $K(\rho)$ are as follows.

- 1: $K(\rho)\sqrt{1 - \rho} \rightarrow 0$, required by Lemmas 7–11, 14, 15.
- 2: $e^{\frac{K^2(\rho)\lambda^2}{2}}\sqrt{1 - \rho} \rightarrow \infty$, required by Lemma 13
- 3: $\frac{K(\rho)}{\log \frac{1}{1 - \rho}} \rightarrow 0$, required by Lemma 15
- 4: $K(\rho) \rightarrow \infty$, required by Lemma 16

If we consider $K(\rho)$ to be of the form $\left(\ln \frac{1}{1 - \rho}\right)^r$, then in order for Item 2 to hold for any λ , it must be that $r > \frac{1}{2}$. Similarly, for Item 3 to hold, it is necessary that $r < 1$. For $\frac{1}{2} < r < 1$, Items 1 and 4 are easily seen to be satisfied. Therefore, we choose

$$K(\rho) = \left\lceil \left(\ln \frac{1}{1 - \rho} \right)^{\frac{3}{4}} \right\rceil.$$

We conclude this proof introduction by mentioning that another key step is the determination of tractable approximations to $f_{X_2|I_1}(x|k)$.

Let $H^\theta(I_2|I_1 = k)$ denote the conditional entropy when the quantizer has offset θ . Define

$$\tilde{H}^\theta(I_2|I_1) = \sum_{k=1}^{\infty} H^\theta(I_2|I_1 = k)P_k^\theta + \frac{H^\theta(I_2|I_1 = 0)P_0^\theta}{2}.$$

We observe that due to the symmetry of the Gaussian pdf and of the uniform quantizers, $P_{l|k}^\theta = P_{-l|-k}^{1-\theta}$ and $P_k^\theta = P_{-k}^{1-\theta}$. There-

fore, it is not hard to see that $H^\theta(I_2|I_1 = k) = H^{1-\theta}(I_2|I_1 = -k)$. Hence

$$H(I_2|I_1) = \tilde{H}^\theta(I_2|I_1) + \tilde{H}^{1-\theta}(I_2|I_1). \quad (7)$$

The proof proceeds by finding an asymptotic expression for $\tilde{H}^\theta(I_2|I_1)$, which only requires focusing on nonnegative k 's. Specifically, letting

$$W_\lambda^\theta = \frac{\sqrt{2}}{\pi} \sum_{k=0}^{\infty} e^{-\frac{(k+1-\theta)^2\lambda^2}{2}} + \frac{\sqrt{2}}{4\pi} \left(e^{-\frac{\theta^2\lambda^2}{2}} - e^{-\frac{(1-\theta)^2\lambda^2}{2}} \right)$$

we show that each term on the right-hand side of the following equation converges to one as $\rho \rightarrow 1$:

$$\begin{aligned} & \frac{\tilde{H}^\theta(I_2|I_1)}{W_\lambda^\theta \mathcal{H}(\sqrt{1 - \rho})} \\ &= \frac{\tilde{H}^\theta(I_2|I_1)}{\sum_{k=1}^{K(\rho)} H(I_2|I_1 = k)P_k + \frac{1}{2}H(I_2|I_1 = 0)P_0} \\ & \times \frac{\sum_{k=1}^{K(\rho)} H(I_2|I_1 = k)P_k + \frac{1}{2}H(I_2|I_1 = 0)P_0}{\sum_{k=1}^{K(\rho)} (H_{k-1,k,k+1|k}) P_k + \frac{1}{2} (H_{-1,0,1|0}) P_0} \\ & \times \frac{\sum_{k=1}^{K(\rho)} (H_{k-1,k,k+1|k}) P_k + \frac{1}{2} (H_{-1,0,1|0}) P_0}{\sum_{k=1}^{K(\rho)} (H_{k-1,k+1|k}) P_k + \frac{1}{2} (H_{-1,1|0}) P_0} \\ & \times \frac{\sum_{k=1}^{K(\rho)} (H_{k-1,k+1|k}) P_k + \frac{1}{2} (H_{-1,1|0}) P_0}{\mathcal{H}(\sqrt{1 - \rho^2}) \left[\sum_{k=1}^{K(\rho)} (W_{L,R,\lambda}^\theta(k)) + \frac{1}{2} (W_{L,R,\lambda}^\theta(0)) \right]} \\ & \times \frac{\mathcal{H}(\sqrt{1 - \rho^2}) \left[\sum_{k=1}^{K(\rho)} (W_{L,R,\lambda}^\theta(k)) + \frac{1}{2} (W_{L,R,\lambda}^\theta(0)) \right]}{W_\lambda^\theta \mathcal{H}(\sqrt{1 - \rho})} \end{aligned} \quad (8)$$

where we define

$$W_{L,R,\lambda}^\theta(k) = \frac{1}{2\pi} e^{-\frac{(k-\theta)^2\lambda^2}{2}} + \frac{1}{2\pi} e^{-\frac{(k+1-\theta)^2\lambda^2}{2}}$$

or otherwise written, $W_{L,R,\lambda}^\theta(k) = W_{L,\lambda}^\theta(k) + W_{R,\lambda}^\theta(k)$, where $W_{L,\lambda}^\theta(k) = \frac{1}{2\pi} e^{-\frac{(k-\theta)^2 \lambda^2}{2}}$ and $W_{R,\lambda}^\theta(k) = \frac{1}{2\pi} e^{-\frac{(k+1-\theta)^2 \lambda^2}{2}}$. Once (8) is established, the theorem follows using (7) and the easily seen fact that $W_\lambda^\theta + W_\lambda^{1-\theta} = M_\lambda^\theta$.

Before proceeding further, we investigate and establish additional properties of $f_{X_2|I_1}(x|k)$, $P_{l|k}$ and other important quantities that we will need. Lemmas will be stated as needed; their proofs will be provided in Section V, or in Appendix D. We begin by finding approximate expressions for the conditional pdf

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

For tractability we would like to drop the smaller of the two Q function terms in the above expression, which leads us to the upper bound

$$\begin{aligned} f_{X_2|I_1}(x|k) &\leq \tilde{f}_{X_2|I_1}(x|k) \triangleq \\ &\begin{cases} \frac{1}{P_k} f_{X_2}(x) Q\left(\frac{t_k - \rho x}{\sigma_\rho}\right), & x < \frac{t_k}{\rho} \\ \frac{1}{P_k} f_{X_2}(x) \left[Q\left(\frac{t_k - \rho x}{\sigma_\rho}\right) - Q\left(\frac{t_{k+1} - \rho x}{\sigma_\rho}\right) \right], & \frac{t_k}{\rho} \leq x \leq \frac{t_{k+1}}{\rho} \\ \frac{1}{P_k} f_{X_2}(x) Q\left(\frac{\rho x - t_{k+1}}{\sigma_\rho}\right), & x > \frac{t_{k+1}}{\rho}. \end{cases} \quad (9) \end{aligned}$$

The following lemma shows that $\tilde{f}_{X_2|I_1}(x|k)$ is an asymptotically tight upper bound to $f_{X_2|I_1}(x|k)$.

Lemma 5:

$$\frac{f_{x_2|I_1}(x|k)}{\tilde{f}_{x_2|I_1}(x|k)} \longrightarrow 1 \text{ as } \rho \longrightarrow 1, \text{ uniformly in } k \text{ and } x$$

and, consequently, for any $0 \leq \gamma < 1$, there exists $\rho_\gamma < 1$ such that for all $\rho > \rho_\gamma$, $f_{X_2|I_1}(x|k) \geq \gamma \tilde{f}_{X_2|I_1}(x|k)$ for all k and x .

We will find it useful to use the following alternate representation of $\tilde{f}_{X_2|I_1}(x|k)$

$$\begin{aligned} \tilde{f}_{X_2|I_1}(x|k) &= \\ &\begin{cases} L_k(\rho) \mathcal{N}_{t_k, \rho, \sigma_\rho^2}(x) g\left(\frac{t_k - \rho x}{\sigma_\rho}\right), & x < \frac{t_k}{\rho} \\ \frac{\mathcal{N}_{0, \sigma^2}(x)}{P_k} - L_k(\rho) \mathcal{N}_{t_k, \rho, \sigma_\rho^2}(x) g\left(\frac{\rho x - t_k}{\sigma_\rho}\right) \\ - R_k(\rho) \mathcal{N}_{t_{k+1}, \rho, \sigma_\rho^2}(x) g\left(\frac{t_{k+1} - \rho x}{\sigma_\rho}\right), & \frac{t_k}{\rho} \leq x \leq \frac{t_{k+1}}{\rho} \\ R_k(\rho) \mathcal{N}_{t_{k+1}, \rho, \sigma_\rho^2}(x) g\left(\frac{\rho x - t_{k+1}}{\sigma_\rho}\right), & x > \frac{t_{k+1}}{\rho} \end{cases} \quad (10) \end{aligned}$$

where

$$L_k(\rho) \triangleq \frac{1}{2} \frac{1}{P_k} e^{-\frac{t_k^2}{2\sigma^2}} \sqrt{1 - \rho^2}$$

and

$$R_k(\rho) \triangleq \frac{1}{2} \frac{1}{P_k} e^{-\frac{t_{k+1}^2}{2\sigma^2}} \sqrt{1 - \rho^2}.$$

Equation (10) is obtained by manipulating exponentials. For example, for the region $x < \frac{t_k}{\rho}$

$$\begin{aligned} &\frac{1}{P_k} f_{X_2}(x) Q\left(\frac{t_k - \rho x}{\sigma_\rho}\right) \\ &= \frac{1}{P_k} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \frac{1}{2} e^{-\frac{(t_k - \rho x)^2}{2\sigma_\rho^2}} g\left(\frac{t_k - \rho x}{\sigma_\rho}\right) \\ &= \left(\frac{1}{2} \frac{1}{P_k} e^{-\frac{t_k^2 - t_k^2 \rho^2}{2\sigma^2(1-\rho^2)}} \sqrt{1 - \rho^2} \right) \\ &\quad \times \left(\frac{1}{\sqrt{2\pi}\sigma_\rho} e^{-\frac{x^2 - 2t_k \rho x + t_k^2 \rho^2}{2\sigma^2(1-\rho^2)}} \right) g\left(\frac{t_k - \rho x}{\sigma_\rho}\right) \\ &= L_k(\rho) \mathcal{N}_{t_k, \rho, \sigma_\rho^2}(x) g\left(\frac{t_k - \rho x}{\sigma_\rho}\right). \end{aligned}$$

Similar algebraic steps can be used for the regions $\frac{t_k}{\rho} \leq x \leq \frac{t_{k+1}}{\rho}$ and $x > \frac{t_{k+1}}{\rho}$. The representation of $\tilde{f}_{X_2|I_1}(x|k)$ in (10) will be useful since it will turn out that $P_{k-1|k}$ and $P_{k+1|k}$ converge to multiples of $L_k(\rho)$ and $R_k(\rho)$, respectively. This motivates us to further examine these quantities. Specifically, the following lemma provides upper and lower bounds to $L_k(\rho)$ and $R_k(\rho)$.

Lemma 6: For $|\rho| < 1$:

- A. $L_k(\rho) < \sqrt{1 - \rho^2} \sqrt{2\pi} k \lambda$ for $k > K_\lambda$
- B. $L_k(\rho) > \sqrt{1 - \rho^2}$ for $k \geq 1$
- C. $R_k(\rho) < \sqrt{1 - \rho^2} \sqrt{\frac{\pi}{\lambda}} e^{\frac{\lambda^2}{2}}$ for $k \geq 0$
- D. $R_k(\rho) > \sqrt{1 - \rho^2} e^{-k\lambda^2} e^{-\frac{\lambda^2}{2}}$ for $k \geq 1$

where $K_\lambda = 1 + \max\{\frac{2}{\lambda}, \frac{2}{\lambda^2}\}$.

Next, when considering the second, third and fourth terms on the right-hand side of (8), $P_{k-1|k}$ and $P_{k+1|k}$ for $0 \leq k \leq K(\rho)$, will play a key role; hence, we shall find expressions for these. (It will turn out that finding an expression for $P_{k|k}$ can be avoided.) From Lemma 5, it will follow that it will be enough to use $\tilde{f}_{X_2|I_1}(x|k)$ rather than $f_{X_2|I_1}(x|k)$, i.e., $\int_{t_{k+1}}^{t_{k+2}} \tilde{f}_{X_2|I_1}(x|k) dx$ will be shown to be a sufficiently good approximation to $P_{k+1|k}$ as $\rho \rightarrow 1$, and similarly it will follow that $\int_{t_{k-1}}^{t_k} \tilde{f}_{X_2|I_1}(x|k) dx$ is a sufficiently good approximation to $P_{k-1|k}$ as $\rho \rightarrow 1$. However, since we are considering $k \geq 0$, one can see from (9) and (10) that the expression for $\tilde{f}_{X_2|I_1}(x|k)$ in the interval $[\frac{t_{k+1}}{\rho}, \frac{t_{k+2}}{\rho}]$ is rather complicated. Therefore, for analytical tractability, we define

$$P_{k+1|k}^* = \int_{\frac{t_{k+1}}{\rho}}^{\frac{t_{k+2}}{\rho}} f_{X_2|I_1}(x|k) dx \quad (11)$$

where we observe that for ρ sufficiently large, $\frac{t_{k+1}}{\rho} < t_{k+2}$; thus, P^* is evaluated on a subset of cell $(k+1)$. Hence, it is easy to see that $P_{k+1|k}^* < P_{k+1|k}$ for all $k \geq 0$. We will show, however, that $P_{k+1|k}^*$ is a sufficiently good approximation of

$P_{k+1|k}$, and consequently, we will be able to use $P_{k+1|k}^*$ instead of $P_{k+1|k}$.

We observe that since $K(\rho)$ grows as $\rho \rightarrow 1$, the number of summands in the numerators and denominators of the second and third terms in (8) grows as well. Consequently, when making convergence statements regarding $P_{k-1|k}$, $P_{k+1|k}$ and $P_{k+1|k}^*$, we will require a special kind of uniform convergence. Thus, we define the following limit notation, which captures the required uniformity.

Definition 2: Let $\lim_{\rho \rightarrow 1}^* f(k, \rho) = c$ mean that $\lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} |f(k, \rho) - c| = 0$, or equivalently, that $\lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} f(k, \rho) = \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} f(k, \rho) = c$.

Lemma D1 of Appendix D shows that certain properties of standard limits hold for \lim^* , as well.

Next, we provide four additional lemmas. Lemma 7 links $L_k(\rho)$ and $R_k(\rho)$ to $P_{k-1|k}$ and $P_{k+1|k}^*$, respectively, which will illuminate the importance of $L_k(\rho)$ and $R_k(\rho)$ and, hence, the reason for decomposing $\hat{f}_{X_2|I_1}(x|k)$ according to (10). Lemma 8 shows that $P_{k+1|k}^*$ is a sufficiently good approximation of $P_{k+1|k}$; Lemma 9 shows that $P_{k-1|k}$ and $P_{k+1|k}$ converge to zero in the \lim^* sense; and Lemma 10 shows how rapidly $P_{k+1|k}$ and $P_{k-1|k}$ decay in terms of $P_{k+1|k}^*$.

Lemma 7:

$$\text{A. } \lim_{\rho \rightarrow 1} \frac{P_{k-1|K}^*}{\frac{\pi}{\lambda} L_k(\rho)} = 1$$

$$\text{B. } \lim_{\rho \rightarrow 1} \frac{P_{k+1|K}^*}{\frac{\pi}{\lambda} R_k(\rho)} = 1$$

and consequently from B

C. for all ρ sufficiently close to one

$$P_{k+1|k}^* < \frac{1}{2} R_k(\rho) < \frac{\sqrt{2\pi}}{4\lambda} e^{\frac{\lambda^2}{2}} \sqrt{1 - \rho^2}$$

for $0 \leq k \leq K(\rho)$

D. for all ρ sufficiently close to one $P_{k+1|k}^* > \frac{1}{5} R_k(\rho)$, for $0 \leq k \leq K(\rho)$

where the second inequality in C follows from Lemma 6 (part C).

Lemma 8:

$$\lim_{\rho \rightarrow 1} \frac{P_{k+1|k}^*}{P_{k+1|k}^*} = 1.$$

Lemma 9:

$$\text{A. } \lim_{\rho \rightarrow 1} p_{k-1|K}^* = 0;$$

$$\text{B. } \lim_{\rho \rightarrow 1} p_{k+1|K}^* = 0.$$

Lemma 10: For all ρ sufficiently close to one, the following holds for $0 \leq k \leq K(\rho)$:

$$\text{A. } P_{k+m|k} < (P_{k+1|k}^*)^m, \text{ for } m \geq 2;$$

$$\text{B. } P_{k-m|k} < (P_{k+1|k}^*)^m, \text{ for } m \geq 2.$$

Now that we have established basic properties of $P_{k-1|k}$, $P_{k+1|k}$ and $P_{k+1|k}^*$, and the rate of decay of certain conditional probabilities, we are ready to proceed with the main part of the

proof and show that each of the five terms in (8) converges to one as $\rho \rightarrow 1$. We consider the terms in the following order: second, first, third, fourth, and fifth. The following lemma shows the second term converges to one as $\rho \rightarrow 1$.

