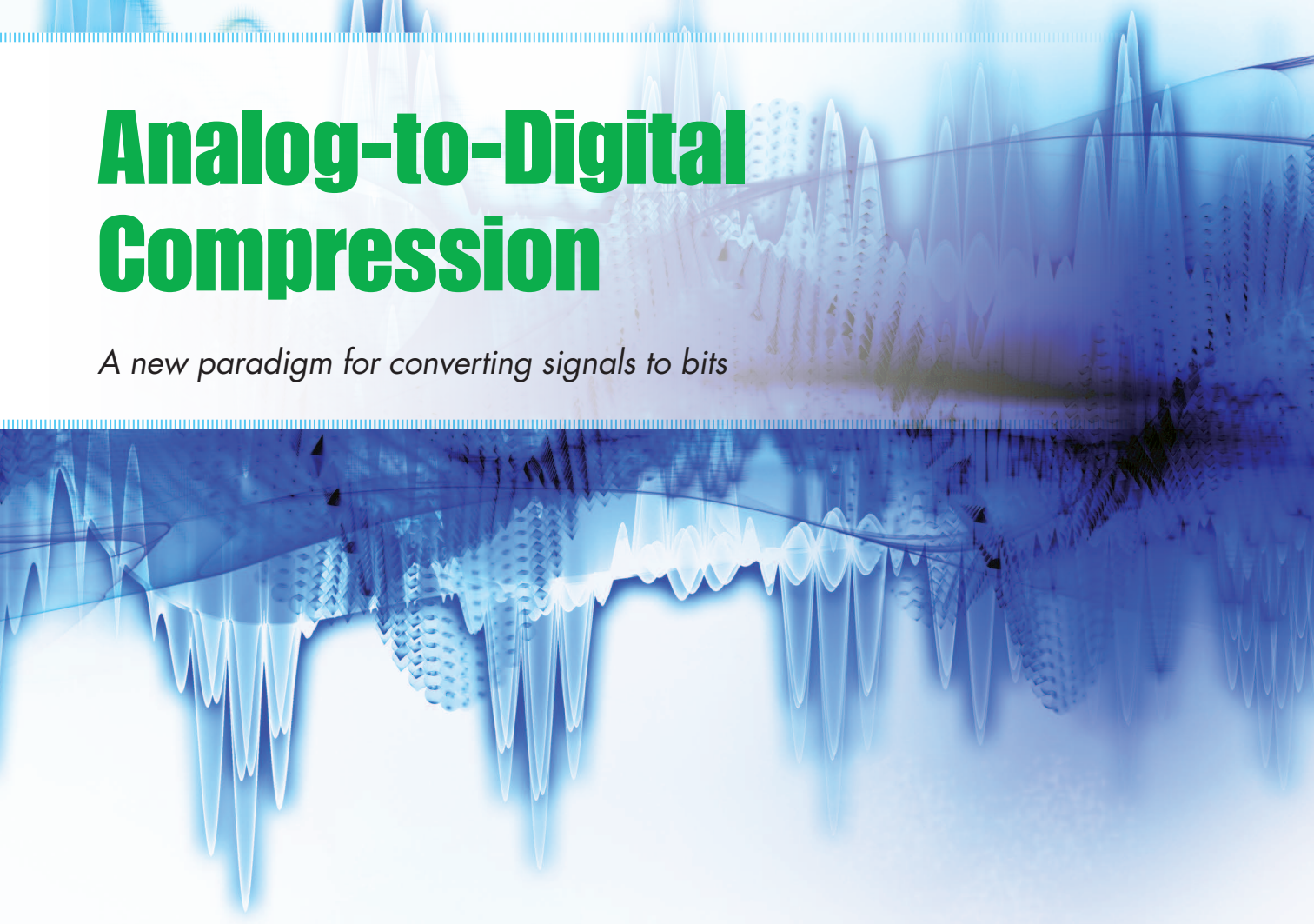
Alon Kipnis, Yonina C. Eldar, and Andrea J. Goldsmith

Analog-to-Digital Compression

A new paradigm for converting signals to bits



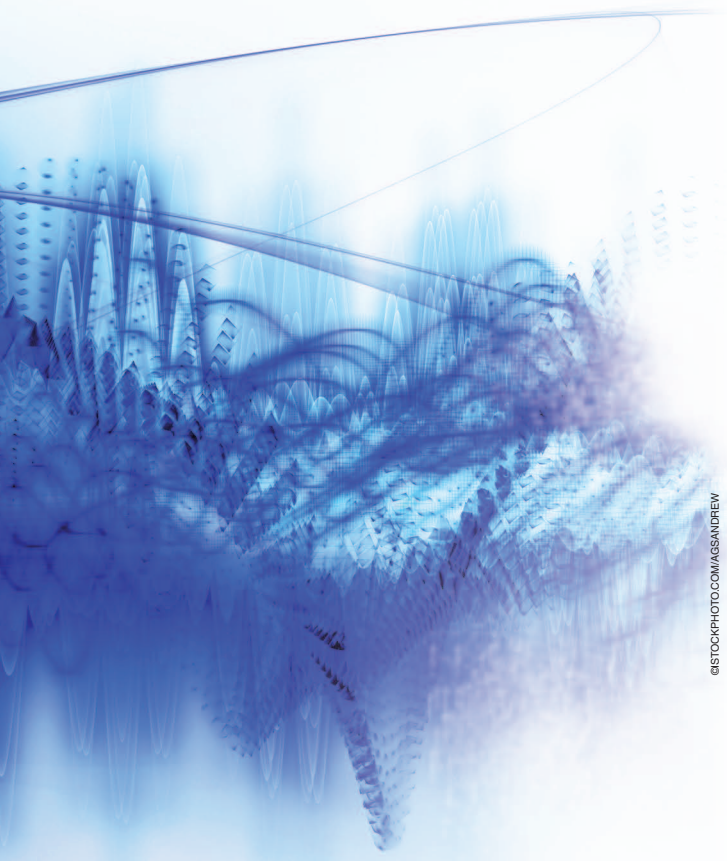
Processing, storing, and communicating information that originates as an analog signal involves converting this information to bits. This conversion can be described by the combined effect of sampling and quantization, as shown in Figure 1. The digital representation is achieved by first sampling the analog signal to represent it by a set of discrete-time samples and then quantizing these samples to a finite number of bits. Traditionally, these two operations are considered separately. The sampler is designed to minimize the information loss due to sampling based on characteristics of the continuous-time input. The quantizer is designed to represent the samples as accurately as possible, subject to a constraint on the number of bits that can be used in the representation. The goal of this article is to revisit this paradigm by illuminating the dependency between these two operations. In particular, we explore the requirements of the sampling system subject to the constraints on the available number of bits for storing, communicating, or processing the analog information.

Digital Object Identifier 10.1109/MSP.2017.2774249
Date of publication: 26 April 2018

Motivation

As a motivation for optimizing sampling and quantization together, consider the minimal sampling rate that arises in classical sampling theory due to Whittaker, Kotelnikov, Shannon, and Landau [1]–[3]. These works establish the Nyquist rate, or the spectral occupancy of the signal, as the critical sampling rate, above which the signal can be perfectly reconstructed from its samples. This statement, however,

 1053-5888/18©2018IEEE



©ISTOCKPHOTO.COM/AGSANDREW

focuses only on the critical sampling rate required to perfectly reconstruct a bandlimited signal from its discrete samples. It does not incorporate the quantization precision of the samples and does not apply to signals that are not bandlimited. It is, in fact, impossible to obtain an exact representation of any continuous-amplitude sequence of samples by a digital sequence of numbers because of finite quantization precision, and, therefore, any digital representation of an analog signal is prone to error. That is, no continuous-amplitude signal can be reconstructed from its quantized samples with zero distortion regardless of the sampling rate, even when the signal is bandlimited.

This limitation raises the following question: In converting a signal to bits via sampling and quantization at a given bit precision, can the signal be reconstructed from these samples with minimal distortion based on sub-Nyquist sampling? In this article, we discuss this question by extending classical sampling theory to account for quantization and for nonbandlimited inputs. That is, for an arbitrary stochastic input and given a total budget of quantization bits, we consider the lowest sampling rate required to sample the signal such that reconstruction of the signal from its quantized samples results in minimal distortion. Without assuming any particular structure of the input analog signal, this sampling rate is often below the signal's Nyquist rate.