Lemma 11:

$$\lim_{\rho \rightarrow 1} \frac{\sum_{k=1}^{K(\rho)} H(I_2|I_1 = k) P_k + \frac{1}{2} H(I_2|I_1 = 0) P_0}{\sum_{k=1}^{K(\rho)} (H_{k-1,k,k+1|k}) P_k + \frac{1}{2} (H_{-1,0,1|0}) P_0} = 1.$$

The next lemma is needed in the proof that the first term in (8) converges to one.

Lemma 12: For all ρ and k

$$H(I_2|I_1 = k) < 2b(\lambda\sqrt{1 - \rho^2}) + 5$$

where $b(\cdot)$ is defined in Lemma D2.

Finally, the four lemmas below show, respectively, that the first, third, fourth, and fifth terms converge to one as $\rho \rightarrow 1$, and conclude the proof of the theorem.

Lemma 13:

$$\lim_{\rho \rightarrow 1} \frac{\tilde{H}^\theta(I_2|I_1)}{\sum_{k=1}^{K(\rho)} H(I_2|I_1 = k) P_k + \frac{1}{2} H(I_2|I_1 = 0) P_0} = 1.$$

Lemma 14:

$$\lim_{\rho \rightarrow 1} \frac{\sum_{k=1}^{K(\rho)} (H_{k-1,k,k+1|k}) P_k + \frac{1}{2} (H_{-1,0,1|0}) P_0}{\sum_{k=1}^{K(\rho)} (H_{k-1,k+1|K}) P_k + \frac{1}{2} (H_{-1,1|0}) P_0} = 1.$$

Lemma 15:

$$\lim_{\rho \rightarrow 1} \frac{\sum_{k=1}^{K(\rho)} (H_{k-1,k+1|k}) P_k + \frac{1}{2} (H_{-1,1|0}) P_0}{\mathcal{H}(\sqrt{1 - \rho^2}) \left[\sum_{k=1}^{K(\rho)} w_{L,R,\lambda}^\theta(k) + \frac{1}{2} W_{L,R,\lambda}^\theta(0) \right]} = 1.$$

Lemma 16:

$$\lim_{\rho \rightarrow 1} \frac{\mathcal{H}(\sqrt{1 - \rho^2}) \left[\sum_{k=1}^{K(\rho)} w_{L,R,\lambda}^\theta(k) + \frac{1}{2} W_{L,R,\lambda}^\theta(0) \right]}{W_\lambda^\theta \mathcal{H}(\sqrt{1 - \rho})} = 1. \quad \square$$

Proof of Theorem 4: The proof is very similar to that of Theorem 3; thus, we shall only point out the differences, rather than repeat the derivations. The main difference lies in the fact that the ratio $\lambda = \frac{\Delta}{\sigma}$ is no longer constant, but rather tends to zero (we attribute this to quantizer cells shrinking, but it can equally be attributed to source variance increasing). Therefore, letting $\lambda(\rho) \rightarrow 0$ as $\rho \rightarrow 1$, as given in the theorem statement, one can restate and reprove Lemmas 5–16, which were needed in the proof of Theorem 3. When doing so, however, some care is needed. Specifically, since the cells shrink as ρ tends to one, it would be necessary to use a different function $K(\rho)$. It is easy to see that in order for the first fraction on the right-hand side of (8) to converge to one, the new $K(\rho)$, denoted $K_{\text{new}}(\rho)$, would have to tend to infinity faster than the old $K(\rho)$, which for emphasis

is denoted $K_{old}(\rho)$. While it is not at all obvious that there exists a choice of $K_{new}(\rho)$ that is sufficient, namely, for which all the new lemmas hold, it turns out that such a choice exists when $\lambda(\rho)$ satisfies the conditions of the theorem. In the following, we list the requirements for $K_{new}(\rho)$ and $\lambda(\rho)$ in order that the lemmas can be applied. As before, we let $x(\rho) \rightarrow a$ mean that the convergence happens as $\rho \rightarrow 1$.

- 1: $\frac{\lambda(\rho)}{\sqrt{1-\rho}} \rightarrow \infty$ required by Lemmas 5, 7–11, 13–15.
- 2: $K_{new}(\rho)\lambda(\rho)\sqrt{1-\rho} \rightarrow 0$, required by Lemmas 7–15.
- 3: $\frac{\lambda^2(\rho)}{\mathcal{H}(1-\rho)} \rightarrow \infty$, required by Lemma 10.
- 4: $\frac{e^{\frac{[K_{new}(\rho)\lambda(\rho)]^2}{2}}\sqrt{1-\rho}}{b(\lambda(\rho)\sqrt{1-\rho^2})} \rightarrow \infty$, Where $b(\cdot)$ is given in Lemma D2, required by Lemma 13.
- 5: $\frac{K_{new}(\rho)\lambda^2(\rho)}{\log \frac{1}{1-\rho}} \rightarrow 0$, required by Lemma 15.
- 6: $\liminf_{\rho \rightarrow 1} \frac{K_{new}(\rho)\lambda(\rho)}{\log \frac{1}{\lambda(\rho)}} > 2$, required by Lemma 16.

We note that the proof of Lemma 16 needs to be modified slightly to work for this theorem as well. Specifically, since λ depends on ρ , the last convergence to W_λ in the proof of Lemma 16 would be shown by showing that the ratio of the term converging to W_λ converges to one. Finally, by choosing $K_{new}(\rho) = \left\lfloor \frac{K_{old}(\rho)}{\lambda(\rho)} \right\rfloor$, we find that the following conditions on $\lambda(\rho)$ are necessary and sufficient for all of the above requirements to hold: (a) $\lim_{\rho \rightarrow 1} \frac{\lambda^2(\rho)}{\mathcal{H}(1-\rho)} = \infty$, and (b) $\liminf_{\rho \rightarrow 1} \frac{K_{old}(\rho)}{\log \frac{1}{\lambda(\rho)}} > 2$. This follows since (a) implies conditions 1 and 3 hold, the definition of K_{new} along with (b) imply conditions 5 and 6 hold, and the definitions of $K_{new}(\rho)$ and $K_{old}(\rho)$ imply conditions 2 and 4 hold. To see that condition 4 holds, we note that in light of condition 1 and the fact (given by Lemma D2) that $\limsup_{\alpha \rightarrow 0} \frac{b(\alpha)}{\log \frac{1}{\alpha}} \leq 1$, it follows

that it suffices to have $\frac{e^{\frac{K_{old}^2(\rho)}{2}}\sqrt{1-\rho}}{\log \frac{1}{1-\rho}} \rightarrow \infty$ as $\rho \rightarrow 1$, which is easily seen to hold. \square

V. LEMMA PROOFS

Proof of Lemma 5: We notice from (9) that there are three regions to consider: $x < \frac{t_k}{\rho}$, $\frac{t_k}{\rho} \leq x \leq \frac{t_{k+1}}{\rho}$ and $x > \frac{t_{k+1}}{\rho}$. Observe that for any k , $\tilde{f}_{X_2|I_1}(x|k) = f_{X_2|I_1}(x|k)$, for $\frac{t_k}{\rho} \leq x \leq \frac{t_{k+1}}{\rho}$. Thus, the lemma holds trivially in this region. Next, we will show that the lemma holds for $x < \frac{t_k}{\rho}$. The result for $x > \frac{t_{k+1}}{\rho}$ can be shown in a similar way

$$\begin{aligned} \frac{f_{X_2|I_1}(x|k)}{\tilde{f}_{X_2|I_1}(x|k)} &= \frac{\frac{1}{P_k} f_{X_2}(x) \left[Q\left(\frac{t_k - \rho x}{\sigma_\rho}\right) - Q\left(\frac{t_{k+1} - \rho x}{\sigma_\rho}\right) \right]}{\frac{1}{P_k} f_{X_2}(x) Q\left(\frac{t_k - \rho x}{\sigma_\rho}\right)} \\ &= 1 - \frac{Q\left(\frac{t_{k+1} - \rho x}{\sigma_\rho}\right)}{Q\left(\frac{t_k - \rho x}{\sigma_\rho}\right)}. \end{aligned}$$

We now show that the fraction on the right-hand side above approaches zero. Specifically

$$\begin{aligned} \frac{Q\left(\frac{t_{k+1} - \rho x}{\sigma_\rho}\right)}{Q\left(\frac{t_k - \rho x}{\sigma_\rho}\right)} &= \frac{\frac{1}{2} e^{-\frac{(t_{k+1} - \rho x)^2}{2\sigma_\rho^2}} g\left(\frac{t_{k+1} - \rho x}{\sigma_\rho}\right)}{\frac{1}{2} e^{-\frac{(t_k - \rho x)^2}{2\sigma_\rho^2}} g\left(\frac{t_k - \rho x}{\sigma_\rho}\right)} \\ &\stackrel{(a)}{<} \frac{e^{-\frac{(t_{k+1} - \rho x)^2}{2\sigma_\rho^2}}}{e^{-\frac{(t_k - \rho x)^2}{2\sigma_\rho^2}}} \\ &= e^{\frac{t_k^2 - t_{k+1}^2 + 2(t_{k+1} - t_k)\rho x}{2\sigma_\rho^2}} \\ &\stackrel{(b)}{<} e^{\frac{t_k^2 - t_{k+1}^2 + 2(t_{k+1} - t_k)t_k}{2\sigma_\rho^2}} \\ &= e^{-\frac{(t_{k+1} - t_k)^2}{2\sigma_\rho^2}} \\ &= e^{-\frac{\lambda^2}{2(1-\rho^2)}} \rightarrow 0 \text{ as } \rho \rightarrow 1 \end{aligned}$$

where (a) follows from $\frac{t_{k+1} - \rho x}{\sigma_\rho} > \frac{t_k - \rho x}{\sigma_\rho} > 0$ and the fact that g is monotonically decreasing (Fact C6); (b) is due to the facts that $t_{k+1} - t_k > 0$ and $x < \frac{t_k}{\rho}$; and the convergence holds for all $k, x < \frac{t_k}{\rho}$. \square

Proof of Lemma 6: We show the statements of the lemma in the following order: A, B, D, and C. Consider Part A

$$\begin{aligned} \frac{L_k(\rho)}{\sqrt{1-\rho^2}} &= \frac{1}{2} \frac{1}{Q((k-\theta)\lambda) - Q((k+1-\theta)\lambda)} e^{-\frac{t_k^2}{2\sigma^2}} \\ &\stackrel{(a)}{<} \sqrt{2\pi}(k-\theta)\lambda e^{\frac{(k-\theta)^2\lambda^2}{2}} e^{-\frac{(k-\theta)^2\lambda^2}{2}} \\ &\leq \sqrt{2\pi}k\lambda \end{aligned}$$

where (a) follows from Fact C4 and having $k - \theta > \max\{\frac{2}{\lambda}, \frac{2}{\lambda^2}\}$.

Next, Parts B and D are shown as follows:

$$\begin{aligned} \frac{L_k(\rho)}{\sqrt{1-\rho^2}} &= \frac{1}{2} \frac{1}{Q\left(\frac{t_k}{\sigma}\right) - Q\left(\frac{t_{k+1}}{\sigma}\right)} e^{-\frac{t_k^2}{2\sigma^2}} \\ &\stackrel{(a)}{>} \frac{1}{2} \frac{1}{\frac{1}{2} e^{-\frac{t_k^2}{2\sigma^2}}} e^{-\frac{t_k^2}{2\sigma^2}} \\ &= 1 \\ \frac{R_k(\rho)}{\sqrt{1-\rho^2}} &= \frac{1}{2} \frac{1}{Q\left(\frac{t_k}{\sigma}\right) - Q\left(\frac{t_{k+1}}{\sigma}\right)} e^{-\frac{t_{k+1}^2}{2\sigma^2}} \\ &\stackrel{(b)}{>} \frac{1}{2} \frac{1}{\frac{1}{2} e^{-\frac{t_{k+1}^2}{2\sigma^2}}} e^{-\frac{t_{k+1}^2}{2\sigma^2}} \\ &\geq e^{-k\lambda^2} e^{-\frac{\lambda^2}{2}} \end{aligned}$$

where (a) and (b) above follow from dropping the smaller Q terms in the denominators and using Fact C1 to upper bound

the larger Q terms (recalling that $k \geq 1$, so that Fact C1 can be used).

Finally, to show Part C, consider first $k \geq 1$

$$\begin{aligned} \frac{R_k(\rho)}{\sqrt{1-\rho^2}} &= \frac{1}{2} \frac{1}{\int_{t_k}^{t_{k+1}} \mathcal{N}_{0,\sigma^2}(x) dx} e^{-\frac{t_{k+1}^2}{2\sigma^2}} \\ &\stackrel{(a)}{<} \frac{1}{\frac{2}{\sqrt{2\pi\sigma}} e^{-\frac{t_{k+1}^2}{2\sigma^2}} \Delta} = \sqrt{\frac{\pi}{2}} \frac{1}{\lambda} \\ &< \sqrt{\frac{\pi}{2}} \frac{1}{\lambda} e^{\frac{\lambda^2}{2}} \end{aligned}$$

where (a) follows by lower bounding the Gaussian pdf with its value at the upper limit of the integral. Notice that the above also holds if $k = 0$ and $\theta \leq \frac{1}{2}$. If, however, $k = 0$ and $\theta > \frac{1}{2}$, then lower bounding the Gaussian pdf with its value at the lower limit of the integral yields (directly) the same upper bound. \square

Proof of Lemma 7: Part B: From the definition (11) of $P_{k+1|k}^*$ and from (9) and (10), we have that for all ρ and k

$$P_{k+1|k}^* < R_k(\rho) \int_{\frac{t_{k+1}}{\rho}}^{t_{k+2}} \mathcal{N}_{t_{k+1}\rho, \sigma_\rho^2}(x) g\left(\frac{\rho x - t_{k+1}}{\sigma_\rho}\right) dx. \quad (12)$$

Similarly, from (11), Lemma 5 and (10), we have that for any $0 \leq \gamma < 1$, there exists $\rho_\gamma < 1$ such that for all $\rho > \rho_\gamma$ and all k

$$P_{k+1|k}^* > R_k(\rho) \int_{\frac{t_{k+1}}{\rho}}^{t_{k+2}} \mathcal{N}_{t_{k+1}\rho, \sigma_\rho^2}(x) g\left(\frac{\rho x - t_{k+1}}{\sigma_\rho}\right) dx. \quad (13)$$

Assuming $k \geq 0$, as needed in this lemma, we evaluate the integrals above as follows:

$$\begin{aligned} &\int_{\frac{t_{k+1}}{\rho}}^{t_{k+2}} \mathcal{N}_{t_{k+1}\rho, \sigma_\rho^2}(x) g\left(\frac{\rho x - t_{k+1}}{\sigma_\rho}\right) dx \\ &= \int_{\frac{t_{k+1} - t_{k+1}\rho}{\sigma_\rho}}^{\frac{t_{k+2} - t_{k+1}\rho}{\sigma_\rho}} \mathcal{N}_{0,1}(x) g\left(\frac{\rho\sigma_\rho x + t_{k+1}\rho^2 - t_{k+1}}{\sigma_\rho}\right) dx \\ &= \int_{\frac{t_{k+1}(1-\rho) + \Delta}{\rho\sigma_\rho}}^{\frac{t_{k+1}(1-\rho^2)}{\rho\sigma_\rho}} \mathcal{N}_{0,1}(x) g\left(\rho x - \frac{t_{k+1}(1-\rho^2)}{\sigma_\rho}\right) dx \\ &= \int_0^\infty \mathcal{N}_{0,1}(x) g(\rho(x - S_{k,\rho})) I_{(S_{k,\rho}, T_{k,\rho})}(x) dx \quad (14) \end{aligned}$$

where $S_{k,\rho} \triangleq \frac{t_{k+1}(1-\rho^2)}{\rho\sigma_\rho}$, $T_{k,\rho} \triangleq \frac{t_{k+1}(1-\rho) + \Delta}{\sigma_\rho}$, and $I_F(x)$ denotes the indicator function of the set F . Next, we wish to apply Lemma D1 (Part A) to the right-hand side above. To justify this, we first observe, using Fact C6, that for all $x \in [0, \infty)$, the integrand is dominated by an integrable function because

$$\sup_{0 \leq k \leq K(\rho)} |\mathcal{N}_{0,1}(x) g(\rho(x - S_{k,\rho})) I_{(S_{k,\rho}, T_{k,\rho})}(x)| \leq \mathcal{N}_{0,1}(x).$$