The minimal distortion achievable in the presence of quantization depends on the particular way the signal is quantized or, more generally, encoded into a sequence of bits. Since we are interested in the fundamental distortion limit in recovering an analog signal from its digital representation, we consider all possible encoding and reconstruction (decoding) techniques. As an example, in Figure 1, the smartphone display may be viewed as a reconstruction of the real-world painting *The Starry Night* from its digital representation. No matter how excellent the quality of a smartphone's high-definition screen may be, this recovery is not perfect, since the digital representation of the analog image is not accurate due to a loss of information occurring during the conversion from analog to bits. Our goal is to analyze this loss as a function of hardware limitations on the sampling mechanism and the number of bits used in the encoding. It is convenient to normalize this number of bits by the signal's free dimensions, i.e., the dimensions along which new information is generated. For example, the free dimensions of a visual signal are usually the horizontal and vertical axes of the frame, and the free dimension of an audio wave is time. For simplicity, we consider analog signals with a single free dimension, i.e., time. Therefore, our restriction on the digital representation is given in terms of its bit rate—the number of bits per unit time.

For an arbitrary continuous-time random signal with known statistics, the fundamental distortion limit due to the encoding of the signal using a limited bit rate is given by Shannon's distortion-rate function (DRF) [4]–[6]. This function provides the optimal tradeoff between the bit rate of the signal's digital representation and the distortion in recovering the original signal from this representation. Shannon's DRF is described only in terms of the distortion criterion, the probability distribution on the continuous-time signal, and the maximal bit rate allowed in the digital representation. Consequently, the optimal encoding scheme that attains Shannon's DRF is a general mapping from continuous-time signal space to bits that does not consider practical constraints in its implementation. In practice, the encoding of an analog signal into bits entails first sampling the signal and then representing the samples using a limited number of bits. Therefore, in practice, the minimal distortion in recovering analog signals from their bit representation considers the digital encoding of the signal samples, with a constraint on both the sampling rate and the bit rate of the system. Here, the sampling rate f_s is defined as the number of samples per

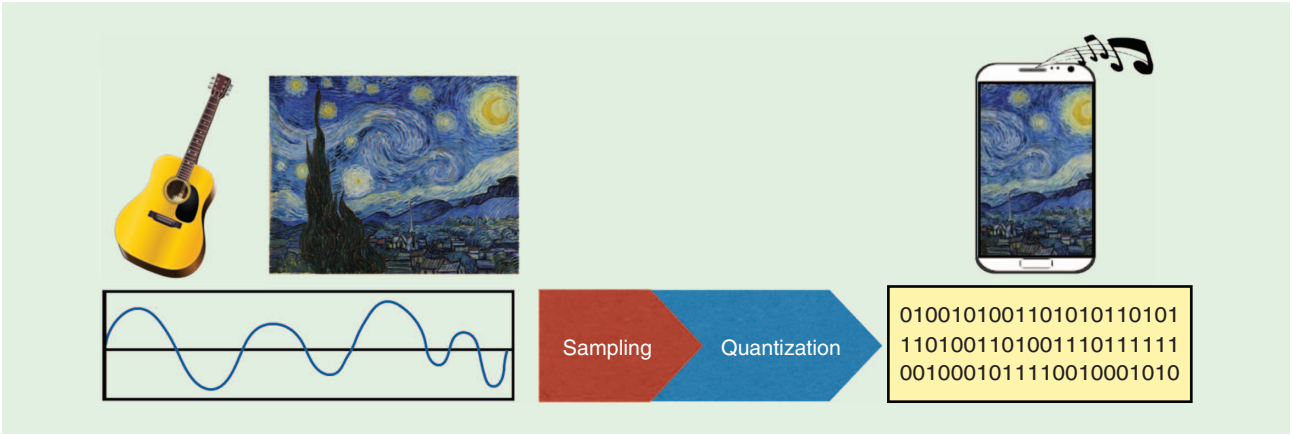


FIGURE 1. An illustration showing that analog-to-digital conversion (ADC) is achieved by combining sampling and quantization. (Image of the guitar courtesy of https://pixabay.com; image of Van Gogh's *The Starry Night* courtesy of Wikipedia.)

unit time of the continuous-time source signal, and the bit rate R is the number of bits per unit time used in the representation of these samples. The resulting system describing our problem is shown in Figure 2 and is referred to as the *analog-to-digital compression* (*ADX*) setting.

The digital representation in this setting is obtained by transforming a continuous-time, continuous-amplitude, random source signal $X(t)$ through a concatenated operation of a sampler and an encoder, resulting in a bit sequence. For instance, when the input signal $X(t)$ is observed over a time interval T, then the sampler produces $\lfloor f_s T \rfloor$ samples, and the encoder maps these samples to $\lfloor TR \rfloor$ bits. The decoder estimates the original analog signal from this bit sequence. The distortion is defined to be the mean squared error (MSE) between the input signal $X(t)$ and its reconstruction $\hat{X}(t)$. Since we are interested in the fundamental distortion limit subject to a sampling constraint, we allow optimization over the encoder, decoder, and time horizon T. In addition, we also explore the optimal sampling mechanism but limit ourselves to the class of linear and continuous deterministic samplers [7]. That is, each sampler in this class is a linear continuous mapping of signals over time lag T to $\mathbb{R}^{\lfloor f_s T \rfloor}$.

The minimal distortion in ADX is bounded from below by two extreme cases of the sampling rate and the bit rate, as shown in Figure 3:

1) When the bit rate R is unlimited, the minimal ADX distortion reduces to the MSE in interpolating a signal from its samples at rate f_s.
2) When the sampling rate f_s is unlimited or above the Nyquist rate of the signal, the ADX distortion reduces to Shannon's DRF of the signal.

Indeed, in this situation, the optimal encoder can recover the original continuous-time signal without distortion and then encode this recovery in an optimal manner according to the scheme that attains Shannon's DRF. Our goal is therefore to characterize the MSE due to the joint effect of a finite bit-rate constraint and sampling at a sub-Nyquist sampling rate or for signals that are not bandlimited. In particular, we are interested in the minimal sampling rate for which Shannon's DRF, describing the minimal distortion subject to a bit-rate constraint, is attained. As shown in Figure 3 and as will be explained in more detail in this article, this sampling rate is usually below the Nyquist rate of the signal. We denote this minimal sampling rate as the *critical sampling rate* subject to a

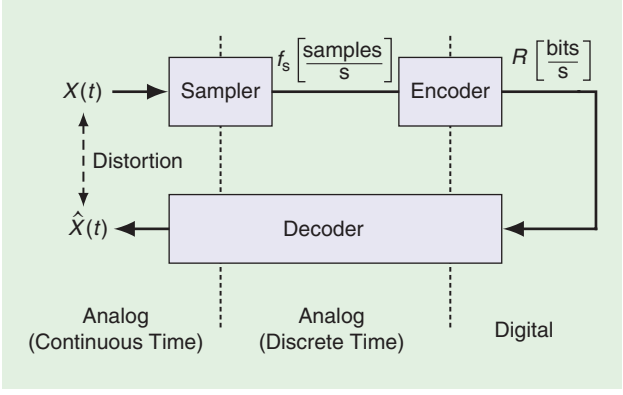


FIGURE 2. The ADX and reconstruction setting. Our goal is to derive the minimal distortion between the signal and its reconstruction from any encoding, at bit rate R of the samples of the signal taken at sampling rate f_s.

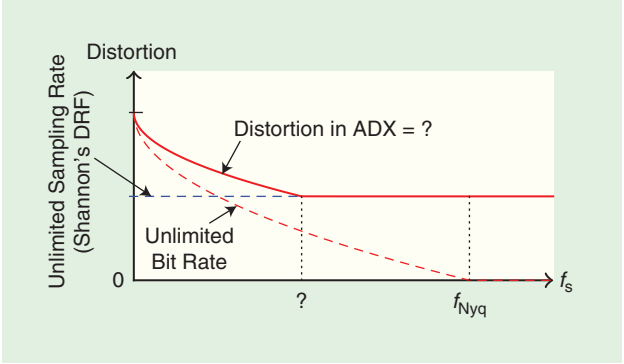


FIGURE 3. The minimal sampling rate for attaining the minimal distortion achievable in the presence of quantization is usually below the Nyquist rate, whereas sampling at the Nyquist rate is necessary to attain zero distortion without quantization constraints.

bit-rate constraint, since it describes the minimal sampling rate required to attain the optimal performance in systems operating under quantization or bit-rate restrictions. Therefore, the critical sampling rate extends the minimal-distortion sampling rate considered by Shannon, Nyquist, and Landau. It is only as the bit rate extends to infinity that sampling at the Nyquist rate is necessary to attain minimal (i.e., zero) distortion for general input distributions.