Second, we need to show that $\lim_{\rho \rightarrow 1}^* \mathcal{N}_{0,1}(x) g(\rho(x - S_{k,\rho})) I_{(S_{k,\rho}, T_{k,\rho})}(x)$ exists for almost all $x \in [0, \infty)$. Defining $\tilde{g}(x) = g(x)$ for $x \geq 0$ and $\tilde{g}(x) = 1$ for $x < 0$, we have for $x \in (0, \infty)$

$$\begin{aligned} &\lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} \mathcal{N}_{0,1}(x) g(\rho(x - S_{k,\rho})) I_{(S_{k,\rho}, T_{k,\rho})}(x) \\ &\leq \mathcal{N}_{0,1}(x) \lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} \tilde{g}(\rho(x - S_{k,\rho})) \\ &\stackrel{(a)}{=} \mathcal{N}_{0,1}(x) \lim_{\rho \rightarrow 1} \tilde{g}(\rho(x - S_{K(\rho), \rho})) \\ &\stackrel{(b)}{=} \mathcal{N}_{0,1}(x) g(x) \quad (15) \end{aligned}$$

where (a) follows from the monotonicity of g (Fact C6), and (b) follows from having $x > 0$ and $\lim_{\rho \rightarrow 1} S_{K(\rho), \rho} = 0$. We also have for $x > 0$

$$\begin{aligned} &\lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} \mathcal{N}_{0,1}(x) g(\rho(x - S_{k,\rho})) I_{(S_{k,\rho}, T_{k,\rho})}(x) \\ &\geq \mathcal{N}_{0,1}(x) \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} \tilde{g}(\rho(x - S_{k,\rho})) \\ &\quad \times \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} I_{(S_{k,\rho}, T_{k,\rho})}(x) \\ &\stackrel{(a)}{=} \mathcal{N}_{0,1}(x) \lim_{\rho \rightarrow 1} \tilde{g}(\rho(x - S_{0,\rho})) \lim_{\rho \rightarrow 1} I_{(S_{K(\rho), \rho}, T_{0,\rho})}(x) \\ &\stackrel{(b)}{=} \mathcal{N}_{0,1}(x) g(x) \quad (16) \end{aligned}$$

where (a) follows from the monotonicity of g (Fact C6), and (b) is due to having $x > 0$, $\lim_{\rho \rightarrow 1} S_{0,\rho} = 0$, $\lim_{\rho \rightarrow 1} S_{K(\rho), \rho} = 0$, and $\lim_{\rho \rightarrow 1} T_{0,\rho} = \infty$. Equations (15) and (16) now imply that $\lim_{\rho \rightarrow 1}^* \mathcal{N}_{0,1}(x) g(\rho(x - S_{k,\rho})) I_{(S_{k,\rho}, T_{k,\rho})}(x) = \mathcal{N}_{0,1}(x) g(x)$ for all $x \in (0, \infty)$. Consequently, we may apply Lemma D1 (Part A) to the right-hand side of (14) and obtain

$$\begin{aligned} &\lim_{\rho \rightarrow 1}^* \int_0^\infty \mathcal{N}_{0,1}(x) g(\rho(x - S_{k,\rho})) I_{(S_{k,\rho}, T_{k,\rho})}(x) dx \\ &= \int_0^\infty \lim_{\rho \rightarrow 1}^* \mathcal{N}_{0,1}(x) g(\rho(x - S_{k,\rho})) I_{(S_{k,\rho}, T_{k,\rho})}(x) dx \\ &= \int_0^\infty \mathcal{N}_{0,1}(x) g(x) dx = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{Q(x)}{\frac{1}{2} e^{-\frac{x^2}{2}}} dx \\ &= \sqrt{\frac{2}{\pi}} \int_0^\infty Q(x) dx \\ &= \frac{1}{\pi}. \quad (17) \end{aligned}$$

Finally, combining the fact that γ in (13) can be chosen arbitrarily close to one, together with (12), (14), and (17), it follows that $\lim_{\rho \rightarrow 1}^* \frac{P_{k+1|k}^*}{R_k(\rho)} = 1$, which concludes the proof of Part B.

Part A: The proof is similar to that of Part B, up to a point. Assume first that $k \geq 1$. Then, $t_k \leq \frac{t_k}{\rho}$, and using Lemma 5 and the definition of $\tilde{f}_{X_2|I_1}(x|k)$, we have that for all ρ

$$P_{k-1|k} < L_k(\rho) \int_{t_{k-1}}^{t_k} \mathcal{N}_{t_k\rho, \sigma_\rho^2}(x) g\left(\frac{t_k - \rho x}{\sigma_\rho}\right) dx \quad (18)$$

and that for any $0 \leq \gamma < 1$, there exists $\rho_\gamma < 1$ such that for all $\rho > \rho_\gamma$

$$P_{k-1|k} > \gamma L_k(\rho) \int_{t_{k-1}}^{t_k} \mathcal{N}_{t_k, \rho, \sigma_\rho^2}(x) g\left(\frac{t_k - \rho x}{\sigma_\rho}\right) dx. \quad (19)$$

The integrals above are evaluated by taking similar steps to (14) and obtaining

$$\begin{aligned} & \int_{t_{k-1}}^{t_k} \mathcal{N}_{t_k, \rho, \sigma_\rho^2}(x) g\left(\frac{t_k - \rho x}{\sigma_\rho}\right) dx \\ &= \int_{-\infty}^0 \mathcal{N}_{0,1}(x) g(T_{k,\rho}(1+\rho) - \rho x) \\ & \quad \times I_{(S_{k,\rho}, T_{k,\rho})}(x) dx \end{aligned} \quad (20)$$

where here $S_{k,\rho} \triangleq \frac{t_k(1-\rho) - \Delta}{\sigma_\rho}$ and $T_{k,\rho} \triangleq \frac{t_k(1-\rho)}{\sigma_\rho}$. We now introduce the operator $\widehat{\lim}^*$, which is the same as \lim^* except that k is greater than or equal to one rather than zero. Namely, if $\widehat{\lim}^* f_k(\rho) = c$ for some function $f_k(\rho)$, then it means that $\lim_{\rho \rightarrow 1} \sup_{1 \leq k \leq K(\rho)} |f_k(\rho) - c| = 0$.

Next, we would like to apply Lemma D1 (Part A) to the right-hand side of (20). Clearly, the result of the lemma holds for the $\widehat{\lim}^*$ operator, as well. To justify using this lemma, we first observe that for all $x \in (-\infty, 0]$, the integrand is dominated by an integrable function because

$$\begin{aligned} & \sup_{1 \leq k \leq K(\rho)} |\mathcal{N}_{0,1}(x) g(T_{k,\rho}(1+\rho) - \rho x) I_{(S_{k,\rho}, T_{k,\rho})}(x)| \\ & \leq \mathcal{N}_{0,1}(x). \end{aligned}$$

Second, we need to show that $\widehat{\lim}_{\rho \rightarrow 1}^* \mathcal{N}_{0,1}(x) g(T_{k,\rho}(1+\rho) - \rho x) I_{(S_{k,\rho}, T_{k,\rho})}(x)$ exists for almost all $x \in (-\infty, 0]$. For $x \in (-\infty, 0]$

$$\begin{aligned} & \lim_{\rho \rightarrow 1} \sup_{1 \leq k \leq K(\rho)} \mathcal{N}_{0,1}(x) g(T_{k,\rho}(1+\rho) - \rho x) I_{(S_{k,\rho}, T_{k,\rho})}(x) \\ & \leq \mathcal{N}_{0,1}(x) \lim_{\rho \rightarrow 1} \sup_{1 \leq k \leq K(\rho)} g(T_{k,\rho}(1+\rho) - \rho x) \\ & \stackrel{(a)}{=} \mathcal{N}_{0,1}(x) \lim_{\rho \rightarrow 1} g(T_{1,\rho}(1+\rho) - \rho x) \\ & \stackrel{(b)}{=} \mathcal{N}_{0,1}(x) g(-x) \end{aligned} \quad (21)$$

where (a) follows from the monotonicity of g (Fact C6), and (b) follows from having $x \leq 0$ and $\lim_{\rho \rightarrow 1} T_{1,\rho} = 0$. We also have for $x \leq 0$

$$\begin{aligned} & \lim_{\rho \rightarrow 1} \inf_{1 \leq k \leq K(\rho)} \mathcal{N}_{0,1}(x) g(T_{k,\rho}(1+\rho) - \rho x) I_{(S_{k,\rho}, T_{k,\rho})}(x) \\ & \geq \mathcal{N}_{0,1}(x) \lim_{\rho \rightarrow 1} \inf_{1 \leq k \leq K(\rho)} g(T_{k,\rho}(1+\rho) - \rho x) \\ & \quad \times \lim_{\rho \rightarrow 1} \inf_{1 \leq k \leq K(\rho)} I_{(S_{k,\rho}, T_{k,\rho})}(x) \\ & \stackrel{(a)}{=} \mathcal{N}_{0,1}(x) \lim_{\rho \rightarrow 1} g(T_{K(\rho), \rho}(1+\rho) - \rho x) \\ & \quad \times \lim_{\rho \rightarrow 1} I_{(S_{K(\rho), \rho}, T_{1,\rho})}(x) \\ & \stackrel{(b)}{=} \mathcal{N}_{0,1}(x) g(-x) \end{aligned} \quad (22)$$

where (a) follows from the monotonicity of g (Fact C6), and (b) is due to having $x \leq 0$, $\lim_{\rho \rightarrow 1} T_{K(\rho), \rho} = 0$, $\lim_{\rho \rightarrow 1} S_{K(\rho), \rho} = -\infty$, and $T_{1,\rho} \geq 0$ for all ρ . Equation (21) and (22) now imply that $\widehat{\lim}_{\rho \rightarrow 1}^* \mathcal{N}_{0,1}(x) g(T_{k,\rho}(1+\rho) - \rho x) I_{(S_{k,\rho}, T_{k,\rho})}(x) = \mathcal{N}_{0,1}(x) g(-x)$ for all $x \in (-\infty, 0]$. Consequently, we may apply Lemma D1 (Part A), which, as mentioned, holds for the \lim^* operator as well, to the right-hand side of (20) and obtain

$$\begin{aligned} & \widehat{\lim}_{\rho \rightarrow 1}^* \int_{-\infty}^0 \mathcal{N}_{0,1}(x) g(T_{k,\rho}(1+\rho) - \rho x) I_{(S_{k,\rho}, T_{k,\rho})}(x) dx \\ &= \int_{-\infty}^0 \widehat{\lim}_{\rho \rightarrow 1}^* \mathcal{N}_{0,1}(x) g(T_{k,\rho}(1+\rho) - \rho x) \\ & \quad \times I_{(S_{k,\rho}, T_{k,\rho})}(x) dx \\ &= \int_{-\infty}^0 \mathcal{N}_{0,1}(x) g(-x) dx \\ &= \int_{-\infty}^0 \mathcal{N}_{0,1}(x) g(x) dx \\ &= \frac{1}{\pi} \end{aligned} \quad (23)$$

where the last equality follows from (17).

Next, combining the fact that γ in (19) can be chosen arbitrarily close to one, together with (18), (20), and (23), it follows that

$$\widehat{\lim}_{\rho \rightarrow 1}^* \frac{P_{k-1|k}}{\frac{1}{\pi} L_k(\rho)} = 1. \quad (24)$$

It remains to consider the $k = 0$ case. Recall that $P_{-1|0}^\theta = P_{1|0}^{1-\theta}$. From Lemma 8, we have $\lim_{\rho \rightarrow 1} \frac{P_{1|0}^{1-\theta}}{P_{1|0}^{*1-\theta}} = 1$ (note that the proof of Lemma 8 uses Part B of the present lemma, which has already been established, so there is no circularity). By Part B, $\lim_{\rho \rightarrow 1} \frac{P_{1|0}^{*1-\theta}}{\frac{1}{\pi} R_0^{1-\theta}(\rho)} = 1$. Combining this with the facts $R_0^{1-\theta}(\rho) = L_0^\theta(\rho)$, $\lim_{\rho \rightarrow 1} \frac{P_{1|0}^{1-\theta}}{P_{1|0}^{*1-\theta}} = 1$, and $P_{-1|0}^\theta = P_{1|0}^{1-\theta}$, we find $\lim_{\rho \rightarrow 1} \frac{P_{-1|0}^\theta}{\frac{1}{\pi} L_0^\theta(\rho)} = 1$. This together with (24) completes the proof of Part A and of the lemma as a whole. \square

Proof of Lemma 8: Let $k \geq 0$, which is sufficient for the purpose of the lemma. We write the following:

$$\begin{aligned} P_{k+1|k} &= \int_{t_{k+1}}^{\frac{t_{k+1}}{\rho}} f_{X_2|I_1}(x|k) dx + P_{k+1|k}^* \\ &< \int_{t_{k+1}}^{\frac{t_{k+1}}{\rho}} \frac{\mathcal{N}_{0,\sigma^2}(x)}{P_k} dx + P_{k+1|k}^* \end{aligned}$$

where the inequality follows from the fact that $f_{X_2|I_1}(x|k) \leq \tilde{f}_{X_2|I_1}(x|k) < \frac{\mathcal{N}_{0,\sigma^2}(x)}{P_k}$ for $t_{k+1} \leq x \leq \frac{t_{k+1}}{\rho}$, as seen by (10). Combining this with the fact, shown in Lemma 7 (Part D),

that for all ρ sufficiently close to one, $P_{k+1|k}^* > \frac{1}{5}R_k(\rho)$ for $0 \leq k \leq K(\rho)$, it follows that it suffices to show

$$\lim_{\rho \rightarrow 1} \int_{t_{k+1}}^{\frac{t_{k+1}}{\rho}} \frac{N_{0,\sigma^2}(x)}{P_k} dx = 0. \quad (25)$$

Simplifying the above expression, we obtain

$$\begin{aligned} & \sup_{0 \leq k \leq K(\rho)} \int_{t_{k+1}}^{\frac{t_{k+1}}{\rho}} \frac{N_{0,\sigma^2}(x)}{P_k} dx \\ &= \sup_{0 \leq k \leq K(\rho)} \frac{\frac{1}{P_k} \int_{t_{k+1}}^{\frac{t_{k+1}}{\rho}} N_{0,\sigma^2}(x) dx}{\frac{1}{5} \frac{1}{2} \frac{1}{P_k} e^{-\frac{t_{k+1}^2}{2\sigma^2}} \sqrt{1-\rho^2}} \\ &\stackrel{(a)}{\leq} \sup_{0 \leq k \leq K(\rho)} \frac{10 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t_{k+1}^2}{2\sigma^2}} \left[t_{k+1} \frac{(1-\rho)}{\rho} \right]}{e^{-\frac{t_{k+1}^2}{2\sigma^2}} \sqrt{1-\rho^2}} \\ &\stackrel{(b)}{\leq} \frac{10\lambda}{\sqrt{2\pi}\rho} (K(\rho) + 1 - \theta) \sqrt{1-\rho} \\ &< \frac{10\lambda}{\sqrt{2\pi}\rho} \left(\ln \frac{1}{1-\rho} \right)^{\frac{3}{4}} \sqrt{1-\rho} \\ &\rightarrow 0 \text{ as } \rho \rightarrow 1 \end{aligned}$$

where (a) follows from substituting the Gaussian pdf with its value at the lower limit of the integral, and (b) is obtained by recalling that $t_{k+1} = (k + 1 - \theta)\Delta$, substituting $K(\rho)$ for k , and noting that $1 < \sqrt{1 + \rho}$. This shows (25) and concludes the proof of the lemma. \square

Proof of Lemma 9: Part B: Lemmas 7 (Part B), 8, and D1 (Part C) imply that $\lim_{\rho \rightarrow 1} \frac{P_{k+1|k}}{\frac{1}{\pi} R_k(\rho)} = 1$. Lemma 6 (Part C) shows that $R_k(\rho) < \sqrt{\frac{\pi}{2}} \frac{1}{\lambda} e^{\frac{\lambda^2}{2}} \sqrt{1-\rho^2}$ for all $k \geq 0$, so that $\lim_{\rho \rightarrow 1}^* R_k(\rho) = 0$. It now follows from Lemma D1 (Part C) that $\lim_{\rho \rightarrow 1}^* P_{k+1|k} = 0$, which is Part B of the lemma.

Part A: To show $\lim_{\rho \rightarrow 1}^* P_{k-1|k} = 0$, from Lemma 7 (Part A) and Lemma D1 (Part C), it suffices to show $\lim_{\rho \rightarrow 1}^* L_k(\rho) = 0$.

Clearly, for any k , $L_k(\rho) = \frac{1}{2} \frac{1}{P_k} e^{-\frac{t_k^2}{2\sigma^2}} \sqrt{1-\rho^2} \rightarrow 0$ as $\rho \rightarrow 1$. That $\lim_{\rho \rightarrow 1}^* L_k(\rho) = 0$, follows directly from the fact that for k such that $K_\lambda < k \leq K(\rho)$, where K_λ is the constant given in Lemma 6 (Part A)

$$L_k(\rho) < \sqrt{2\pi} k \lambda \sqrt{1-\rho^2} \leq \sqrt{2\pi} K(\rho) \lambda \sqrt{1-\rho^2} \rightarrow 0 \text{ as } \rho \rightarrow 1. \quad \square$$

Proof of Lemma 10: Part A: We need to show that for all ρ sufficiently close to one and for $0 \leq k \leq K(\rho)$

$$P_{k+m|k} < (P_{k+1|k}^*)^m \text{ for } m \geq 2.$$

This follows easily from the steps below that hold for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$.