Figure 2 represents a general block diagram for systems that process information through sampling and are limited in the number of bits they can transmit per unit time, the amount of memory they use, or the number of states they can assume. Therefore, the critical sampling rate that arises in this setting describes the fundamental limit of sampling in systems like audio and video recorders, radio receivers, and digital cameras. Moreover, this model also includes signal processing techniques that use sampling and operate under bit-rate constraints, such as artificial neural networks [8], financial markets analyzers [9], and techniques to accelerate operations over large data sets by sampling [10]. In "System Constraints on Bit Rate," we list a few scenarios where sampling and bit-rate restrictions arise in practice. Other utilizations of the ADX paradigm will be discussed in the "Applications" section.

To derive the critical sampling rate, we rely on the following two steps:

1) Given the output of the sampler, derive the optimal way to encode these samples subject to the bit rate R so as to minimize the MSE distortion in reconstructing the original continuous-time signal.
2) Derive the optimal sampling scheme that minimizes the MSE in the first step subject to the sampling rate constraint.

When the analog signal can be perfectly recovered from the output of the sampler, the fundamental distortion limit in step 1 depends only on the bit-rate constraint and leads to Shannon's DRF. In this article, we explore this function as well as the optimal encoding to attain it. Applications of the ADX framework and the critical sampling rate that attains the minimal distortion are also discussed.

Before exploring the minimal distortion limit in the ADX setting, it is instructive to consider the distortion in pulse-code modulation, which is a particular system that is implementing a simple version of a sampler, an encoder, and a decoder. Although this system does not implement the optimal sampling

System Constraints on Bit Rate

The analog-to-digital compression setting of Figure 2 is relevant to any system that processes information by sampling and is subject to a bit-rate constraint. Three possible restrictions on a system's bit rate that arise in practice are as follows:

- *Memory*: Digital systems often operate under a constraint on the amount of memory or the states they can assume. Under such a restriction, the bit rate is the normalized amount of memory used over time (or the dimension of the source signal). For example, consider a system of K states that analyzes information obtained by observing an analog signal for T seconds. The maximal bit rate of the system is $R = \log_2(K)/T$.
- *Power*: Emerging sensor network technologies, such as those developed for biomedical applications and smart cities, use many low-cost sensors to collect and transmit data to remote locations [11]. These sensors must operate under severe power restrictions and, hence, are limited by the number of comparisons in their analog-to-digital conversion (ADC) operation. These comparisons are typically the most energy-consuming part of the ADC unit, so that the total power consumption in an ADC unit is proportional to the number of comparisons [12, Sec. 2.1]. In general, the number of comparisons is proportional to the bit rate, since any output of bit rate R is generated by at least R comparisons (although the exact number depends on the particular implementation of the ADC and may even grow exponentially in the bit rate [13]). Therefore, power restrictions lead to a bit-rate constraint and to an MSE distortion floor given by Shannon's distortion-rate function of the analog input signal.

  An important scenario of power-restricted ADC units arises in wireless communication using millimeter waves [14]. Severe path loss of electromagnetic waves in these frequencies is compensated for by using a large number of receiver antennas. Each antenna is associated with a radio-frequency chain that includes an ADC unit. Because of the resulting large number of ADC units, power consumption is one of the major engineering challenges in millimeter-wave communication.
- *Communication*: Low-power sensors may also be limited by the rates of communication available to send their digital sensed information to a remote location. For example, consider a low-energy device collecting medical signals and transmitting its measurements wirelessly to a central processor (e.g., a smartphone). The communication rate from the sensor to the central processor depends on the capacity of the channel between them, which is a function of the available transmit power for communication. When the transmit power is limited, so is the capacity. As a result, the data rate associated with the digital representation of the sensed information cannot exceed this capacity limit, since, without additional processing, there is no point in collecting more information than what can be communicated.

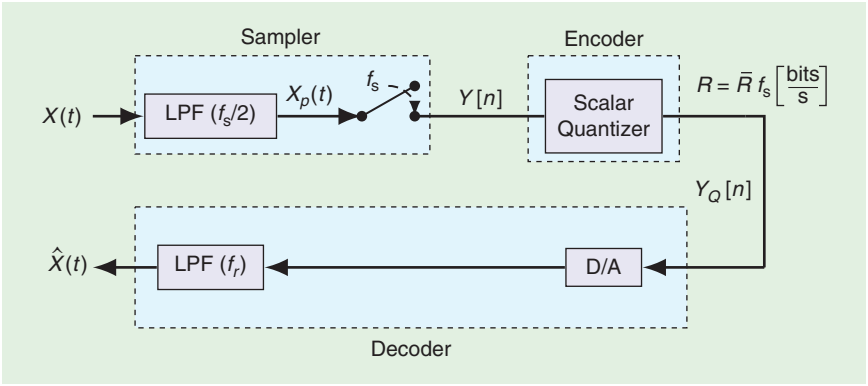


FIGURE 4. PCM and reconstruction system.

and encoding scheme, it illustrates an instance where, as a result of the bit-rate constraint, sampling below the Nyquist rate is optimal. In addition, this analysis provides a simple way to introduce the notions of sampling, quantization, and bit rate and serves as a basis for the generalization of the sampling and for encoding operations into optimal ones.

ADX via pulse-code modulation

A particular example for a system incorporating a sampler, an encoder, and a decoder is given in Figure 4. This system converts the analog signal $X(t)$ to a digital representation $Y_Q[n]$ by a uniform sampler followed by a scalar quantizer. This conversion technique is known as *pulse-code modulation* (*PCM*) [15], [16]; refer to [17, Sec. I.A] for its historical overview. The bit rate in this system is defined as the average number of bits per unit time required to represent the process $Y_Q[n]$. The goal of our analysis is to derive the MSE distortion in recovering the analog input signal $X(t)$ under a constraint R on this bit rate, assuming a particular sampling rate f_s of the sampler. We denote this distortion by $D_{\mathrm{PCM}}(f_s, R)$. Since the system in Figure 4 is a special case of Figure 2, the function $D_{\mathrm{PCM}}(f_s, R)$ is lower-bounded by the minimal distortion in the ADX, obtained by optimizing over all of the encoders and decoders, subject only to a sampling rate constraint f_s and a bit-rate constraint R.

We analyze the system of Figure 4 assuming a stochastic continuous-time, continuous-amplitude source signal $X(t)$ at its input. This signal is first filtered using a presampling low-pass filter (LPF) to yield $X_{\mathrm{p}}(t)$. The filtered signal is then sampled uniformly at rate f_s samples/s. Each sample $Y[n]$ is mapped using a scalar quantizer to $Y_Q[n]$, which is the nearest value to $Y[n]$ among a prescribed set of K quantization levels. More details on the operation of the scalar quantizer are provided in "Scalar Quantization." Since each of the quantization levels can be assigned a finite digital number, we say that the process $Y_Q[n]$ is a digital representation of $Y[n]$. As explained in "Scalar Quantization," the selection of the quantization levels and the length of the digital number assigned to each of them may also be subject to optimization. Subsequently, we assume that $\bar{R}$ is the expected number of bits per sample assigned to represent the quantization levels (the expectation is with respect to the distribution of the source signal). Using this notation, the bit rate of the digital representation, i.e., the number of bits per unit time required to represent the process $Y_Q[n]$, is defined as $R = \bar{R} f_s$.

The process of recovering the analog source signal $X(t)$ from the digital sequence $Y_Q[n]$ is described at the bottom of Figure 4. The digital discrete-time sequence of quantized values $Y_Q[n]$ is first converted to a continuous-time impulse train using a digital-to-analog (D/A) unit and then filtered using an ideal LPF with cutoff frequency f_r. In the time domain, this LPF is equivalent to an ideal sinc interpolation between the analog sample values to create a continuous-time signal bandlimited to $(-f_r, f_r)$. The result of this interpolation is denoted by $\hat{X}(t)$. We measure the distortion of the system by the MSE between $X(t)$ and $\hat{X}(t)$ averaged over time as

$$D_{\mathrm{PCM}}(f_s, R) \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \mathbb{E}\left(X(t) - \hat{X}(t)\right)^2 dt. \tag{1}$$

Note that letting the time grow symmetrically in both directions simplifies some of the expressions, but our results remain valid even if time grows in one direction. It is, in general, possible to use a different decoding scheme that would lead to a lower MSE under the same sampling and bit-rate constraint. Indeed, (1) is minimized by using the conditional expectation of $X(t)$ given $Y_Q[n]$ as the reconstruction signal rather than using $\hat{X}(t)$. However, the nonlinearity introduced by the scalar quantizer makes the exact analysis of the distortion under the conditional expectation a difficult task [17], and, therefore, for simplicity, we focus here on interpolation by low-pass filtering.

We now turn to analyze the distortion in (1) as a function of the sampling rate f_s and the bit rate R. We assume that $X(t)$ is a stationary stochastic process with a symmetric power spectral density (PSD) $S_X(f)$, and we denote its bandwidth by $f_{\mathrm{Nyq}}/2$. If $X(t)$ is not bandlimited, then we use the notation $f_{\mathrm{Nyq}} = \infty$. In either case, we assume that $X(t)$ is bounded in energy and denote

$$\sigma^2 = \mathrm{var} X(t) = \int_{\mathbb{R}} S_X(f)\, df.$$

We further assume that the PSD $S_X(f)$ is unimodal, in the sense that its energy distribution is decreasing as one moves away from the origin, as given, for example, in Figure 5. Under this assumption, the presampling filter that minimizes the distortion, among all linear time-invariant filters, is an LPF with a cutoff frequency of $f_s/2$ [18]. Henceforth, we assume that this filter is used. Finally, we pick the cutoff frequency f_r of the reconstruction filter to match the bandwidth of the low-pass filtered signal. This cutoff frequency is therefore the minimum between $f_s/2$ and the bandwidth of $X(t)$, which equals $f_{\mathrm{Nyq}}/2$.

As a result of these assumptions, the only distortion introduced in the sampling process is due to the presampling filter,

Scalar Quantization

Consider the problem of representing a random number X drawn from a continuous distribution using another number taken from a finite alphabet of K elements $X_Q \in \{x_1, \ldots, x_K\}$. Since an exact representation of X cannot be attained due to cardinality limitations, the goal is to minimize

$$\mathbb{E}(X - X_Q)^2. \tag{S1}$$

The mapping of X to X_Q is called *quantization*. When the representation of a sequence of random numbers is considered, we use the term *scalar quantization* to denote that the same quantization mapping is applied to each element of the sequence, independently of the previous elements.

Assuming that the quantizer inputs are independent, the estimation of each input sample from the output of the quantizer is based on only one of these K states $\{x_1, \ldots, x_K\}$. Evidently, the minimal estimation error is attained by mapping X to the reconstruction value x_i that minimizes (S1). As a result, the procedure of optimizing a scalar quantizer of K states can be described by selecting the optimal K reconstruction values. Given the distribution of the input, this optimal set may be attained by an iterative procedure known as the *Lloyd algorithm* or, more commonly, the *K-means algorithm* [30], [31].

The number of bits or the bit resolution of the quantizer is the number of binary digits that represent X at its output by assigning a different label to each state. Clearly, the output of a K-state quantizer can be encoded with $\lceil \log_2 K \rceil$ binary digits. However, this number may be reduced on average if the labels of the states consist of binary numbers of different length. For example, by using uniform quantization levels to quantize a nonuniformly distributed input, we may label those states that are more likely with binary numbers shorter than those numbers assigned to less likely states. These numbers must satisfy the condition that no member is a prefix of another member, so that the sequence of states can be uniquely decoded. This procedure is denoted as *variable-length scalar quantization*, distinguished from fixed-length quantization, in which the labels are all binary numbers of the same length.

Interestingly, the average mean squared error (MSE) over an independent and identically distributed sequence using a variable-length scalar quantizer may be strictly smaller than with a fixed-length scalar quantizer for the same average number of bits, even if the levels in the latter were optimized for the input distribution using the Lloyd algorithm. For example, with input taken from a standard normal distribution, the average MSE attained by a variable-length scalar quantizer with equally spaced reconstruction levels and an optimal labeling of these levels converges to $(\pi e/6)2^{-2\bar{R}} \approx 1.42 \times 2^{-2\bar{R}}$ as $\bar{R}$ becomes large [32]. In fact, it is also shown in [32] that a uniform quantizer with optimal labeling converges to the optimal variable-length quantizer as $\bar{R}$ increases. However, the distortion attained by a fixed-length quantizer under an optimal selection of the $K = 2^{\bar{R}}$ reconstruction levels converges to $(\sqrt{3}\,\pi/2)2^{-2\bar{R}} \approx 2.72 \times 2^{-2\bar{R}}$ [30].

As explained in "Source Coding and the DRF," a lower MSE for the same average number of bits per source sample $\bar{R}$ can be attained by using a vector quantizer, i.e., by considering the joint encoding of multiple samples from a sequence of samples, rather than one sample at a time.

and only in the case where f_s is smaller than the Nyquist rate of $X(t)$. In fact, this distortion is exactly the energy in the part of the spectrum of $X(t)$ blocked by the presampling filter. We therefore have

$$D_{\text{smp}}(f_s) \triangleq \sigma^2 - \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} S_X(f)\,df.$$

Note that $D_{\text{smp}}(f_s)$ equals zero when f_s is above the Nyquist rate of $X(t)$.

To analyze the distortion due to quantization, we represent the output of the quantizer as

$$Y_Q[n] = Y[n] + \eta[n], \quad n = 0, 1, \ldots, \tag{2}$$

where $\eta[n] = Y_Q[n] - Y[n]$ is the quantization noise. Since there is no aliasing in the sampling operation, the reconstruction filter applied to $Y[n]$ leads to the signal $X_p(t)$ at the output of the first LPF. Since the quantizer is a deterministic function of $Y[n]$, the process $\eta[n]$ is stationary, and we denote its PSD by $S_\eta(f)$ [note that $S_\eta(f)$ is periodic, with a period of f_s]. Nevertheless, an exact description of the statistics of $\eta[n]$ turns out to be a surprisingly difficult task. As a result, many approximations of its statistics have been developed [17], [19]. Most of these approximations provide conditions under which the spectrum of $\eta[n]$ is white (i.e., different elements of $\eta[n]$ are uncorrelated) [20]. One of the widely used approximations was provided by Bennet [21], who showed that, when the distribution of the input to the quantizer $Y[n]$ is continuous and the quantization levels are uniformly distributed, the spectrum of the quantization noise $S_\eta(f)$ converges to a constant as the quantizer resolution $\bar{R}$ increases. Another way to achieve the uniform spectral distribution of $\eta[n]$ is by dithering the signal at the input to the quantizer, i.e., by adding a pseudorandom noise signal [22]. For simplicity, our following analysis assumes that $S_\eta(f)$ is a

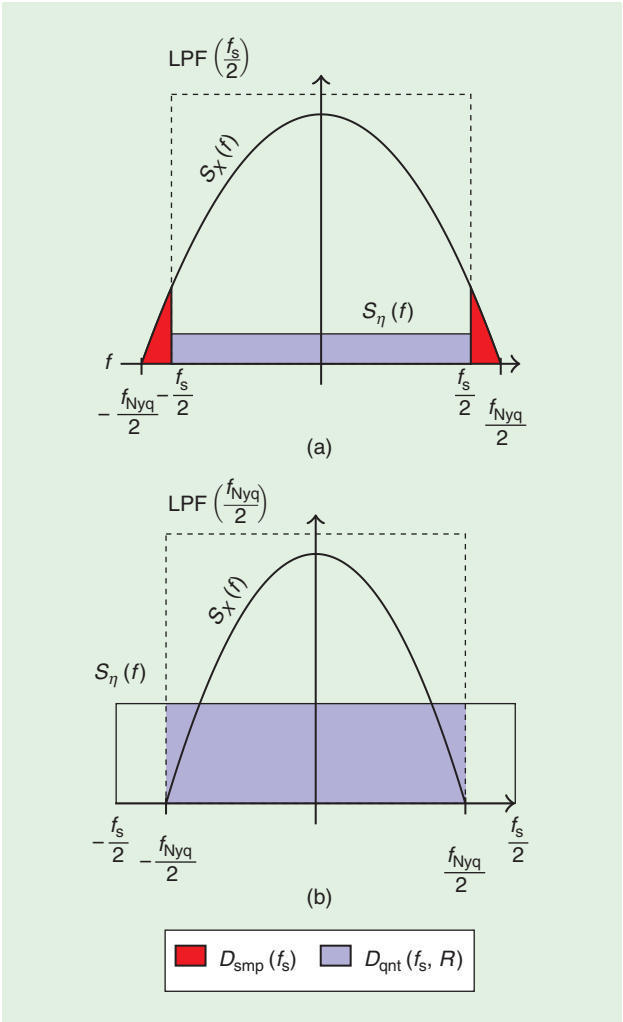


FIGURE 5. A spectral representation of the distortion in PCM (1). (a) Sampling below the Nyquist rate ($f_s > f_{Nyq}$) introduces sampling distortion $D_{smp}(f_s, R)$. (b) Sampling distortion vanishes when sampling above the Nyquist rate ($f_s > f_{Nyq}$), but the contribution of the in-band quantization noise $D_{qnt}(f_s, R)$ increases because of the lower bit precision of each sample.