A1: $P_{l+1|k} < P_{l|k} P_{k+1|k}^*$, $l \geq k + 2$.

A2: $P_{k+2|k} < (P_{k+1|k}^*)^2$.

Step A1: We observe that $t_l > \frac{t_{k+1}}{\rho}$ for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$. Thus, since $f_{X_2|I_1}(x|k) < \tilde{f}_{X_2|I_1}(x|k)$ for all $x > \frac{t_{k+1}}{\rho}$, it follows from (10) and the monotonicity of g (Fact C6) that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} P_{l+1|k} &< R_k(\rho) \int_{t_{l+1}}^{t_{l+2}} N_{t_{k+1}, \sigma_\rho^2}(x) g\left(\frac{\rho x - t_{k+1}}{\sigma_\rho}\right) dx \\ &< R_k(\rho) g\left(\frac{\rho t_{l+1} - t_{k+1}}{\sigma_\rho}\right) Q\left(\frac{t_{l+1} - t_{k+1}\rho}{\sigma_\rho}\right). \end{aligned} \quad (26)$$

Next, using Lemma 5 with $\gamma = \frac{1}{2}$, (10), and the monotonicity of g (Fact C6), we obtain that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} P_{l|k} &> \frac{1}{2} R_k(\rho) \int_{t_l}^{t_{l+1}} N_{t_{k+1}, \sigma_\rho^2}(x) g\left(\frac{\rho x - t_{k+1}}{\sigma_\rho}\right) dx \\ &> \frac{1}{2} R_k(\rho) g\left(\frac{\rho t_{l+1} - t_{k+1}}{\sigma_\rho}\right) \\ &\quad \times \left[Q\left(\frac{t_l - t_{k+1}\rho}{\sigma_\rho}\right) - Q\left(\frac{t_{l+1} - t_{k+1}\rho}{\sigma_\rho}\right) \right]. \end{aligned} \quad (27)$$

Using (26) and (27), it suffices to show that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} & Q\left(\frac{t_{l+1} - t_{k+1}\rho}{\sigma_\rho}\right) \\ & < \frac{1}{2} \left[Q\left(\frac{t_l - t_{k+1}\rho}{\sigma_\rho}\right) - Q\left(\frac{t_{l+1} - t_{k+1}\rho}{\sigma_\rho}\right) \right] P_{k+1|k}^* \end{aligned}$$

or alternatively, since $P_{k+1|k}^* \leq 1$, it suffices to have

$$\frac{3Q\left(\frac{t_{l+1} - t_{k+1}\rho}{\sigma_\rho}\right)}{Q\left(\frac{t_l - t_{k+1}\rho}{\sigma_\rho}\right)} < P_{k+1|k}^*. \quad (28)$$

Next, we observe that the argument of the Q function in the numerator can be written as

$$\frac{t_{l+1} - t_{k+1}\rho}{\sigma_\rho} = \frac{t_l - t_{k+1}\rho + \Delta}{\sigma_\rho} = \frac{\Delta \left(\frac{t_l - t_{k+1}\rho}{\Delta} + 1 \right)}{\sigma_\rho}.$$

Using Fact C5, with $a = \frac{t_l - t_{k+1}\rho}{\Delta}$ and $z = \frac{\Delta}{\sigma_\rho} = \frac{\lambda}{\sqrt{1-\rho^2}}$, where we notice that $l \geq k + 2$ implies that $a \geq 1$, it follows that for all ρ sufficiently close to one

$$\frac{3Q\left(\frac{t_{l+1} - t_{k+1}\rho}{\sigma_\rho}\right)}{Q\left(\frac{t_l - t_{k+1}\rho}{\sigma_\rho}\right)} < 6e^{-\frac{\lambda^2}{2(1-\rho^2)}}. \quad (29)$$

To obtain a lower bound to $P_{k+1|k}^*$, we apply Lemmas 7 (Part D) and 6 (Part D) to obtain that for all ρ sufficiently close to one and $1 \leq k \leq K(\rho)$

$$\begin{aligned} P_{k+1|k}^* &> \frac{1}{5} \sqrt{1-\rho^2} e^{-k\lambda^2} e^{-\frac{\lambda^2}{2}} \\ &\geq \frac{1}{5} \sqrt{1-\rho^2} e^{-K(\rho)\lambda^2} e^{-\frac{\lambda^2}{2}} \end{aligned} \quad (30)$$

where the second inequality derives from substituting $K(\rho)$ for k . If $k = 0$, then from Lemma 7 (Part D),

$P_{1|0}^* > \frac{1}{5} \frac{1}{2} \frac{1}{P_0} e^{-\frac{t_1^2}{2\sigma^2}} \sqrt{1-\rho^2}$, where we notice that P_0 and $e^{-\frac{t_1^2}{2\sigma^2}}$ do not depend on ρ . Combining this and (30) with (29), it is easy to see that (28) holds for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$, which completes Step A1.

Step A2: Applying (26) with $l = k+1$, we obtain that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$P_{k+2|k} < R_k(\rho)g\left(\frac{\rho t_{k+2} - t_{k+1}}{\sigma_\rho}\right)Q\left(\frac{t_{k+2} - t_{k+1}\rho}{\sigma_\rho}\right).$$

Next, we observe (using Fact C1) that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} Q\left(\frac{t_{k+2} - t_{k+1}\rho}{\sigma_\rho}\right) &< \frac{1}{2}e^{-\frac{(t_{k+2} - t_{k+1}\rho)^2}{2\sigma_\rho^2}} < \frac{1}{2}e^{-\frac{\lambda^2}{2(1-\rho^2)}} \\ &< \frac{1}{12}P_{k+1|k}^* \end{aligned}$$

where the last inequality was shown in Step A1. Substituting this bound into the right-hand side of the previous equation, we find that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$P_{k+2|k} < R_k(\rho)g\left(\frac{\rho t_{k+2} - t_{k+1}}{\sigma_\rho}\right)\frac{1}{12}P_{k+1|k}^*. \quad (31)$$

Using the same ideas that lead to (27) with $l = k+1$, we find that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} P_{k+1|k}^* &> \frac{1}{2}R_k(\rho)g\left(\frac{\rho t_{k+2} - t_{k+1}}{\sigma_\rho}\right) \\ &\times \left[Q\left(\frac{\frac{t_{k+1}}{\rho} - t_{k+1}\rho}{\sigma_\rho}\right) - Q\left(\frac{t_{k+2} - t_{k+1}\rho}{\sigma_\rho}\right)\right] \\ &> \frac{1}{2}R_k(\rho)g\left(\frac{\rho t_{k+2} - t_{k+1}}{\sigma_\rho}\right)\frac{1}{6}, \end{aligned} \quad (32)$$

where the second inequality follows from having $\lim_{\rho \rightarrow 1} \frac{\frac{t_{k+1}}{\rho} - t_{k+1}\rho}{\sigma_\rho} = 0$ and $\lim_{\rho \rightarrow 1} \frac{t_{k+2} - t_{k+1}\rho}{\sigma_\rho} = \infty$ so that the left Q function above goes to $\frac{1}{2}$ and the right Q function above goes to 0 as ρ goes to one, uniformly for $0 \leq k \leq K(\rho)$. Step A2 now follows from (31) and (32).

Part B: We need to show that for all ρ sufficiently close to one and for $0 \leq k \leq K(\rho)$

$$P_{k-m|k} < (P_{k+1|k}^*)^m \text{ for } m \geq 2.$$

The proof is similar to that of Part A, yet not identical. The difference lies in the fact that B has $P_{k+1|k}^*$ in its expression instead of, say, $P_{k-1|k}$. Part B follows easily from the steps below that hold for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$.

B1: $P_{l-1|k} < P_{l|k}P_{k+1|k}^*$ for $l \leq k-2$.

B2: $P_{k-2|k} < (P_{k+1|k}^*)^2$.

Step B1: We observe that $t_{l+1} < \frac{t_k}{\rho}$ for all ρ and $l \leq k-2$. Thus, since $f_{X_2|I_1}(x|k) < \tilde{f}_{X_2|I_1}(x|k)$ for all $x < \frac{t_k}{\rho}$, it follows

using derivations similar to those leading to (26) that for all ρ and $0 \leq k \leq K(\rho)$

$$P_{l-1|k} < L_k(\rho)g\left(\frac{t_k - \rho t_l}{\sigma_\rho}\right)Q\left(\frac{t_k\rho - t_l}{\sigma_\rho}\right). \quad (33)$$

A derivation similar to that of (27) shows that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} P_{l|k} &> \frac{L_k(\rho)}{2}g\left(\frac{t_k - \rho t_l}{\sigma_\rho}\right) \\ &\times \left[Q\left(\frac{t_k\rho - t_{l+1}}{\sigma_\rho}\right) - Q\left(\frac{t_k\rho - t_l}{\sigma_\rho}\right)\right]. \end{aligned}$$

Using the above two equations together with similar steps to those used to obtain (28), it follows that it suffices to show

$$\frac{3Q\left(\frac{t_k\rho - t_l}{\sigma_\rho}\right)}{Q\left(\frac{t_k\rho - t_{l+1}}{\sigma_\rho}\right)} < P_{k+1|k}^*. \quad (34)$$

Following the same derivation as that which was used to obtain (29), we have

$$\frac{3Q\left(\frac{t_k\rho - t_l}{\sigma_\rho}\right)}{Q\left(\frac{t_k\rho - t_{l+1}}{\sigma_\rho}\right)} < 6e^{-\frac{\lambda^2}{2(1-\rho^2)}}.$$

Combining this with (34) and with the lower bound to $P_{k+1|k}^*$ given in (30) and with the derivation thereafter, completes the proof of Step B1.

Step B2: Applying (33) with $l = k-1$, we obtain that for all ρ and $0 \leq k \leq K(\rho)$

$$\begin{aligned} P_{k-2|k} &< L_k(\rho)g\left(\frac{t_k - t_{k-1}\rho}{\sigma_\rho}\right)Q\left(\frac{t_k\rho - t_{k-1}}{\sigma_\rho}\right) \\ &= L_k(\rho)g\left(\frac{(k-\theta)\Delta(1-\rho) + \Delta\rho}{\sigma_\rho}\right) \\ &\times Q\left(\frac{t_k\rho - t_{k-1}}{\sigma_\rho}\right). \end{aligned} \quad (35)$$

From (32), we have that for all ρ sufficiently close to one and $1 \leq k \leq K(\rho)$ with any θ , and for $k=0$ and $\theta \leq \frac{1}{2}$

$$\begin{aligned} P_{k+1|k}^* &\stackrel{(a)}{>} \frac{R_k(\rho)}{2}g\left(\frac{\rho t_{k+2} - t_{k+1}}{\sigma_\rho}\right) \\ &\times \left[\frac{1}{4} - Q\left(\frac{t_{k+2} - t_{k+1}\rho}{\sigma_\rho}\right)\right] \\ &\stackrel{(b)}{>} \frac{1}{16}R_k(\rho)g\left(\frac{\rho t_{k+2} - t_{k+1}}{\sigma_\rho}\right) \\ &= \frac{1}{16}R_k(\rho)g\left(\frac{-(k+1-\theta)\Delta(1-\rho) + \Delta\rho}{\sigma_\rho}\right) \\ &\stackrel{(c)}{\geq} \frac{1}{16}R_k(\rho)g\left(\frac{(k-\theta)\Delta(1-\rho) + \Delta\rho}{\sigma_\rho}\right) \end{aligned} \quad (36)$$

where (a) follows since the first Q function in (32) tends to $\frac{1}{2}$ as $\rho \rightarrow 0$ (since $\lim_{\rho \rightarrow 1}^* \frac{t_{k+1} - t_{k+1}\rho}{\sigma_\rho} = 0$) and is in particular smaller than $\frac{1}{4}$ for sufficiently large ρ , (b) derives from the fact that the Q term tends to zero in the \lim^* sense, thus, in particular, it is less than $1/8$ for all ρ sufficiently close to one, and (c) is due to having $-(k+1-\theta) \leq k-\theta$, along with the monotonicity of g (Fact C6). When $k=0$ and $\theta > \frac{1}{2}$, (b) above does not hold, however, $P_{-2|0}$ in (35) can be further upper bounded as $P_{-2|0} < L_0(\rho)g\left(\frac{-(1-\theta)\Delta(1-\rho)+\Delta\rho}{\sigma_\rho}\right)Q\left(\frac{t_0\rho-t_{-1}}{\sigma_\rho}\right)$, by using the fact that $-(1-\theta) > -\theta$, along with the monotonicity of g (Fact C6). It now follows from this and (35) and (36) that it suffices to show that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$16 \frac{L_k(\rho)}{R_k(\rho)} Q\left(\frac{t_k\rho - t_{k-1}}{\sigma_\rho}\right) < P_{k+1|k}^* \tag{37}$$

Using the definitions of $L_k(\rho)$ and $R_k(\rho)$, we obtain that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} & 16 \frac{L_k(\rho)}{R_k(\rho)} Q\left(\frac{t_k\rho - t_{k-1}}{\sigma_\rho}\right) \\ &= \frac{16 \frac{1}{2} \frac{1}{P_k} e^{-\frac{t_k^2}{2\sigma^2}} \sqrt{1-\rho^2}}{\frac{1}{2} \frac{1}{P_k} e^{-\frac{t_{k+1}^2}{2\sigma^2}} \sqrt{1-\rho^2}} Q\left(\frac{[(k-\theta)\rho - (k-1-\theta)]\Delta}{\sigma\sqrt{1-\rho^2}}\right) \\ &= 16e^{\frac{\lambda^2}{2}} e^{k\lambda^2} Q\left(\frac{(1-(k-\theta)(1-\rho))\lambda}{\sqrt{1-\rho^2}}\right) \\ &\stackrel{(a)}{\leq} 16e^{\frac{\lambda^2}{2}} e^{k\lambda^2} Q\left(\frac{\lambda}{2\sqrt{1-\rho^2}}\right) \\ &\stackrel{(b)}{\leq} 16e^{\frac{\lambda^2}{2}} e^{k\lambda^2} \frac{1}{2} e^{-\frac{\lambda^2}{8(1-\rho^2)}} \\ &\stackrel{(c)}{\leq} 8e^{\frac{\lambda^2}{2}} e^{K(\rho)\lambda^2} e^{-\frac{\lambda^2}{8(1-\rho^2)}} \end{aligned} \tag{38}$$

where (a) follows from having $\lim_{\rho \rightarrow 1}^* (k-\theta)(1-\rho) = 0$, (b) is due to Fact C1, and (c) is obtained by substituting $K(\rho)$ for k .

Finally, for the case that $1 \leq k \leq K(\rho)$, (37) can be seen to hold for all ρ sufficiently close to one, by combining (38) and (30). For the case that $k=0$, combining (38) with Lemma 7 (Part B), from which it follows that $\lim_{\rho \rightarrow 1} \frac{P_{1|0}}{\frac{1}{2\pi} \frac{1}{P_0} e^{-\frac{(1-\theta)^2\lambda^2}{2}} \sqrt{1-\rho^2}} = 1$, implies (37) for all ρ sufficiently close to one. This concludes Step B2 and the proof of the lemma. \square

Proof of Lemma 11: It follows from Lemma D1 (Part D) that it suffices to show

$$\lim_{\rho \rightarrow 1}^* \frac{H(I_2|I_1 = k)}{H_{k-1,k,k+1|k}} = 1.$$

Since $H(I_2|I_1 = k) = H_{k-1,k,k+1|k} + \sum_{l=-\infty}^{-2} H_{k+l|k} + \sum_{l=2}^{\infty} H_{k+l|k}$, from Lemma D1 (Part B), it suffices to show

$$\lim_{\rho \rightarrow 1}^* \frac{\sum_{l=-\infty}^{-2} H_{k+l|k} + \sum_{l=2}^{\infty} H_{k+l|k}}{H_{k-1|k} + H_{k|k} + H_{k+1|k}} = 0. \tag{39}$$