constant, although deviation from this rule would not affect our general conclusions. Regardless of this assumption and as explained in "Scalar Quantization," the variance of this noise $\eta[n]$ is proportional to the variance of the process $Y[n]$ at the input to the quantizer and decreases exponentially with the number of quantization bits $\bar{R}$:

$$\text{var}(\eta[n]) = c_Q \text{var}(Y[n]) 2^{-2\bar{R}}. \tag{3}$$

The proportionality constant c_Q depends on the actual digital label assigned to each quantization level. At high quantization precision $\bar{R} = R/f_s$ and using a uniform quantizer, the value of the constant corresponding to optimal encoding converges to $c_Q = (\pi e/6)$. This value of c_Q is used in our figures.

Under the assumption that the PSD of $\eta[n]$ is constant over the entire discrete-time frequency range with variance (3) and using the fact that the variance of $Y[n]$ equals the variance of the low-pass filtered version of $X(t)$, the contribution of the quantization to the distortion in (1) is given by

$$\begin{aligned} D_{qnt}(f_s, R) &\triangleq \int_{-f_r}^{f_r} S_\eta(f)\, df \\ &= c_Q \left(\frac{\min\{f_s, f_{Nyq}\}}{f_s} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} S_X(f)\, df \right) 2^{-2R/f_s}, \end{aligned} \tag{4}$$

where the term in the braces represents the variance of $Y[n]$ or the energy of the signal at the output of the reconstruction LPF (the min is present because the LPF at the sampler is in use only if the sampling rate is lower than Nyquist). The overall distortion in PCM is, therefore,

$$D_{PCM}(f_s, R) = D_{smp}(f_s) + D_{qnt}(f_s, R). \tag{5}$$

The important observation from this expression is that, under a fixed bit rate R, the distortion due to quantization increases as the sampling rate f_s increases. This increase in f_s means fewer quantization bits are available to represent each sample, and, therefore, the distortion due to quantization is larger. Alternatively, the distortion due to sampling decreases as f_s increases and, in fact, vanishes as f_s exceeds the Nyquist rate. A spectral interpretation of the function $D_{PCM}(f_s, R)$ is shown in Figure 5. This figure shows the spectrum of the sampled source signal and the spectrum of the quantization noise under the high-resolution approximation for two representative cases of the sampling frequency:

1) *Sub-Nyquist sampling*: The distortion due to sampling $D_{smp}(f_s)$ is the part of $S_X(f)$ not included in the sampling interval $(-f_s/2, f_s/2)$. The distortion due to quantization is relatively low, since the small value of f_s allows the quantization of each sample with the relatively high resolution of $\bar{R} = R/f_s$ bits.
2) *Super-Nyquist sampling*: The distortion due to sampling $D_{smp}(f_s)$ is zero, but the distortion due to quantization $D_{qnt}(f_s)$ is affected by the reduction in the bit-resolution that decreases linearly in f_s, since $\bar{R} = R/f_s$.

It follows from the prior description that there exists a sampling rate that balances the two error contributions from quantization and sampling to minimize the total distortion in (5). This sampling rate can be seen in Figure 6, where the distortion $D_{PCM}(f_s, R)$ is shown versus the relative sampling rate f_s/f_{Nyq} for two PSDs. For the PSD $S_\Pi(f)$ with uniform energy distribution, the sampling rate that minimizes the distortion is exactly the Nyquist rate. For the triangular PSD $S_\Lambda(f)$, the optimal sampling rate is below the Nyquist rate. In general, it is shown in [18] that, under similar assumptions, the sampling rate that minimizes the distortion in PCM is always at or below the Nyquist rate. This rate is, in fact, strictly smaller than the Nyquist rate when the energy of the signal is not uniformly distributed over its spectral support, as in $S_\Lambda(f)$ of Figure 6. Going back to our general question, PCM illustrates an instance where, as a result of a bit-rate constraint, sampling below the Nyquist rate is optimal.

Another conclusion from our analysis is that, under a fixed bit rate, the distortion in PCM increases as a result of oversampling. This phenomenon is explained by the increasing correlation between consecutive time samples at a super-Nyquist sampling

rate, since the covariance function of a bandlimited signal is continuous [23], [24]. This correlation is not exploited by the quantizer, which maps two similar samples to the same digital value, leading to a redundant digital representation of the analog signal. Since the overall bit rate is limited, this redundancy in representation is translated to a higher distortion compared to the distortion in a less-redundant representation obtained at a lower sampling rate. In fact, it is well known that the sampling rate that minimizes the distortion in PCM also maximizes the entropy rate of the process postquantization, i.e., of $Y_Q[n]$ [17]. Therefore, we conclude that the most efficient representation of the analog signal in PCM under a bit-rate constraint is attained by sampling at or below the Nyquist rate.

The previously discussed conclusions imply that we can readily improve the performance of PCM by providing a more compact representation of the signal in terms of bit rate under the same distortion level, and we can do so in one of the following ways:

1) reduce the correlation between consecutive quantizer outputs by using a whitening transformation as in transform coding [17] or by a delta feedback loop as in sigma-delta modulation [25], [26]
2) compress the digital process $Y_Q[n]$ using a universal lossless compressor, such as Lempel–Ziv [27], [28] or context-tree weighting [29]
3) aggregate a large block of, e.g., N samples of $Y[n]$ and represent these samples using a single index out of $2^{\bar{R}N}$ possible values.

This last technique, commonly known as *vector quantization* [17], does not assume any restrictions on the mapping from the samples to the digital representation, except the size of the block. It therefore covers a wide range of quantization techniques operating at bit rate R and includes 1) and 2) as special cases. This technique leads to the most general way to encode any discrete-time process to a digital representation, subject only to a bit-rate constraint. Moreover, combined with an optimal mechanism to represent the analog signal as a bit sequence, this encoding technique attains the minimal distortion in encoding $X(t)$, described by Shannon's DRF $D(R)$.

Minimal distortion subject to a bit-rate constraint

We now go back to the ADX setting of Figure 2. In this section, we consider the minimal distortion that can be attained when the only restriction is the bit rate R of the resulting digital representation. In other words, we consider the minimal distortion assuming that the encoder operates directly on the continuous-time process $X(t)$, as shown in Figure 7.

This encoder observes a realization $x(t)$ of the process $X(t)$ over some finite time horizon T and then represents its observation using $\lfloor TR \rfloor$ bits. The number of possible states this encoding can take is, therefore, $2^{\lfloor TR \rfloor}$. As shown in Figure 8, without

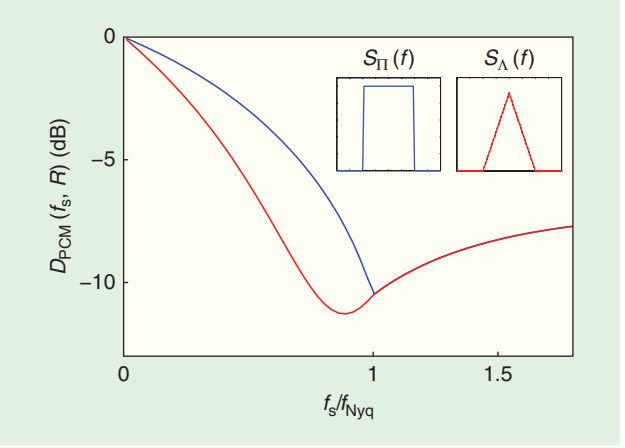


FIGURE 6. The distortion in PCM as a function of the sampling rate f_s for a fixed bit rate R and the PSDs in the small frames. With a nonuniform energy distribution, the optimal sampling rate of PCM is below the Nyquist rate.

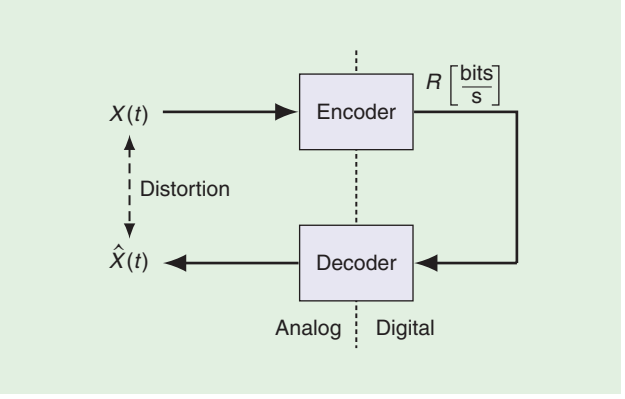


FIGURE 7. Encoding with full continuous-time source signal information.

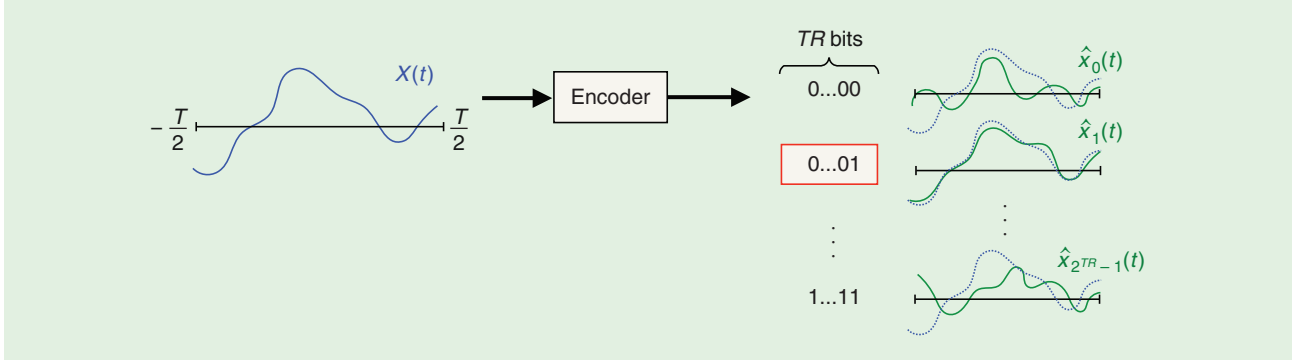


FIGURE 8. The optimal encoding with TR bits is obtained by mapping the source signal realization to the index of the predetermined reconstruction waveform closest to this realization. The optimal set of reconstruction waveforms and the resulting average distortion are given by Shannon's source coding theorem.

Source Coding and the DRF

The source coding problem addresses the encoding of a random source sequence so as to attain the minimal distortion over all possible encoding and reconstruction schemes, under a constraint on the average bits per source symbol in this encoding. In "Scalar Quantization," we considered the encoding of such sequences subject to the additional restriction that each source symbol is encoded independently of the other. By removing this restriction and considering the joint encoding of n independent source symbols, we can attain smaller distortion using the same average number of bits. For this reason, the source coding problem with respect to a real independent and identically distributed (i.i.d.) sequence $X_1, \ldots, X_n$ is defined as determining the minimum mean squared error attainable under all possible encoder mappings of a realization of this sequence to an index out of $2^{\lfloor n\bar{R} \rfloor}$ possible indices, as well as all reconstruction decoder mappings from this set of indices back to $\mathbb{R}^n$. This minimal value is called the *operational distortion-rate function* (*DRF*) of the i.i.d. distribution of the sequence at code rate $\bar{R}$ and is denoted by $\delta_n(\bar{R})$.

In his source coding theorem, Shannon showed that, as the number of jointly described source symbols n extends to infinity, the operational DRF $\delta_n(\bar{R})$ converges to the informational DRF. The latter is defined as

$$D(\bar{R}) \triangleq \inf \mathbb{E}(X - \hat{X})^2, \tag{S2}$$

where the infimum is over all joint probability distributions $p(x, \hat{x})$ such that their marginal over the x coordinate coincides with the distributions of X_1 and their mutual information does not exceed $\bar{R}$. For example, when the source sequence is drawn from a standard normal distribution, the result of the prior optimization leads to

$$D(\bar{R}) = 2^{-2\bar{R}}. \tag{S3}$$

Comparing with the distortion under scalar quantization in "Scalar Quantization," this value is strictly smaller than the minimal distortion in encoding the same sequence using either fixed or variable bit-length scalar quantization. This difference is explained by the fact that as n extends to infinity, the law of large numbers implies that the probability mass of n i.i.d. copies of a random variable of bounded variance concentrates around the edges of an n-dimensional sphere of radius equal to the square root of this variance. Thus, these n copies can be represented in a more compact manner than with independent representations of each coordinate, as in scalar quantization [39].

losing generality we can assume that each reconstruction waveform produced by the decoder is only a function of one of these states, so there are at most $2^{\lfloor TR \rfloor}$ possible reconstruction waveforms. Moreover, any encoder that strives to attain the minimum MSE (MMSE) in this system would map the input signal to the state i associated with the reconstruction waveform $\hat{x}_i(t)$ that is closest to the input in the distance defined by the L_2 norm over the interval $[-T/2, T/2]$, as derived from our distortion criterion. Therefore, the only freedom in designing the optimal encoding scheme is in deciding on the set of reconstruction waveforms $\{\hat{x}_i(t), t \in [-T/2, T/2], i = 1, \ldots, 2^{\lfloor TR \rfloor}\}$, which we denote as *codewords*.

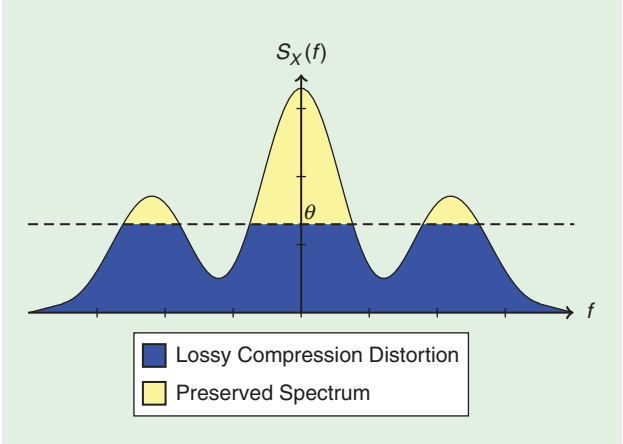


FIGURE 9. The reverse water-filling interpretation of (6). Water is poured into the area bounded by the graph of $S_X(f)$ up to level θ. The bit rate R is tied to the water level θ through the preserved part of the spectrum (6b). The lossy compression distortion D is given by (6a).

The procedure for selecting these codewords and the resulting MMSE distortion are given by Shannon's classical source coding theorem [4], [5] and its extensions to continuous alphabets [33], [34]. According to this theorem, a near-optimal set of codewords is obtained by $2^{\lfloor TR \rfloor}$ independent random draws from a distribution on the set of functions over $[-T/2, T/2]$ with a finite L_2 norm, such that the mutual information of the joint distribution of the input and the reconstruction waveforms is limited to $\lfloor TR \rfloor$ bits. Moreover, Shannon's theorem also provides the asymptotic MMSE obtained by using this set of codewords, denoted as *Shannon's function* or the *information DRF* of the source signal $X(t)$ at bit rate R.

Shannon's source coding theorem with respect to a discrete-time independent and identically distributed process is explained in "Source Coding and the DRF." In the case of a continuous-time Gaussian stationary input signal with a PSD of $S_X(f)$, a closed-form expression for Shannon's DRF was derived by Pinsker and Kolmogorov [35] and is given by the following parametric form:

The Water-Filling Scheme

In "Source Coding and the DRF," we explored the encoding of an independent and identically distributed (i.i.d.) Gaussian sequence using a code of rate $\bar{R}$ bits per sample. We now extend this source coding problem to consider the joint encoding of m i.i.d. sequences taken from m Gaussian distributions with variances $\sigma_1^2, \ldots, \sigma_m^2$, using a total of $\bar{R}$ bits and under a sum mean squared error criterion.