We proceed by upper bounding the second sum in the numerator above. In a similar way, the same expression can be shown to be an upper bound for the first term. First, let us write the second sum more explicitly as $\sum_{l=2}^{\infty} H_{k+l|k} = -\sum_{l=2}^{\infty} P_{k+l|k} \log P_{k+l|k}$. Next, recall from Lemma 7 (Part C) that for all ρ sufficiently close to one, $P_{k+1|k}^* < \frac{\sqrt{2\pi}}{4\lambda} e^{\frac{\lambda^2}{2}} \sqrt{1-\rho^2}$, for $0 \leq k \leq K(\rho)$. Thus, for all ρ sufficiently close to one, $P_{k+1|k}^* < \frac{1}{e}$ for $0 \leq k \leq K(\rho)$. We will assume for the rest of the proof that $P_{k+1|k}^* < \frac{1}{e}$. Combining this with Fact C8 and with the fact that $P_{k+l|k} < (P_{k+1|k}^*)^l$ for $l \geq 2$ and $0 \leq k \leq K(\rho)$, as shown in Lemma 10 (Part A), it follows that $-P_{k+l|k} \log P_{k+l|k} < -(P_{k+1|k}^*)^l \log (P_{k+1|k}^*)^l$ for $l \geq 2$ and $0 \leq k \leq K(\rho)$. Therefore

$$\begin{aligned} & -\sum_{l=2}^{\infty} P_{k+l|k} \log P_{k+l|k} \\ &< -\sum_{m=0}^{\infty} (P_{k+1|k}^*)^{m+2} \log (P_{k+1|k}^*)^{m+2} \\ &= -(P_{k+1|k}^*)^2 (\log P_{k+1|k}^*) \sum_{m=0}^{\infty} (m+2)(P_{k+1|k}^*)^m \\ &= -(P_{k+1|k}^*)^2 (\log P_{k+1|k}^*) \\ &\quad \times \left[\frac{P_{k+1|k}^*}{(1-P_{k+1|k}^*)^2} + \frac{2}{1-P_{k+1|k}^*} \right] \\ &< -P_{k+1|k}^* (\log P_{k+1|k}^*) \frac{4P_{k+1|k}^*}{1-P_{k+1|k}^*} \end{aligned}$$

where the last equality follows from having $\frac{P_{k+1|k}^*}{(1-P_{k+1|k}^*)^2} < \frac{2}{1-P_{k+1|k}^*}$, since $P_{k+1|k}^* < \frac{1}{e}$. As mentioned, $-\sum_{l=-\infty}^{-2} P_{k+l|k} \log P_{k+l|k}$ can be upper bounded by the same expression (using Lemma 10 (Part B)). Thus, it follows that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} & -\sum_{l=-\infty}^{-2} P_{k+l|k} \log P_{k+l|k} - \sum_{l=2}^{\infty} P_{k+l|k} \log P_{k+l|k} \\ &< -P_{k+1|k}^* (\log P_{k+1|k}^*) \frac{8P_{k+1|k}^*}{1-P_{k+1|k}^*}. \end{aligned}$$

Using this upper bound in the numerator of (39), we obtain that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned}
& \frac{\sum_{l=-\infty}^{-2} H_{k+l|k} + \sum_{l=2}^{\infty} H_{k+l|k}}{H_{k-1,k,k+1|k}} \\
& < \frac{-P_{k+1|k}^* \left(\log P_{k+1|k}^* \right) \frac{8P_{k+1|k}^*}{1-P_{k+1|k}^*}}{-\sum_{l=k-1}^{k+1} P_{l|k} \log P_{l|k}} \\
& \stackrel{(a)}{<} \frac{-P_{k+1|k}^* \left(\log P_{k+1|k}^* \right) \frac{8P_{k+1|k}^*}{1-P_{k+1|k}^*}}{-P_{k+1|k}^* \log P_{k+1|k}^*} = \frac{8P_{k+1|k}^*}{1-P_{k+1|k}^*} \\
& \stackrel{(b)}{<} 16P_{k+1|k}^* \stackrel{(c)}{<} 16 \frac{\sqrt{2\pi}}{4\lambda} e^{\frac{\lambda^2}{2}} \sqrt{1-\rho^2} \\
& \longrightarrow 0 \text{ as } \rho \longrightarrow 1
\end{aligned}$$

where (a) derives from the fact that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$, $-P_{k+1|k}^* \log P_{k+1|k}^* < -P_{k+1|k} \log P_{k+1|k}$, which follows from Fact C8, given that $P_{k+1|k}^* < P_{k+1|k}$ and $P_{k+1|k} < \frac{1}{e}$, as shown by Lemmas 8 and 9 (Part B). (b) is due to having $P_{k+1|k}^* < \frac{1}{e}$ and so $\frac{1}{1-P_{k+1|k}^*} < 2$. Finally, (c) is obtained using Lemma 7 (Part C). This shows (39) and completes the proof of the lemma. \square

Proof of Lemma 12: We start by recalling from (6) that

$$f_{X_2|I_1}(x|k) = \int_{t_k}^{t_{k+1}} f_{X_2|X_1}(x|y) \frac{f_{X_1}(y)}{P_k} dy$$

where $f_{X_2|X_1}(x|y)$ is a Gaussian density whose mean is ρy . Thus, it is easy to see that for $x < t_k\rho$

$$\begin{aligned}
f_{X_2|I_1}(x|k) & < \int_{t_k}^{t_{k+1}} f_{X_2|X_1}(x|t_k) \frac{f_{X_1}(y)}{P_k} dy \\
& = f_{X_2|X_1}(x|t_k) \triangleq \bar{f}_k^L(x).
\end{aligned}$$

For tractability, let $\bar{f}_k^L(x) = f_{X_2|X_1}(x|t_k)$ for all x (not just for $x < t_k\rho$). Similarly, for $x > t_{k+1}\rho$

$$\begin{aligned}
f_{X_2|I_1}(x|k) & < \int_{t_k}^{t_{k+1}} f_{X_2|X_1}(x|t_{k+1}) \frac{f_{X_1}(y)}{P_k} dy \\
& = f_{X_2|X_1}(x|t_{k+1}) \\
& \triangleq \bar{f}_k^R(x).
\end{aligned}$$

As before, let $\bar{f}_k^R(x) = f_{X_2|X_1}(x|t_{k+1})$ for all x (not just for $x > t_{k+1}\rho$).

Letting $\bar{f}_k^C(x) = f_{X_2|I_1}(x|k)$ for $t_k\rho \leq x \leq t_{k+1}\rho$, and zero for all other x 's, we define

$$\bar{f}_k(x) = \begin{cases} \bar{f}_k^L(x), & x < t_k\rho \\ \bar{f}_k^C(x), & t_k\rho \leq x \leq t_{k+1}\rho \\ \bar{f}_k^R(x), & x > t_{k+1}\rho \end{cases}$$

so that $f_{X_2|I_1}(x|k) \leq \bar{f}_k(x)$. We break $f_{X_2|I_1}(x|k)$ in a similar manner. Specifically

$$f_{X_2|I_1}(x|k) = \begin{cases} f_k^L(x), & x < t_k\rho \\ f_k^C(x), & t_k\rho \leq x \leq t_{k+1}\rho \\ f_k^R(x), & x > t_{k+1}\rho \end{cases}$$

where f_k^L , f_k^C and f_k^R are zero outside their respective regions. Next, we have

$$\begin{aligned}
H(I_2|I_1 = k) & = H_q(f_{X_2|I_1=k}) \\
& < H_q(f_k^L) + H_q(f_k^C) + H_q(f_k^R) \quad (40)
\end{aligned}$$

where the inequality follows from the fact that for those quantization cells where two of the three functions f_k^L , f_k^C and f_k^R are nonzero (i.e., the quantization cells containing $t_k\rho$ and $t_{k+1}\rho$), the right-hand side is larger, as shown by Fact C9. Similarly

$$H_q(\bar{f}_k) < H_q(\bar{f}_k^L) + H_q(\bar{f}_k^C) + H_q(\bar{f}_k^R) \quad (41)$$

where the inequality follows similarly to (40). Specifically, since \bar{f}_k^L and \bar{f}_k^R are nonzero outside the regions $x < t_k\rho$ and $x > t_{k+1}\rho$, respectively, it follows that for all cells there are at least two functions that are nonzero, and in the cells containing $t_k\rho$ and $t_{k+1}\rho$, three functions are nonzero. In all cases, Fact C9 implies that the right-hand side is larger.

Clearly, $H_q(f_k^C) = H_q(\bar{f}_k^C)$, since the two functions are equal. We claim that $H_q(f_k^L) < H_q(\bar{f}_k^L) - 3\frac{1}{e} \log \frac{1}{e}$ and that $H_q(f_k^R) < H_q(\bar{f}_k^R) - 3\frac{1}{e} \log \frac{1}{e}$. We will show the first claim and note that the second claim follows via the same arguments. By definition $f_k^L < \bar{f}_k^L$. Thus, letting n be such that $t_k\rho \in S_n$, and letting $\bar{P}_{l|k} \triangleq \int_{S_l} \bar{f}_k(x) dx$, we have that $P_{l|k} < \bar{P}_{l|k}$, for any $l < n$. For those l 's for which $\bar{P}_{l|k} < \frac{1}{e}$, Fact C8 implies that $-P_{l|k} \log P_{l|k} < -\bar{P}_{l|k} \log \bar{P}_{l|k}$. Since \bar{f}_k^L integrates to one (and is nonzero), it follows that there can be at most two cells for which $\bar{P}_{l|k} \geq \frac{1}{e}$. Thus, for all $l < n$, except for at most two cells, we have $-P_{l|k} \log P_{l|k} < -\bar{P}_{l|k} \log \bar{P}_{l|k}$. The contribution to $H_q(f_k^L)$ of those cells for which $-P_{l|k} \log P_{l|k} \geq -\bar{P}_{l|k} \log \bar{P}_{l|k}$ is upper bounded by $-2\frac{1}{e} \log \frac{1}{e}$, where $\max_p \{-p \log p\} = -\frac{1}{e} \log \frac{1}{e}$. Lastly, the contribution of cell n is upper bounded by $-\frac{1}{e} \log \frac{1}{e}$. Consequently, we obtain that $H_q(f_k^L) < H_q(\bar{f}_k^L) - 3\frac{1}{e} \log \frac{1}{e}$. It follows from the three upper bounds in this paragraph that

$$H(I_2|I_1 = k) < H_q(\bar{f}_k) - 6\frac{1}{e} \log \frac{1}{e}. \quad (42)$$

It remains to upper bound $H_q(\bar{f}_k)$. Recalling from (41) that $H_q(\bar{f}_k) < H_q(\bar{f}_k^L) + H_q(\bar{f}_k^C) + H_q(\bar{f}_k^R)$, we observe that (i) since \bar{f}_k^L and \bar{f}_k^R are Gaussian densities with variance $\sigma_\rho^2 = \sigma^2(1-\rho^2)$, it follows from Lemma D2 of the Appendix that $H_q(\bar{f}_k^L)$ and $H_q(\bar{f}_k^R)$ are each upper bounded by $b(\lambda\sqrt{1-\rho^2})$, where $b(\cdot)$ is given in Lemma D2, and (ii) since \bar{f}_k^C is nonzero in at most two cells, $H_q(\bar{f}_k^C) \leq -2\frac{1}{e} \log \frac{1}{e}$. From this and

(42), it follows that $H_q(\tilde{f}_k) < 2b(\lambda\sqrt{1-\rho^2}) - 8\frac{1}{e} \log \frac{1}{e} < 2b(\lambda\sqrt{1-\rho^2}) + 5$, for all ρ and k . \square

Proof of Lemma 13: We begin by lower and upper bounding the expression in the lemma statement as follows:

$$1 < \frac{\tilde{H}(I_2|I_1)}{\sum_{k=1}^{K(\rho)} H(I_2|I_1 = k)P_k + \frac{1}{2}H(I_2|I_1 = 0)P_0} < 1 + \frac{\sum_{k=K(\rho)+1}^{\infty} H(I_2|I_1 = k)P_k}{\frac{1}{2}H(I_2|I_1 = 0)P_0}. \quad (43)$$

It suffices to show that $\lim_{\rho \rightarrow 1} \frac{\sum_{k=K(\rho)+1}^{\infty} H(I_2|I_1 = k)P_k}{\frac{1}{2}H(I_2|I_1 = 0)P_0} = 0$. To do so, we upper bound the numerator and lower bound the denominator. Consider the numerator first. For all ρ sufficiently close to one

$$\begin{aligned} \sum_{k=K(\rho)+1}^{\infty} H(I_2|I_1 = k)P_k &\stackrel{(a)}{<} B(\lambda\sqrt{1-\rho^2}) \sum_{k=K(\rho)+1}^{\infty} P_k \\ &= B(\lambda\sqrt{1-\rho^2})Q\left(\frac{tK(\rho)+1}{\sigma}\right) \\ &\stackrel{(b)}{\leq} \frac{B(\lambda\sqrt{1-\rho^2})}{2} e^{-\frac{(K(\rho)+1-\theta)^2\lambda^2}{2}} \\ &\leq \frac{B(\lambda\sqrt{1-\rho^2})}{2} e^{-\frac{K(\rho)^2\lambda^2}{2}} \quad (44) \end{aligned}$$

where (a) is due to Lemma 12 with $B(\alpha) \triangleq 2b(\alpha) + 5$ and $b(\cdot)$ defined in Lemma D2, and (b) uses Fact C1.

Next, we lower bound the denominator. Specifically, for all ρ sufficiently close to one

$$\begin{aligned} H(I_2|I_1 = 0)P_0 &= \left(-\sum_{l=-\infty}^{\infty} P_{l|0} \log P_{l|0}\right) P_0 \\ &> (-P_{1|0} \log P_{1|0}) P_0 \\ &\stackrel{(a)}{>} (-P_{1|0}^* \log P_{1|0}^*) P_0 \\ &\stackrel{(b)}{>} \left(\frac{1}{5}R_0(\rho) \log\left(\frac{1}{5}R_0(\rho)\right)\right) P_0 \\ &= \left[-\frac{1}{10} \frac{1}{P_0} e^{-\frac{(1-\theta)^2\lambda^2}{2}} \sqrt{1-\rho^2} \times \log\left(\frac{1}{10} \frac{1}{P_0} e^{-\frac{(1-\theta)^2\lambda^2}{2}} \sqrt{1-\rho^2}\right)\right] P_0 \\ &\stackrel{(c)}{>} \frac{1}{10} e^{-\frac{(1-\theta)^2\lambda^2}{2}} \sqrt{1-\rho} \quad (45) \end{aligned}$$

where (a) follows from having $P_{1|0} < \frac{1}{e}$ for all ρ sufficiently close to one and $P_{1|0}^* < P_{1|0}$, which imply (using Fact C8) that $-P_{1|0}^* \log P_{1|0}^* < -P_{1|0} \log P_{1|0}$; (b) is due to Lemma 7 (Part D), and (c) derives from having $-\log\left(\frac{1}{10} \frac{1}{P_0} e^{-\frac{(1-\theta)^2\lambda^2}{2}} \sqrt{1-\rho^2}\right) > 1$, for all ρ sufficiently close to one, and having $1 < \sqrt{1+\rho}$.

Finally, plugging the upper bound given by (44) and the lower bound given by (45) into the last fraction in (43), we obtain that for all ρ sufficiently close to one

$$\frac{\sum_{k=K(\rho)+1}^{\infty} H(I_2|I_1 = k)P_k}{H(I_2|I_1 = 0)P_0} < \frac{B(\lambda\sqrt{1-\rho^2}) e^{-\frac{K^2(\rho)\lambda^2}{2}}}{\frac{1}{10} e^{-\frac{(1-\theta)^2\lambda^2}{2}} \sqrt{1-\rho}} \longrightarrow 0 \text{ as } \rho \longrightarrow 1$$

where the convergence holds since for all ρ sufficiently close to one

$$\begin{aligned} &\frac{\log \frac{1}{1-\rho}}{e^{\frac{K^2(\rho)\lambda^2}{2}} \sqrt{1-\rho}} \\ &\leq \frac{\log \frac{1}{1-\rho}}{e^{\left[\left(\ln \frac{1}{1-\rho}\right)^{\frac{3}{2}} - 2\left(\ln \frac{1}{1-\rho}\right)^{\frac{3}{4}} + 1\right] \frac{\lambda^2}{2}} e^{-\frac{1}{2}\left(\ln \frac{1}{1-\rho}\right)}} \\ &= \frac{\log \frac{1}{1-\rho}}{e^{\left(\ln \frac{1}{1-\rho}\right) \left[\frac{\lambda^2}{2} \left(\left(\ln \frac{1}{1-\rho}\right)^{\frac{1}{2}} - 2\left(\ln \frac{1}{1-\rho}\right)^{-\frac{1}{4}} + \left(\ln \frac{1}{1-\rho}\right)^{-1}\right) - \frac{1}{2}\right]}} \\ &\longrightarrow 0 \text{ as } \rho \longrightarrow 1 \end{aligned}$$

where the $\log \frac{1}{1-\rho}$ term in the numerator replaces $B(\lambda(\sqrt{1-\rho^2}))$, which tends to ∞ at this rate as $\rho \rightarrow 1$ as follows from the definition of $B(\cdot)$ and Lemma D2. This concludes the proof of the lemma. \square

Proof of Lemma 14: It follows from Lemma D1 (Part D) that it suffices to show

$$\lim_{\rho \rightarrow 1}^* \frac{H_{k-1,k,k+1|k}}{H_{k-1,k+1|k}} = 1.$$

Next, using Lemma D1 (Part B) and Fact C9, we obtain that it is enough to show

$$\lim_{\rho \rightarrow 1}^* \frac{\mathcal{H}(P_{k|k})}{\mathcal{H}(P_{k-1|k} + P_{k+1|k})} = 0. \quad (46)$$

In order to show (46), we write $P_{k|k} = 1 - P_{k-1|k} - P_{k+1|k} - P_{tail|k} = 1 - \tilde{P}_k - P_{tail|k}$, where $\tilde{P}_k \triangleq P_{k-1|k} + P_{k+1|k}$ and $P_{tail|k} \triangleq \sum_{l=-\infty}^{-2} P_{k+l|k} + \sum_{l=2}^{\infty} P_{k+l|k}$. With this notation, (46) becomes

$$\lim_{\rho \rightarrow 1}^* \frac{\mathcal{H}(1 - \tilde{P}_k - P_{tail|k})}{\mathcal{H}(\tilde{P}_k)} = 0. \quad (47)$$

We proceed by upper bounding the numerator in (47). To do so, we upper bound $P_{tail|k}$ in terms of \tilde{P}_k and then use Fact C8. From Lemma 10, we have that for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$

$$\begin{aligned} P_{tail|k} &= \sum_{l=-\infty}^{k-2} P_{l|k} + \sum_{l=k+2}^{\infty} P_{l|k} < 2 \sum_{l=0}^{\infty} (P_{k+1|k}^*)^{l+2} \\ &= \frac{2(P_{k+1|k}^*)^2}{1 - P_{k+1|k}^*} < \frac{2(P_{k+1|k}^*)}{1 - P_{k+1|k}^*} \tilde{P}_k \\ &< \tilde{P}_k \end{aligned}$$

where the last inequality follows from Lemma 7 (Part C), which implies that $\lim_{\rho \rightarrow 1}^* P_{k+1|k}^* = 0$.

Next, we would like to use Fact C8 in order to get an upper bound for the numerator in (47). We now show that the conditions required by Fact C8 are indeed met. Lemma D1 (Part B) and Lemma 9 imply that $\lim_{\rho \rightarrow 1}^* \tilde{P}_k = 0$. Therefore, for all ρ sufficiently close to one and $0 \leq k \leq K(\rho)$, $1 - \tilde{P}_k - P_{tail|k} > 1 - 2\tilde{P}_k > \frac{1}{e}$. Consequently, it follows from Fact C8 that $\mathcal{H}(1 - \tilde{P}_k - P_{tail|k}) < \mathcal{H}(1 - 2\tilde{P}_k)$. Combining this with (47), it follows that it is enough to show

$$\lim_{\rho \rightarrow 1}^* \frac{\mathcal{H}(1 - 2\tilde{P}_k)}{\mathcal{H}(\tilde{P}_k)} = 0. \quad (48)$$

Finally, (48) can be seen to hold using the fact, shown earlier, that $\lim_{\rho \rightarrow 1}^* \tilde{P}_k = 0$, together with Lemmas D3 and D1 (Part F). This concludes the proof of the lemma. \square

Proof of Lemma 15: It follows from Lemma D1 (Part D) that it suffices to show

$$\lim_{\rho \rightarrow 1}^* \frac{H_{k-1, k+1|k} P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{L, R, \lambda}(k)} = 1.$$

Next, recalling that $W_{L, R, \lambda}(k) = W_{L, \lambda}(k) + W_{R, \lambda}(k)$ and applying Lemma D1 (Part E), it follows that it suffices to show the following:

$$\text{A. } \lim_{\rho \rightarrow 1}^* \frac{H_{k-1|k} P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{L, \lambda}(k)} = 1;$$

$$\text{B. } \lim_{\rho \rightarrow 1}^* \frac{H_{k+1|k} P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{R, \lambda}(k)} = 1.$$

To show A, we write

$$\begin{aligned} & \frac{H_{k-1|k} P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{L, \lambda}(k)} \\ &= \frac{H_{k-1|k} P_k}{\mathcal{H}(\frac{1}{\pi} L_k(\rho)) P_k} \frac{\mathcal{H}(\frac{1}{\pi} L_k(\rho)) P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{L, \lambda}(k)}. \end{aligned}$$

Lemmas 7 (Part A), D4, and D1 (Part G) (with $G(x, y) = \frac{\mathcal{H}(x)}{\mathcal{H}(y)}$) show that the first term on the right-hand side has \lim^* equal one. Therefore, by Lemma D1 (Part C), it suffices to show

$$\lim_{\rho \rightarrow 1}^* \frac{\mathcal{H}(\frac{1}{\pi} L_k(\rho)) P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{L, \lambda}(k)} = 1. \quad (49)$$

To show this, we observe that by the definitions of $W_{L, \lambda}(k)$ and $L_k(\rho)$ one can show

$$\begin{aligned} \frac{\mathcal{H}(\frac{1}{\pi} L_k(\rho)) P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{L, \lambda}(k)} &= \frac{-\frac{1}{\pi} L_k(\rho) \log(\frac{1}{\pi} L_k(\rho))}{-\frac{1}{\pi} L_k(\rho) \log \sqrt{1 - \rho^2}} \\ &= \frac{\log(\frac{1}{\pi} L_k(\rho))}{\log \sqrt{1 - \rho^2}} \\ &= 1 + \frac{\log\left(\frac{1}{\pi} \frac{L_k(\rho)}{\sqrt{1 - \rho^2}}\right)}{\log \sqrt{1 - \rho^2}}. \end{aligned}$$

We now show that the fraction on the right-hand side above converges to zero as $\rho \rightarrow 1$ in the \lim^* sense, from which (49) will follow via Lemma D1 (Part B).

Consider first $K_\lambda < k \leq K(\rho)$, where K_λ is as defined in Lemma 6. For all ρ sufficiently close to one

$$\begin{aligned} \frac{\log\left(\frac{1}{\pi} \frac{L_k(\rho)}{\sqrt{1 - \rho^2}}\right)}{\left|\log \sqrt{1 - \rho^2}\right|} &\stackrel{(a)}{<} \frac{\log\left(\sqrt{\frac{2}{\pi}} k \lambda\right)}{\left|\log \sqrt{1 - \rho^2}\right|} \\ &\stackrel{(b)}{<} \frac{\log\left(\sqrt{\frac{2}{\pi}} K(\rho) \lambda\right)}{\left|\log \sqrt{1 - \rho^2}\right|} \\ &\longrightarrow 0 \text{ as } \rho \longrightarrow 1 \end{aligned} \quad (50)$$

where (a) is due to Lemma 6 (Part A) and observing that since $k > K_\lambda > \frac{2}{\lambda}$ the logarithm in the numerator is positive for all such k 's, (b) is due to upper bounding k by $K(\rho)$, and the convergence can be easily seen to hold. We also have from Lemma 6 (Part B), which can be used since $k > K_\lambda \geq 1$, that $\frac{1}{\pi} \frac{L_k(\rho)}{\sqrt{1 - \rho^2}} > \frac{1}{\pi}$. Thus

$$\frac{\log\left(\frac{1}{\pi} \frac{L_k(\rho)}{\sqrt{1 - \rho^2}}\right)}{\left|\log \sqrt{1 - \rho^2}\right|} > \frac{\log \frac{1}{\pi}}{\left|\log \sqrt{1 - \rho^2}\right|} \longrightarrow 0 \text{ as } \rho \longrightarrow 1. \quad (51)$$

Equations (50) and (51) show

$$\begin{aligned} \frac{\log\left(\frac{1}{\pi} \frac{L_k(\rho)}{\sqrt{1 - \rho^2}}\right)}{\log \sqrt{1 - \rho^2}} &\longrightarrow 0 \text{ as } \rho \longrightarrow 1 \\ &\text{uniformly for } K_\lambda < k \leq K(\rho). \end{aligned} \quad (52)$$

Consider now $0 \leq k \leq K_\lambda$. For such k 's

$$\log\left(\frac{1}{\pi} \frac{L_k(\rho)}{\sqrt{1 - \rho^2}}\right) = \log\left(\frac{1}{2\pi} \frac{1}{P_k} e^{-\frac{t_k^2}{2\sigma^2}}\right)$$

can assume only finitely many values. Therefore, it follows that

$$\frac{\log\left(\frac{1}{\pi} \frac{L_k(\rho)}{\sqrt{1 - \rho^2}}\right)}{\log \sqrt{1 - \rho^2}} \rightarrow 0 \text{ as } \rho \rightarrow 1 \text{ uniformly for } 0 \leq k \leq K_\lambda. \quad (53)$$

Together, (52) and (53) imply

$$\lim_{\rho \rightarrow 1}^* \frac{\log\left(\frac{1}{\pi} \frac{L_k(\rho)}{\sqrt{1 - \rho^2}}\right)}{\log \sqrt{1 - \rho^2}} = 0$$

which completes the proof of (49).

Similarly, to show B, we write

$$\begin{aligned} & \frac{H_{k+1|k} P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{R, \lambda}(k)} \\ &= \frac{H_{k+1|k} P_k}{\mathcal{H}(\frac{1}{\pi} R_k(\rho)) P_k} \frac{\mathcal{H}(\frac{1}{\pi} R_k(\rho)) P_k}{\mathcal{H}(\sqrt{1 - \rho^2}) W_{R, \lambda}(k)}. \end{aligned}$$

Lemmas 7 (Part B), 8, D1 (Parts C and G), and D4 show that the first term on the right-hand side has \lim^* equal one. Therefore, by Lemma D1 (Part C), it suffices to show

$$\lim_{\rho \rightarrow 1}^* \frac{\mathcal{H}\left(\frac{1}{\pi} R_k(\rho)\right) P_k}{\mathcal{H}(\sqrt{1-\rho^2}) W_{R,\lambda}(k)} = 1. \quad (54)$$

From the definitions of $W_{R,\lambda}(k)$ and $R_k(\rho)$, one can show

$$\begin{aligned} \frac{\mathcal{H}\left(\frac{1}{\pi} R_k(\rho)\right) P_k}{\mathcal{H}(\sqrt{1-\rho^2}) W_{R,\lambda}(k)} &= \frac{-\frac{1}{\pi} R_k(\rho) \log\left(\frac{1}{\pi} R_k(\rho)\right)}{-\frac{1}{\pi} R_k(\rho) \log \sqrt{1-\rho^2}} \\ &= \frac{\log\left(\frac{1}{\pi} R_k(\rho)\right)}{\log \sqrt{1-\rho^2}} \\ &= 1 + \frac{\log\left(\frac{1}{\pi} \frac{R_k(\rho)}{\sqrt{1-\rho^2}}\right)}{\log \sqrt{1-\rho^2}}. \end{aligned}$$

We now show that the fraction on the right-hand side above converges to zero as $\rho \rightarrow 1$ in the \lim^* sense, from which (54) will follow via Lemma D1 (Part B).

On the one hand, we have from Lemma 6 (Part C) that $\frac{1}{\pi} \frac{R_k(\rho)}{\sqrt{1-\rho^2}} < \frac{1}{\lambda \sqrt{2\pi}} e^{\frac{\lambda^2}{2}}$, for $k \geq 0$. On the other hand, Lemma 6 (Part D) implies that for $1 \leq k \leq K(\rho)$

$$\frac{1}{\pi} \frac{R_k(\rho)}{\sqrt{1-\rho^2}} > \frac{1}{\pi} e^{-\frac{\lambda^2}{2}} e^{-k\lambda^2} \geq \frac{1}{\pi} e^{-\frac{\lambda^2}{2}} e^{-K(\rho)\lambda^2}.$$

Therefore, it follows that

$$\frac{\log\left(\frac{1}{\pi} \frac{R_k(\rho)}{\sqrt{1-\rho^2}}\right)}{\log \sqrt{1-\rho^2}} \rightarrow 0 \text{ as } \rho \rightarrow 1 \text{ uniformly for } 1 \leq k \leq K(\rho) \quad (55)$$

as can be seen from the fact that

$$\begin{aligned} \frac{\log e^{-K(\rho)\lambda^2}}{\log \sqrt{1-\rho^2}} &= \frac{4K(\rho)\lambda^2}{\ln \frac{1}{1-\rho^2}} \leq \frac{4\lambda^2 \left(\ln \frac{1}{1-\rho}\right)^{\frac{3}{4}}}{\ln \frac{1}{1-\rho^2}} \\ &\rightarrow 0 \text{ as } \rho \rightarrow 1. \end{aligned}$$

Since (55) also holds for $k = 0$, it follows that the convergence holds in the \lim^* sense. This completes the proof of (54) and concludes the proof of the lemma. \square

Proof of Lemma 16: It needs to be shown that

$$\lim_{\rho \rightarrow 1} \frac{\mathcal{H}(\sqrt{1-\rho^2}) \left[\sum_{k=1}^{K(\rho)} W_{L,R,\lambda}(k) + \frac{1}{2} W_{L,R,\lambda}(0) \right]}{W_\lambda \mathcal{H}(\sqrt{1-\rho})} = 1.$$

First, we observe that

$$\begin{aligned} \frac{\mathcal{H}(\sqrt{1-\rho^2})}{\mathcal{H}(\sqrt{1-\rho})} &= \frac{-\sqrt{1-\rho^2} \log \sqrt{1-\rho^2}}{-\sqrt{1-\rho} \log \sqrt{1-\rho}} \\ &\rightarrow \sqrt{2} \text{ as } \rho \rightarrow 1. \end{aligned}$$

Therefore, it needs to be shown that

$$\lim_{\rho \rightarrow 1} \sqrt{2} \left[\sum_{k=1}^{K(\rho)} W_{L,R,\lambda}(k) + \frac{1}{2} W_{L,R,\lambda}(0) \right] = W_\lambda.$$

Recalling that $W_{L,\lambda}(k) = \frac{1}{2\pi} e^{-\frac{(k-\theta)^2 \lambda^2}{2}}$, $W_{R,\lambda}(k) = \frac{1}{2\pi} e^{-\frac{(k+1-\theta)^2 \lambda^2}{2}}$, and $W_\lambda = \frac{\sqrt{2}}{\pi} \sum_{k=0}^{\infty} e^{-\frac{(k+1-\theta)^2 \lambda^2}{2}} + \frac{\sqrt{2}}{4\pi} \left(e^{-\frac{\theta^2 \lambda^2}{2}} - e^{-\frac{(1-\theta)^2 \lambda^2}{2}} \right)$, the above can straightforwardly be shown in the following way:

$$\begin{aligned} &\sqrt{2} \sum_{k=1}^{K(\rho)} (W_{L,\lambda}(k) + W_{R,\lambda}(k)) \\ &\quad + \frac{\sqrt{2}}{2} (W_{L,\lambda}(0) + W_{R,\lambda}(0)) \\ &= \sqrt{2} \sum_{k=1}^{K(\rho)} \frac{1}{2\pi} \left(e^{-\frac{(k-\theta)^2 \lambda^2}{2}} + e^{-\frac{(k+1-\theta)^2 \lambda^2}{2}} \right) \\ &\quad + \frac{\sqrt{2}}{4\pi} \left(e^{-\frac{\theta^2 \lambda^2}{2}} + e^{-\frac{(1-\theta)^2 \lambda^2}{2}} \right) \\ &= \sum_{k=0}^{K(\rho)} \frac{\sqrt{2}}{2\pi} e^{-\frac{(k+1-\theta)^2 \lambda^2}{2}} + \sum_{k=1}^{K(\rho)} \frac{\sqrt{2}}{2\pi} e^{-\frac{(k+1-\theta)^2 \lambda^2}{2}} \\ &\quad + \frac{\sqrt{2}}{4\pi} \left(e^{-\frac{\theta^2 \lambda^2}{2}} + e^{-\frac{(1-\theta)^2 \lambda^2}{2}} \right) \\ &= \frac{\sqrt{2}}{\pi} \sum_{k=0}^{K(\rho)} e^{-\frac{(k+1-\theta)^2 \lambda^2}{2}} + \frac{\sqrt{2}}{4\pi} \left(e^{-\frac{\theta^2 \lambda^2}{2}} - e^{-\frac{(1-\theta)^2 \lambda^2}{2}} \right) \\ &\xrightarrow{\rho \rightarrow 1} W_\lambda \end{aligned}$$

as follows from the fact that $K(\rho) \rightarrow \infty$ as $\rho \rightarrow 1$. \square

VI. CONCLUDING REMARKS

The principal result of the first part of this paper, Theorem 1, shows that the joint entropy of identically scalar quantized samples of a stationary continuous-time random process, taken over a finite time interval, converges to infinity as the sampling rate increases. The only requirement is that the process cross a quantizer threshold with positive probability. Note that if the latter does not occur, then the quantizer is essentially useless.

The contrast between the infinite entropy limit and the finite rate of ideal lossy source coding at the same distortion was described. It is interesting to ask how much the considered system would need to change in order for the entropy in bits/sec to remain bounded. For example, suppose that instead of losslessly coding just the N quantized samples in the unit time interval at rate equal to their joint entropy $H(I_1, \dots, I_N)$ bits/sec, one losslessly codes the infinite sequence of quantized samples at rate $NH_\infty(I)$ bits/sec, which is the sampling rate times the entropy-rate of the sequence. $NH_\infty(I)$ can be no larger than $H(I_1, \dots, I_N)$, which goes to infinity. But does $NH_\infty(I)$ also tend to infinity? Though we conjecture that ordinarily it does, a detailed verification is required.

Next, suppose that the scalar quantizer is replaced by a K -dimensional vector quantizer (VQ), for some fixed K . Does the joint entropy of the quantized samples in the unit time interval again increase to infinity? Informal arguments, somewhat like those in Section II, suggest that it does. This would lead one to believe that to attain a bounded rate (in bits/sec), the dimension of the VQ should increase as the sampling rate increases.