From (S3) we see that it is possible to describe the ith sequence using $\bar{R}_i$ bits per symbol, such that $\sum_i R_i \leq \bar{R}$ and the overall distortion with respect to all sequences is

$$D(R_1, \ldots, R_m) = \sum_{i=1}^{m} \sigma_i^2 2^{-2R_i}. \tag{S4}$$

The problem we consider next is how to allocate the total bit budget $\bar{R}$ in a way that minimizes the overall distortion. This is a convex problem whose solution can be expressed by the following parametric expression [42, Ex. 5.2]:

$$R_i^* = \frac{1}{2}\log_2^+[\sigma_i^2/\theta],$$

where θ is chosen to satisfy the constraint $R = \sum_{i=1}^{m} R_i^*$. The resulting distortion-rate function (DRF) is

$$D(\bar{R}) = D(R_1^*, \ldots, R_m^*) = \sum_{i=1}^{m} \min\{\sigma_i^2, \theta\}.$$

This parametric expression for the DRF is referred to as a *water-filling scheme*. The parameter θ may be interpreted as a water level, such that $D(\bar{R})$ is obtained by summing the part of the variances that are below this level (see Figure S1).

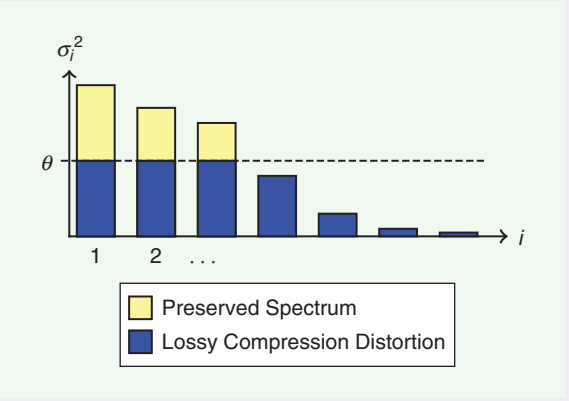


FIGURE S1. The total distortion equals the sum of the part of the variances that lie below the water-level θ.

Intuitively, components with higher variance are described with more bits, since they have a higher impact on the total distortion. An interesting property of the water-filling scheme is that, when R is small, the optimal coding does not allocate any bit budget to some of the components with the lowest variance. This means that no information is sent on these low-variance components.

When the source is a stationary process, the DRF is described by water-filling over the power spectral density of the process, as in (6). In this case, different frequency subbands correspond to different independent signal components, and (6) is obtained by solving an optimization similar to that of minimizing (S4) over $R_1, \ldots, R_m$ [36].

$$D(R_\theta) = \int_{-\infty}^{\infty} \min\{S_X(f), \theta\} df \tag{6a}$$

$$R_\theta = \frac{1}{2}\int_{-\infty}^{\infty} \log^+[S_X(f)/\theta] df, \tag{6b}$$

where $[x]^+$ is the maximum between x and zero. The parametric form of (6) has the graphical interpretation given by Figure 9, denoted as the *water-filling scheme*. The distortion in (6a) may be seen as if water is being poured into the area bounded by the graph of $S_X(f)$ up to level θ. The distortion in (6a) is the total volume of the water. The bit rate is determined by integration over the preserved part through (6b). As explained in "The Water-Filling Scheme," this approach is obtained as the solution of an optimization problem involving the allocation of the rate of the codes to describe different frequency components of the signal according to their respective energy (components with higher energy are given a higher code rate). As a result, in addition to the minimal distortion subject only to the bit-rate constraint, the water-filling interpretation provides the optimal coding scheme that attains this minimal distortion [36]. Independent spectral components of the signal are represented using independent bitstreams, where the rate of each bitstream is determined according to the water-filling principle.

The Pinsker–Kolmogorov expression (6) is easily adjusted to account for a distortion criterion that assigns different weights $W(f) \geq 0$ to each spectral component. This spectral weighting is useful in applications where some tones are of different importance than others, such as in psychoacoustic consideration in the digital encoding of audio signals [37]. The adjustment of the expression for the minimal distortion required because of this importance weighting is achieved by evaluating the distortion equation (6a) with respect to $W(f)S_X(f)$ rather than $S_X(f)$, in a way similar to the procedure explained in [38]. This different weighting emphasizes the generality of the lossy compression principle. Under a strict bit-rate budget, part of the analog signal must be removed due to lossy compression, and this part is the least important in our application.

The Pinsker–Kolmogorov expression, with a possible spectral reweighting, provides a mechanism to determine those parts of the signal that should be removed in an optimal encoding subject to the bit-rate constraint.

This provides the minimal distortion in any system that is used to recover a length T realization of $X(t)$ having no more than $2^{\lfloor TR \rfloor}$ states. A special case of such a system is PCM of the section "ADX Via Pulse-Code Modulation," and, therefore, when $X(t)$ is a Gaussian process, the distortion in (5) is bounded from below by (6).

In general, the optimal encoder that attains Shannon's DRF operates in continuous time. Upon receiving a realization of $X(t)$ over $[-T/2, T/2]$, the encoder compares this realization to each of the $2^{\lfloor TR \rfloor}$ reconstruction waveforms [33], [34]. We note, however, that Shannon's DRF is attainable even if this encoder is required to first map the analog waveform to a discrete-time sequence. Indeed, this discrete-time sequence can be the random coefficients in the analog signal's expansion according to some predetermined orthogonal basis. Consequently, encoding and decoding may be performed with respect to this discrete sequence without changing the fundamental distortion limit described by the DRF in (6). We emphasize that the equivalence between analog signals and coefficients in their basis expansion holds regardless of whether the original process $X(t)$ is bandlimited or not [40].

One commonly used example for such an orthogonal basis is the Karhunen–Loèeve (KL) basis [41]. The latter's functions are chosen as the eigenfunctions of the bilinear kernel defined by the covariance of $X(t)$. As a result, the coefficients in this expansion are orthogonal to each other and, in fact, independent in our case of Gaussian signals. This fact implies that the KL expansion decomposes the process $X(t)$ over the interval $[-T/2, T/2]$ into a discrete Gaussian sequence of independent random variables, where the variance of each element is proportional to the eigenvalue associated with the eigenfunction. Since $X(t)$ is stationary, multiple sequences of this type obtained from different length T blocks of $X(t)$ are identically distributed and, therefore, can be encoded using the same block T encoder that essentially encodes multiple discrete Gaussian sequences. The optimal encoding of such a sequence using $\lfloor TR \rfloor$ bits is achieved according to the water-filling principle, as described in "The Water-Filling Scheme." Moreover, as T extends to infinity, the density of the KL eigenvalues is described by the PSD $S_X(f)$ of $X(t)$, and the average distortion in encoding each block converges to (6) [41]. The prior described coding procedure is one way to show that Pinsker and Kolmogorov's water-filling expression (6) is attainable.

To implement any of the optimal encoding schemes of the analog signal described previously, it is required to represent it first by a discrete sequence of coefficients. However, the implementation of this transformation is subject to practical limitations. In particular, realizable hardware such as filters and pointwise samplers are limited in the number of coefficient values they produce per unit time [7]. That is, for a time lag T, there exists a number f_s such that any system consisting of these operations does not produce more than $T f_s$ analog samples. In the next section, we explore the minimal distortion that can be attained under this restriction. We are especially interested in the minimal sampling rate f_s that is required to achieve Shannon's DRF.

ADX via sampling

We have seen that the optimal tradeoff between MSE distortion and bit rate in the digital representation of an analog signal is described by Shannon's DRF of the signal. In this section, we explore the minimal distortion under the additional constraint that the digital representation must be a function of the samples of the analog signal, rather than the analog signal itself.

Lossy compression from samples

In the ADX setting of Figure 2, the encoder observes samples of the source signal $X(t)$, and is required to encode these samples so that $X(t)$ can be estimated from this encoding using MMSE. Specifically, assuming that the sampler observes $X(t)$ for $t \in [-T/2, T/2]$, we denote by $\mathbf{Y}$ the $\lfloor Tf_s \rfloor$-dimensional random vector resulting from sampling $X(t)$ at rate f_s. The encoder maps the vector $\mathbf{Y}$ to a digital word of length $\lfloor TR \rfloor$ and delivers this sequence without errors to the decoder. The latter provides an estimate $\hat{X}(t)$ for $X(t)$, $t \in [-T/2, T/2]$, based on only the digital sequence and the statistics of $X(t)$. The distortion between $X(t)$ and its reconstruction for a fixed sampler S is defined by

$$D_S(f_s, R) = \inf \frac{1}{T} \int_{-T/2}^{T/2} \mathbb{E}\left(X(t) - \hat{X}(t)\right)^2 dt. \tag{7}$$

The infimum in (7) is over encoders, decoders, and time horizons T. We note that, under the assumption that $X(\cdot)$ and its samples are stationary, any finite time-horizon encoding strategy may be transformed into an infinite time-horizon strategy by applying it to consecutive blocks. As a result, increasing the time horizon cannot increase the distortion, and the minimum over the time horizon in (7) can be replaced by the limit $T \to \infty$.