The dithered scalar quantization schemes considered in [7], [8], and [10] use periodic dither with a period (in samples) that increases to infinity as the sampling rate increases. In effect, these are VQs with dimension increasing to infinity. Arguments like those in Section II suggest that the entropy of the quantized samples in the unit time interval again goes to infinity. Now, though, distortion goes to zero. However, when the target distortion is fixed, it is not known if dithered scalar quantization has entropy in bits/sec remaining finite or increasing to infinity as the sampling rate increases.

For the motivating sensor network scenario, in which the samples are separated by space rather than time and for which distributed coding is required, it is interesting to ask whether allowing ideal distributed lossy coding (rather than scalar quantization) at some fixed target MSE would again have rate in bits/sec tending to infinity. For the important case that the underlying random process is Gaussian, this question was answered by Kashyap *et al.* [26], who found a finite upper bound to the rates attainable by distributed lossy coding over the unit space interval as sampling rate increases. This shows that at high sampling rates, ideal distributed lossy coding overcomes the principal shortcoming of scalar quantization. When coding is not restricted to a fixed finite space interval, a tighter upper bound was found in [27].

To gain some understanding of the rate at which joint entropy approaches infinity, Theorem 3 of the second part of this paper finds the asymptotic form of the conditional entropy $H(I_2|I_1)$, for a Gaussian process and a uniform threshold quantizer. As expressed in (1), this factors into one term that depends on the sampling rate N and the behavior of the autocorrelation function near the origin, and a second term that depends on the ratio of the quantizer step size to the process variance and (typically just a little) on the location of the mean of the process within the quantization cell that contains it. For large N , $NH(I_2|I_1)$ is essentially an upper bound to $H(I_1, \dots, I_N)$. Theorem 4 also found the asymptotic form of $H(I_2|I_1)$ for a situation in which the quantizer cells shrink as $N \rightarrow \infty$. It would be interesting to have a complementary expression that applies in the case of expanding quantizer cells. For example, one would like to know how fast the cells would need to expand as $N \rightarrow \infty$ in order that the limiting joint entropy $H(I_1, \dots, I_N)$ be finite. Another interesting problem is to find a similar formula to that of Theorem 3 when conditioning on more than one quantized sample.

APPENDIX A

The discussion below appears in more detail in [28 (pp. 41–45)] and in [29 (pp. 86–92)]. A random process $\{X_t(\omega), t \in T, \omega \in \Omega\}$, abbreviated X , defined on a probability space (Ω, \mathcal{F}, P) is said to be *separable* if there exists a

countable set $S \subseteq T$ and a fixed null event Λ such that for any closed set $K \subseteq [-\infty, \infty]$ and any open interval I , the two sets $\{\omega : X_t(\omega) \in K, t \in I \cap T\}$ and $\{\omega : X_t(\omega) \in K, t \in I \cap S\}$ differ by a subset of Λ . The countable set S is called a *separating set* or *separant*.

It follows from the definition of separability that when the underlying probability space is complete, for any $a \in \mathbb{R}$ the set $\{\omega : X_t(\omega) < a, t \in I \cap T\}$ is an event with the same probability as the event $\{\omega : X_t(\omega) < a, t \in I \cap S\}$, where S is a separating set.

The process X is continuous in probability, if for any $\varepsilon > 0$, $\Pr(|X_t - X_s| \leq \varepsilon) \rightarrow 1$ as $s \rightarrow t$. If X is both separable and continuous in probability, then every countable dense subset of T is a separating set. Thus, when X is also stationary (i.e., for any $N \in \mathbb{Z}_+$, $t_1, \dots, t_N \in T$, and $\tau \in \mathbb{R}$ such that $t_1 + \tau, \dots, t_N + \tau \in T$, $X_{t_1+\tau}, \dots, X_{t_N+\tau}$ has the same probability distribution as X_{t_1}, \dots, X_{t_N}), the two events $\{\omega : X_t(\omega) \leq r, t_1 < t < t_2\}$ and $\{\omega : X_t(\omega) \leq r, t_1 + s < t < t_2 + s\}$ have the same probability since one can find a countable set, which is dense in (t_1, t_2) and is a separating set, and use it to compute the probability of both events (since the shifted set is also a separating set). Therefore, stationarity along with separability and continuity in probability imply that shifting events of the above form does not alter their probabilities.

Next, a random process X is said to be *measurable* if $X_t(\omega)$ is a (t, ω) function measurable with respect to $\mathcal{B} \otimes \mathcal{F}$, where \mathcal{B} is the σ -algebra of Lebesgue measurable sets in T , and $\mathcal{B} \otimes \mathcal{F}$ is the σ -algebra generated by all sets of the form $B \times F$, $B \in \mathcal{B}$, $F \in \mathcal{F}$, i.e., for any $x \in (-\infty, \infty)$, $\{(t, \omega) : X_t(\omega) \leq x\} \in \mathcal{B} \otimes \mathcal{F}$.

Let $\Phi(t)$ be an indicator function, which is one if some specified event occurs in X at time t and zero otherwise. Let I be some Lebesgue measurable set, for example, an interval. If Φ is a measurable function with respect to $\mathcal{B} \otimes \mathcal{F}$, then it follows from Fubini's theorem that $E[\int_I \Phi(t) dt] = \int_I E[\Phi(t)] dt$. Thus, the measurability of the process X ensures the correctness of swapping integration and expectation above.

Finally, if the process X is continuous in probability, and T is an interval, then there exists a process \tilde{X} defined on the same probability space, which is separable, measurable, and equivalent to X in the sense that $\Pr(\tilde{X}_t = X_t) = 1$, for all $t \in T$, which implies that \tilde{X} and X have the same probability distribution for any finite collection of random variables. This implies that continuity in probability suffices to guarantee that basic operations on stationary continuous-time random processes such as preservation of probability under shifting and exchange of integration and expectation can be performed.

APPENDIX B

We show in this appendix that $\lim_{\lambda \rightarrow 0} \frac{M_\lambda^\theta}{\sqrt{\pi}} = 1$, where $M_\lambda^\theta = \frac{\sqrt{2}}{\pi} \sum_{k=-\infty}^{\infty} e^{-\frac{(k+1-\theta)^2 \lambda^2}{2}}$. We first observe that M_λ^θ can be rewritten as

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

Next

$$\begin{aligned} \sum_{k=0}^{\infty} e^{-\frac{(k+\theta)^2\lambda^2}{2}} &= \frac{\sqrt{2\pi}}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{(k+\theta)^2\lambda^2}{2}} \\ &> \frac{\sqrt{2\pi}}{\lambda} \int_{\theta}^{\infty} \mathcal{N}_{0, \frac{1}{\lambda^2}}(x) dx \\ &= \frac{\sqrt{2\pi}}{\lambda} Q(\theta\lambda). \end{aligned}$$

Similarly

$$\begin{aligned} \sum_{k=0}^{\infty} e^{-\frac{(k+\theta)^2\lambda^2}{2}} &= \frac{\sqrt{2\pi}}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{(k+\theta)^2\lambda^2}{2}} \\ &< \frac{\sqrt{2\pi}}{\lambda} \int_{\theta}^{\infty} \mathcal{N}_{0, \frac{1}{\lambda^2}}(x) dx + e^{-\frac{\theta^2\lambda^2}{2}} \\ &= \frac{\sqrt{2\pi}}{\lambda} Q(\theta\lambda) + e^{-\frac{\theta^2\lambda^2}{2}}. \end{aligned}$$

Since $\lim_{\lambda \rightarrow 0} Q(\theta\lambda) = \frac{1}{2}$ and $\lim_{\lambda \rightarrow 0} e^{-\frac{\theta^2\lambda^2}{2}} = 1$, it follows that $\lim_{\lambda \rightarrow 0} \frac{\sum_{k=0}^{\infty} e^{-\frac{(k+\theta)^2\lambda^2}{2}}}{\sqrt{\frac{\pi}{2}\lambda}} = 1$. Similarly, the same limit applies for $\sum_{k=0}^{\infty} e^{-\frac{(k+1-\theta)^2\lambda^2}{2}}$, as needed to show.

APPENDIX C

The following are elementary facts about the Q , g and \mathcal{H} functions, where we recall that $Q(\alpha) = \int_{\alpha}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ and $g(\alpha) = \frac{Q(\alpha)}{\frac{1}{2}e^{-\frac{\alpha^2}{2}}}$, $\alpha \geq 0$.

- Fact C1: $Q(x) \leq \frac{1}{2}e^{-\frac{x^2}{2}}$, for $x \geq 0$.
- Fact C2: $Q(x) < \frac{1}{\sqrt{2\pi x}} e^{-\frac{x^2}{2}}$, for $x > 0$.
- Fact C3: $Q(x) > \frac{1}{\sqrt{2\pi x}} (1 - \frac{1}{x^2}) e^{-\frac{x^2}{2}}$, for $x > 0$.
- Fact C4: $Q(x\lambda) - Q((x+1)\lambda) > \frac{1}{2\sqrt{2\pi x\lambda}} e^{-\frac{x^2\lambda^2}{2}}$, for $x > \max\{\frac{2}{\lambda}, \frac{2}{\lambda^2}\}$ and $\lambda > 0$.
- Fact C5: $\frac{Q((a+1)z)}{Q(az)} < 2e^{-\frac{z^2}{2}}$, when $a, z > 0$ and $az > \frac{1}{2}$.
- Fact C6: $g(0) = 1, g(\infty) = 0$ and $g(x)$ is a strictly decreasing function of x , for $x \geq 0$.
- Fact C7: $-p \log p$ is concave and attains its maximum at $p = \frac{1}{e}$.
- Fact C8: $p < p' < \frac{1}{e}$ implies $-p \log p < -p' \log p'$, and $p > p' > \frac{1}{e}$ implies $-p \log p < -p' \log p'$.
- Fact C9: For any $\{a_k\} \in \mathbb{R}^+$, $\mathcal{H}(\sum_k a_k) < \sum_k \mathcal{H}(a_k)$.

Facts C1, C2, and C3 are shown in [24 (pp. 82–83)]. Facts C4 and C5 can be shown straightforwardly using Facts C2 and C3. The first two parts of Fact C6 are immediate, and the monotonicity part follows from having $g'(x) < 0$ for all $x \geq 0$, which can be shown using Fact C2. Fact C7 is well-known. Fact C8 is a direct consequence of Fact C7, and Fact C9 is due to the concavity of \mathcal{H} .

APPENDIX D

We recall first that $\lim_{\rho \rightarrow 1}^* f(k, \rho) = c$ means that $\lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} |f(k, \rho) - c| = 0$. Equiv-

alently, we may write $\lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} f(k, \rho) = \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} f(k, \rho) = c$.

Lemma D1: For any nonnegative choice of $K(\rho)$

- A. (Dominated Convergence Theorem) If $\sup_{0 \leq k \leq K(\rho)} |f(k, \rho, x)| \leq G(x)$ a.e., for some integrable function G and $\lim_{\rho \rightarrow 1}^* f(k, \rho, x)$ exists a.e., then $\lim_{\rho \rightarrow 1}^* \int f(k, \rho, x) dx = \int \lim_{\rho \rightarrow 1}^* f(k, \rho, x) dx$.
- B. If $\lim_{\rho \rightarrow 1}^* a_k(\rho)$ and $\lim_{\rho \rightarrow 1}^* b_k(\rho)$ exist, then $\lim_{\rho \rightarrow 1}^* (a_k(\rho) + b_k(\rho)) = \lim_{\rho \rightarrow 1}^* a_k(\rho) + \lim_{\rho \rightarrow 1}^* b_k(\rho)$.
- C. Let $a_k(\rho)$ and $b_k(\rho)$ be nonnegative for all ρ and $0 \leq k \leq K(\rho)$. If $\lim_{\rho \rightarrow 1}^* a_k(\rho)$ and $\lim_{\rho \rightarrow 1}^* b_k(\rho)$ exist, then $\lim_{\rho \rightarrow 1}^* a_k(\rho)b_k(\rho) = \lim_{\rho \rightarrow 1}^* a_k(\rho) \times \lim_{\rho \rightarrow 1}^* b_k(\rho)$.
- D. Let $b_k(\rho)$ be positive for all ρ and $0 \leq k \leq K(\rho)$. If $\lim_{\rho \rightarrow 1}^* \frac{a_k(\rho)}{b_k(\rho)} = 1$, then $\lim_{\rho \rightarrow 1} \frac{\sum_{k=0}^{K(\rho)} a_k(\rho)}{\sum_{k=0}^{K(\rho)} b_k(\rho)} = 1$.
- E. Let $c_k(\rho)$ and $d_k(\rho)$ be positive for all ρ and $0 \leq k \leq K(\rho)$. If $\lim_{\rho \rightarrow 1}^* \frac{a_k(\rho)}{c_k(\rho)} = 1$ and $\lim_{\rho \rightarrow 1}^* \frac{b_k(\rho)}{d_k(\rho)} = 1$, then $\lim_{\rho \rightarrow 1}^* \frac{a_k(\rho)+b_k(\rho)}{c_k(\rho)+d_k(\rho)} = 1$.
- F. If $\lim_{z \rightarrow z_0} G(z) = c$ and $\lim_{\rho \rightarrow 1}^* a_k(\rho) = z_0$, then $\lim_{\rho \rightarrow 1}^* G(a_k(\rho)) = c$.
- G. If $\lim_{\frac{x}{y} \rightarrow z} G(x, y) = c$ and $\lim_{\rho \rightarrow 1}^* \frac{a_k(\rho)}{b_k(\rho)} = z$, then $\lim_{\rho \rightarrow 1}^* G(a_k(\rho), b_k(\rho)) = c$.

Proof: We show the statements of the lemma in the following order: G, F, C, B, D, E, and A. Consider Part G. Let $\varepsilon > 0$ be given. Then by assumption there exists $\delta > 0$ such that if $|\frac{x}{y} - z| < \delta$, then $|G(x, y) - c| < \varepsilon$. Similarly, by assumption, there exists ρ_0 such that for all $\rho > \rho_0$, $\sup_{0 \leq k \leq K(\rho)} \left| \frac{a_k(\rho)}{b_k(\rho)} - z \right| < \delta$. Combining the last two statements it follows that for all $\rho > \rho_0$, $\sup_{0 \leq k \leq K(\rho)} |G(a_k(\rho), b_k(\rho)) - c| < \varepsilon$. Since ε is arbitrary, the result follows.

Part F is a special case of Part G with, for example, $x = z_0$, $y = 1$, $\lim_{\rho \rightarrow 1}^* a_k(\rho) = z_0$ and $b_k(\rho) = 1$ for all ρ and for $0 \leq k \leq K(\rho)$.

Next, we show Part C. Let $\lim_{\rho \rightarrow 1}^* a_k(\rho) = a$ and $\lim_{\rho \rightarrow 1}^* b_k(\rho) = b$. Then

$$\begin{aligned} \lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} a_k(\rho)b_k(\rho) &\leq \lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} a_k(\rho) \lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} b_k(\rho) \\ &= ab, \end{aligned}$$

and

$$\begin{aligned} \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} a_k(\rho)b_k(\rho) &\geq \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} a_k(\rho) \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} b_k(\rho) \\ &= ab, \end{aligned}$$

where the equalities in the two equations above follow from the definition of \lim^* . Combining the two equations above and using the definition of \lim^* , we obtain that

$\lim_{\rho \rightarrow 1}^* a_k(\rho)b_k(\rho) = ab$, which is what was needed to be shown. Part B can be shown in a similar way.

We proceed with Part D. Since $\lim_{\rho \rightarrow 1}^* \frac{a_k(\rho)}{b_k(\rho)} = 1$, it follows that for any $\varepsilon > 0$, there exists ρ_o such that for all $\rho > \rho_o$

$$b_k(\rho)(1 - \varepsilon) < a_k(\rho) < b_k(\rho)(1 + \varepsilon)$$

for all $0 \leq k \leq K(\rho)$. Consequently, for all such ρ

$$1 - \varepsilon < \frac{\sum_{k=0}^{K(\rho)} a_k(\rho)}{\sum_{k=0}^{K(\rho)} b_k(\rho)} < 1 + \varepsilon.$$

Since ε is arbitrary, the result follows. Part E can be shown in a similar way.

Finally, we show Part A as follows:

$$\begin{aligned} & \lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} \int f(k, \rho, x) dx \\ & \leq \lim_{\rho \rightarrow 1} \int \sup_{0 \leq k \leq K(\rho)} f(k, \rho, x) dx \\ & \stackrel{(a)}{=} \int \lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} f(k, \rho, x) dx \\ & \stackrel{(b)}{=} \int \lim_{\rho \rightarrow 1}^* f(k, \rho, x) dx \end{aligned} \quad (D1)$$

where (a) follows from having

$$\left| \sup_{0 \leq k \leq K(\rho)} f(k, \rho, x) \right| \leq \sup_{0 \leq k \leq K(\rho)} |f(k, \rho, x)| \leq G(x) \text{ a.e.}$$

and having $\lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} f(k, \rho, x)$ exist a.e. (due to the fact that if $\lim_{\rho \rightarrow 1}^* f(k, \rho, x)$ exists, it equals $\lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} f(k, \rho, x)$), and applying the dominated convergence theorem⁶ [30 (p. 209)], and (b) is due to the just mentioned fact that if $\lim_{\rho \rightarrow 1}^* f(k, \rho, x)$ exists, it equals $\lim_{\rho \rightarrow 1} \sup_{0 \leq k \leq K(\rho)} f(k, \rho, x)$. In a similar way, we also have

$$\begin{aligned} & \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} \int f(k, \rho, x) dx \\ & \geq \lim_{\rho \rightarrow 1} \int \inf_{0 \leq k \leq K(\rho)} f(k, \rho, x) dx \\ & \stackrel{(a)}{=} \int \lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} f(k, \rho, x) dx \\ & \stackrel{(b)}{=} \int \lim_{\rho \rightarrow 1}^* f(k, \rho, x) dx \end{aligned} \quad (D2)$$

where (a) follows from having

$$\left| \inf_{0 \leq k \leq K(\rho)} f(k, \rho, x) \right| \leq \sup_{0 \leq k \leq K(\rho)} |f(k, \rho, x)| \leq G(x) \text{ a.e.}$$

⁶It is easily shown that the theorem applies when the integrand is parameterized by some t converging continuously to some t_o , rather than some integer n converging to ∞ .

and having $\lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} f(k, \rho, x)$ exist a.e. (due to the fact that if $\lim_{\rho \rightarrow 1}^* f(k, \rho, x)$ exists, it equals $\lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} f(k, \rho, x)$), and applying the dominated convergence theorem, and (b) is due to the just mentioned fact that if $\lim_{\rho \rightarrow 1}^* f(k, \rho, x)$ exists, it equals $\lim_{\rho \rightarrow 1} \inf_{0 \leq k \leq K(\rho)} f(k, \rho, x)$. Combining (D1) and (D2) concludes the proof of Part A and of the lemma. \square

Lemma D2: Let X be a Gaussian random variable with variance σ^2 and any mean, quantized by quantizer q with step size Δ . Then

$$\begin{aligned} H(q(X)) & < \log \frac{1}{\lambda} \\ & + 2N(\lambda) + \frac{\log e}{\sqrt{2\pi}} \left(6\lambda + 3\sqrt{2}\lambda^2 + \lambda^3 \right) + 6 \\ & \triangleq b(\lambda) \end{aligned}$$

where $N(\lambda) = 0$ if $\lambda \leq \frac{\sqrt{2\pi}}{e}$, or $N(\lambda) = \left\lceil \frac{\sqrt{2 - \ln \frac{2\pi}{\lambda^2}}}{\lambda} + 1 \right\rceil$ otherwise, and $\lambda = \frac{\sigma}{\Delta}$. It follows that if λ tends to zero, then

$$\limsup_{\lambda \rightarrow 0} \frac{H(q(X))}{\log \frac{1}{\lambda}} \leq 1.$$

Proof: Let the mean of X be contained in cell zero, and let f denote the pdf of X . We write f as follows:

$$f(x) = \begin{cases} f^L(x), & x < t_0 \\ f^C(x), & t_0 \leq x < t_1 \\ f^R(x), & x \geq t_1 \end{cases}$$

where $f^L(x)$, $f^C(x)$ and $f^R(x)$ are zero outside of their respective regions. We now have

$$\begin{aligned} H(q(X)) & = H_q(f) = H_q(f^L) + H_q(f^C) + H_q(f^R) \\ & \leq H_q(f^L) + H_q(f^R) + \mathcal{H}\left(\frac{1}{e}\right) \end{aligned} \quad (D3)$$

where the inequality follows by upper bounding $H_q(f^C)$ by $\mathcal{H}\left(\frac{1}{e}\right)$.

Next, we upper bound $H_q(f^R)$. For $k \geq 1$

$$\begin{aligned} P_k & = \int_{(k-\theta)\Delta}^{(k+1-\theta)\Delta} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx < \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{(k-1)^2\lambda^2}{2}} \\ & \triangleq P'_k. \end{aligned}$$

It is not hard to see that $P'_k < \frac{1}{e}$ for all $k > N(\lambda)$, where $N(\lambda)$ is given in the Lemma statement. Thus, using Fact C8, we have that $\mathcal{H}(P_k) < \mathcal{H}(P'_k)$, for all $k > N(\lambda)$. It now follows that

$$\begin{aligned} & H_q(f^R) \\ & = - \sum_{k=1}^{\infty} P_k \log P_k \\ & < \mathcal{H}\left(\frac{1}{e}\right)(N(\lambda) + 1) \\ & + \sum_{k=N(\lambda)+2}^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{(k-1)^2\lambda^2}{2}} \end{aligned}$$

$$\begin{aligned} & \times \left(\frac{(k-1)^2 \lambda^2}{2} \log e + \log \frac{\sqrt{2\pi}}{\lambda} \right) \\ & < (N(\lambda) + 1) + \log \frac{\sqrt{2\pi}}{\lambda} \sum_{k=1}^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{k^2 \lambda^2}{2}} \\ & + \log e \sum_{k=1}^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{k^2 \lambda^2}{2}} \frac{k^2 \lambda^2}{2} \end{aligned} \quad (D4)$$

where the last inequality follows since $\mathcal{H}(\frac{1}{e}) < 1$ and $N(\lambda) \geq 0$. We now consider the last two summations in the last equation. The first summation of the two is upper bounded as follows:

$$\begin{aligned} \log \frac{\sqrt{2\pi}}{\lambda} \sum_{k=1}^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{k^2 \lambda^2}{2}} & < \log \frac{\sqrt{2\pi}}{\lambda} \int_0^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{x^2 \lambda^2}{2}} dx \\ & = \frac{1}{2} \log \frac{\sqrt{2\pi}}{\lambda}. \end{aligned} \quad (D5)$$

Next, we upper bound the second summation as follows:

$$\begin{aligned} & \sum_{k=1}^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{k^2 \lambda^2}{2}} \frac{k^2 \lambda^2}{2} \\ & \stackrel{(a)}{<} \sum_{k=1}^{\lceil \frac{\sqrt{2}}{\lambda} \rceil} \frac{\lambda}{\sqrt{2\pi}} \frac{k^2 \lambda^2}{2} + \sum_{k=\lceil \frac{\sqrt{2}}{\lambda} \rceil + 1}^{\infty} \frac{\lambda}{\sqrt{2\pi}} e^{-\frac{k^2 \lambda^2}{2}} \frac{k^2 \lambda^2}{2} \\ & \stackrel{(b)}{<} \left(\frac{\sqrt{2}}{\lambda} + 1 \right) \frac{\lambda}{\sqrt{2\pi}} \frac{(\frac{\sqrt{2}}{\lambda} + 1)^2 \lambda^2}{2} \\ & + \frac{1}{2} \int_0^{\infty} \frac{\lambda}{\sqrt{2\pi}} x^2 \lambda^2 e^{-\frac{x^2 \lambda^2}{2}} dx \\ & \stackrel{(c)}{=} \frac{1}{\sqrt{\pi}} \left(1 + \frac{3\lambda}{\sqrt{2}} + \frac{3\lambda^2}{2} + \frac{\lambda^3}{2\sqrt{2}} \right) + \frac{1}{4} \end{aligned} \quad (D6)$$

where (a) uses the fact that $e^{-\frac{k^2 \lambda^2}{2}} < 1$, (b) follows since the integrand is monotonically decreasing for $x > \frac{\sqrt{2}}{\lambda}$, and by integrating over a larger interval than $(\lceil \frac{\sqrt{2}}{\lambda} \rceil, \infty)$, and (c) is derived using simple algebra. It follows from (D4), (D5), and (D6) that

$$\begin{aligned} H_q(f^R) & < \frac{1}{2} \log \frac{\sqrt{2\pi}}{\lambda} + (N(\lambda) + 1) \\ & + \left[\frac{1}{\sqrt{\pi}} \left(1 + \frac{3\lambda}{\sqrt{2}} + \frac{3\lambda^2}{2} + \frac{\lambda^3}{2\sqrt{2}} \right) + \frac{1}{4} \right] \log e. \end{aligned}$$

The lemma now follows from (D3), the last equation, and the fact that the above upper bound can be shown to hold for $H_q(f^L)$ as well. \square

Lemma D3: Let $\alpha \in \mathbb{R}$ be given. Then

$$\lim_{p \rightarrow 0} \frac{\mathcal{H}(1 - \alpha p)}{\mathcal{H}(p)} = 0.$$

This is similar to Lemma 2 of [25].

Proof: We need to show that $\lim_{p \rightarrow 0} \frac{-(1-\alpha p) \ln(1-\alpha p)}{-p \ln p} = 0$. The following string of equalities proves the lemma:

$$\begin{aligned} & \frac{-(1 - \alpha p) \ln(1 - \alpha p)}{-p \ln p} \\ & = \left[\frac{-(1 - \alpha p) \ln(1 - \alpha p)}{(1 - \alpha p) \alpha p} \right] \left[\frac{-p \ln p}{-p \ln p} \right] \\ & = \left[\frac{\ln(1 - \alpha p)}{-\alpha p} \right] \left[(1 - \alpha p) \alpha \frac{1}{-\ln p} \right] \rightarrow 0 \text{ as } p \rightarrow 0 \end{aligned}$$

where we used the well-known fact that $\lim_{x \rightarrow 0} \frac{\ln(1-x)}{-x} = 1$. \square

Lemma D4: Let $a(s)$ and $b(s)$ be positive functions on \mathbb{R} such that $\lim_{s \rightarrow s_0} \frac{a(s)}{b(s)} = 1$ and $\lim_{s \rightarrow s_0} b(s) = b_0 \neq 1$. Then

$$\lim_{s \rightarrow s_0} \frac{\mathcal{H}(a(s))}{\mathcal{H}(b(s))} = 1.$$

This is a slightly weaker version of Lemma 4 of [25].

Proof: To keep notation short, we omit the parameter s from $a(s)$ and $b(s)$. The following string of equalities proves the lemma:

$$\begin{aligned} \frac{\mathcal{H}(a)}{\mathcal{H}(b)} & = \frac{-a \log a}{-b \log b} = \frac{a \log \left[\frac{a}{b} \right]}{b \log b} = \frac{a}{b} \left[\frac{\log \frac{a}{b}}{\log b} + 1 \right] \\ & = \frac{a}{b} + \frac{a}{b} \frac{\log \frac{a}{b}}{\log b} \xrightarrow{s \rightarrow s_0} 1 + \frac{1 \log 1}{\log b_0} = 1. \end{aligned}$$

\square

REFERENCES

- [1] H. Gish and J. N. Pierce, "Asymptotically efficient quantization," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 5, pp. 676–683, Sep. 1968.
- [2] J. Ziv, "On universal quantization," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 344–347, May 1985.
- [3] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 10, pp. 2325–2383, Oct. 1998.
- [4] S. Shamai, "Information rates by oversampling the sign of a bandlimited process," *IEEE Trans. Inf. Theory*, vol. 40, no. 7, pp. 1230–1236, Jul. 1994.
- [5] Z. Cvetkovic and M. Vetterli, "Error-rate characteristics of oversampled analog-to-digital conversion," *IEEE Trans. Inf. Theory*, vol. 44, no. 9, pp. 1961–1964, Sep. 1998.
- [6] Z. Cvetkovic and M. Vetterli, "On simple oversampled A/D conversion in $L^2(\mathbb{R})$," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 146–154, Jan. 2001.
- [7] Z. Cvetkovic and I. Daubechies, "Single-bit oversampled A/D conversion with exponential accuracy in the bit-rate," in *Proc. DCC*, Snowbird, UT, Mar. 2000, pp. 343–352.
- [8] P. Ishwar, A. Kumar, and K. Ramchandran, "Distributed sampling for dense sensor networks: A bit-conservation principle," in *Proc. IPSN*, Palo Alto, CA, Apr. 2003, pp. 17–31.
- [9] I. Bar-David, "An implicit sampling theorem for bounded bandlimited functions," *Inf. Control*, vol. 24, pp. 36–44, 1974.
- [10] A. Kumar, P. Ishwar, and K. Ramchandran, "On distributed sampling of smooth non-bandlimited fields," in *Proc. IPSN*, Berkeley, CA, Apr. 2004, pp. 89–98.
- [11] N. T. Thao and M. Vetterli, "Reduction of the MSE in R-times oversampled A/D conversion $O(1/R)$ to $O(1/R^2)$," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 200–203, Jan. 1994.
- [12] N. T. Thao and M. Vetterli, "Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates," *IEEE Trans. Signal Process.*, vol. 42, no. 3, pp. 519–531, Mar. 1994.
- [13] Z. Cvetkovic and M. Vetterli, "Error-rate characteristics of oversampled analog-to-digital conversion," *IEEE Trans. Inf. Theory*, vol. 44, no. 9, pp. 1961–1964, Sep. 1998.
- [14] I. Bar-David, "Sample functions of a Gaussian random process cannot be reconstructed from their zero crossings," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 86–87, Jan. 1975.

- [15] D. Slepian, "Estimation of the Gauss-Markov process from observation of its sign," *Stoch. Proc. Appl.*, vol. 14, pp. 249–265, 1983.
- [16] D. Marco, "Markov random processes are neither bandlimited nor recoverable from samples or after quantization," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 900–905, Feb. 2009.
- [17] A. Scaglione and S. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks," in *Proc. ACM MobiCom*, Atlanta, GA, Sep. 2002, pp. 140–147.
- [18] D. Marco, E. J. Duarte-Melo, M. Liu, and D. L. Neuhoff, "On the many-to-one capacity of a dense wireless sensor network," in *Proc. IPSN*, Palo Alto, CA, Apr. 2003, pp. 1–16.
- [19] E. J. Duarte-Melo and M. Liu, "Data-gathering wireless sensor networks: Organization and capacity," *Wireless Sensor Netw.*, vol. 43, pp. 419–537, Nov. 2003.
- [20] D. Ganesan, R. Cristescu, and B. Beferull-Lozano, "Power-efficient sensor placement and transmission structure for data gathering under distortion constraints," in *Proc. ISPN*, Berkeley, CA, Apr. 2004, pp. 142–150.
- [21] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.
- [22] M. R. Leadbetter, G. Lindgren, and H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*. New York: Springer-Verlag, 1983.
- [23] D. Marco, "Asymptotic Quantization and Applications to Sensor Networks," Ph.D. dissertation, Univ. Michigan, Ann Arbor, 2004.
- [24] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1967.
- [25] D. Marco and D. L. Neuhoff, "Low resolution scalar quantization for Gaussian sources and squared error," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1689–1697, Apr. 2006.
- [26] A. Kashyap, L. A. Lastras-Montano, C. Xia, and Z. Liu, "Distributed source coding in dense sensor networks," in *Proc. DCC*, Snowbird, UT, Apr. 2005, pp. 13–22.
- [27] D. L. Neuhoff and S. Pradhan, "An upper bound to the rate of ideal distributed lossy source coding of densely sampled data," in *Proc. ICASSP*, Toulouse, May 2006, pp. V–1137–1140.
- [28] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*, 2nd ed. New York: Springer-Verlag, 1985.
- [29] J. Neveu, *Mathematical Foundations of the Calculus of Probability*. San Francisco, CA: Holden-Day, 1965.
- [30] P. Billingsley, *Probability and Measure*, 3rd ed. New York: Wiley, 1995.

Daniel Marco (S'02–M'04) received the B.Sc. degree with distinction in computer engineering from the Technion—Israel Institute of Technology, Haifa, in 1999, and the M.S. degree in electrical engineering, the M.S. degree in mathematics, and the Ph.D. degree in electrical engineering, all from the University of Michigan, Ann Arbor, in 2001, 2003, and 2004, respectively.

From 2004 to 2006, he was with the California Institute of Technology, Pasadena, as a postdoctoral scholar. His research interests lie primarily in information theory, high- and low-resolution quantization theory, rate-distortion theory, coding with side information, and sensor networks. Since 2008, he has been with Final Israel, Ltd.

David L. Neuhoff (F'94) received the B.S.E. degree from Cornell University, Ithaca, NY, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA.

Since graduation, he has been a faculty member at the University of Michigan, Ann Arbor, where he is now the Joseph E. and Anne P. Rowe Professor of Electrical Engineering. From 1984 to 1989, he was the Associate Chair of the Systems Science and Engineering Division of the EECS Department, and since September 2008, he has been the Associate Chair for the ECE Division of the Department. He spent two sabbaticals at Bell Laboratories, Murray Hill, NJ, and one at Northwestern University, Evanston, IL. His research and teaching interests are in communications, information theory, and signal processing, especially data compression, quantization, image coding, source-channel coding, half-toning, ad hoc network scaling, distributed coding for sensor networks, and Markov random fields.

Dr. Neuhoff was an Associate Editor for Source Coding for the IEEE TRANSACTIONS ON INFORMATION THEORY from 1986 to 1989. He served on the Board of Governors of the IEEE Information Theory Society from 1988 to 1990, as well as from 2002 to 2008. He was President of the Information Theory Society in 2006. Also, he cochaired the 1986 IEEE International Symposium on Information Theory in Ann Arbor.