As an example, in the PCM encoding described in the "ADX Via Pulse-Code Modulation" section, S is a pointwise sampler at sampling rate f_s preceded by an LPF. The particular encoder and decoder used in PCM was described in Figure 4. Therefore, since the optimization in (7) is over all encoders and decoders, for any signal for which pointwise sampling is well defined, we have $D_S(f_s, R) \le D_{\mathrm{PCM}}(f_s, R)$.

Characterizing $D_S(f_s, R)$ gives rise to a source coding problem in which the encoder has no direct access to the source signal it is required to describe. Source coding problems of this type are referred to as *remote* or *indirect* source coding problems [6]. More details on this class of problems is provided in "Indirect Source Coding." Under the MSE criterion (7), the optimal encoding scheme of most indirect source coding problems is obtained by a simple two-step procedure [43], [44], [6]:

1) Estimate $X(t)$ from its samples $\mathbf{Y}$ subject to the MSE criterion (7), i.e., compute the conditional expectation

The characterization of the optimal encoding scheme and the resulting minimal distortion in Figure 2 can be seen as a special case of a family of source coding problems in which the encoder does not observe the source process X directly. Instead, it observes another process Y, statistically correlated with X, where the relation between the two processes is given by a conditional probability distribution $P_{Y|X}$, as in Figure S2.

This setting describes a compression problem in which the encoder is required to describe the source X using a code of rate R bits per source symbol, but with only partial information on X as provided by the signal Y. In information theory, this problem is referred to as the *indirect*, *remote*, or *noisy* source coding problem, which was first introduced in [38]. The optimal tradeoff between code rate and distortion in this setting is denoted as the *indirect distortion-rate function* (*iDRF*). For example, when the source is an independent and identically distributed (i.i.d.) Gaussian process $X = X_1, X_2, \ldots$ and the observable process at the encoder is $Y_n = X_n + W_n$, where W_n is an i.i.d. Gaussian noise sequence independent of X, the iDRF is given by

$$D_{X|Y}(R) = \mathrm{mmse}(X \mid Y) + \mathrm{Var}\left(\mathbb{E}[X \mid Y]\right)2^{-2R}, \qquad \text{(S5)}$$

where $\mathrm{mmse}(X \mid Y)$ is the minimum mean squared error (MMSE) in estimating X_n from Y_n, and $\mathrm{Var}(\mathbb{E}[X \mid Y])$ is the variance of this estimator. Comparing (S5) with Shannon's distortion-rate function (DRF) of X in (S3), we see that the first term in (S5) is the MMSE in estimating the source from its observations, and the second term is Shannon's DRF of the mean squared error estimator. The decomposition of the iDRF into an MMSE term plus the DRF of the estimator is a general property of the indirect source coding setting for any ergodic source pair (X, Y) under quadratic distortion [43]. In the analog-to-digital compression setting of Figure 2, this decomposition takes on the form of (8).

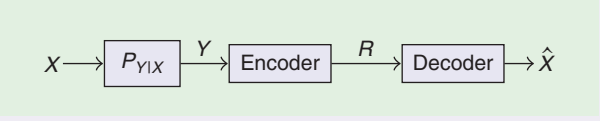


FIGURE S2. Indirect source coding: the source process X is not directly observed.

$\tilde{X}(t) = \mathbb{E}[X(t) \mid \mathbf{Y}]$, where $\mathbf{Y}$ is the output of the sampler with input $X(t)$, $t \in [-T/2, T/2]$.

2) Encode the estimated signal as in a standard (direct) source coding problem at rate R, i.e., encode $\tilde{X}(t)$ as the source signal to the system in Figure 8.

These two steps are shown in Figure 10. We note that although the encoding in step 2 is with respect to an analog signal and, hence, prone to the same sampling limitation in processing analog signals that was mentioned in the "Minimal Distortion Subject to a Bit-Rate Constraint" section, the input to step 1 is a discrete-time process. Therefore, the composition of steps 1 and 2 is a valid coding scheme for the encoder in the ADX setting, since it takes as its input a discrete-time sequence of samples and outputs a binary word.

As explained in "Indirect Source Coding," the two-step encoding procedure leads to the following decomposition:

$$D_S(f_s, R) = \mathrm{mmse}_S(f_s) + D_{\tilde{X}}(R), \qquad (8)$$

where $\mathrm{mmse}_S(f_s)$ is the asymptotic noncausal MMSE in estimating $X(t)$ from the output $\mathbf{Y}$ of the sampler S, and $D_{\tilde{X}}(R)$ is Shannon's DRF of the estimated process $\tilde{X}(t)$.

The decomposition in (8) has a few important consequences. First, it reduces the characterization of $D_S(f_s, R)$ to the evaluation of the MMSE in sampling plus the evaluation of Shannon's DRF of another signal, defined as the noncausal instantaneous MMSE estimator of $X(t)$, given its samples. In particular, these two quantities are independent of the time horizon T, and the MMSE term $\mathrm{mmse}_S(f_s)$ is independent of the bit rate R. In addition, this decomposition implies that, for any sampler S, the minimal distortion is always bounded from below by the MMSE in this estimation, as shown in Figure 3. Moreover, it follows from (8) that whenever the sampling operation is such that $X(t)$ can be recovered with zero MSE from its samples, then $D_S(f_s, R)$ reduces to Shannon's DRF of the source signal $X(t)$. For example, this last situation occurs when $X(t)$ is bandlimited and the sampling is uniform at any sampling rate exceeding the Nyquist rate of $X(t)$, as seen in the "ADX Via Pulse-Code Modulation" section.

This last property implies that oversampling cannot increase $D_S(f_s, R)$, as opposed to the PCM distortion of the "ADX Via Pulse-Code Modulation" section, which increases when the

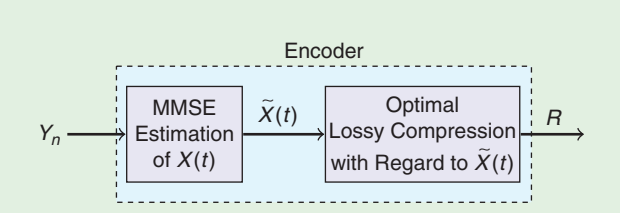


FIGURE 10. The optimal encoder in the ADX setting first estimates the analog source from its samples $\mathbf{Y}$ and then encodes this estimate in an optimal manner.

sampling rate goes above the Nyquist rate of the input signal. This fact highlights an important distinction between the optimal encoder we consider in the definition of $D_S(f_s, R)$ and the encoder in PCM. While the scalar quantizer in PCM encodes each sample instantaneously and independently, the optimal encoder can observe an unlimited number of samples by increasing the time horizon T before deciding on a single index out of $2^{\lfloor TR \rfloor}$. This index is chosen to best describe the realization of $X(t)$ based on the samples stacked in its buffer up until time T. Oversampling $X(t)$ provides the encoder with redundant information to make this choice, which cannot result in a worse choice and, hence, cannot result in worse performance.

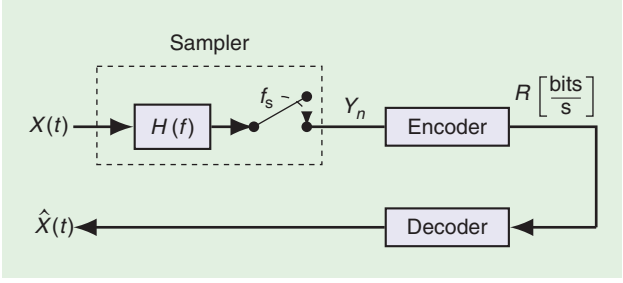


FIGURE 11. ADX with an SI uniform sampler.

Next, we study the behavior of $D_S(f_s, R)$ under various classes of samplers. We begin with samplers that can be described by the concatenation of a linear time-invariant filter and a uniform pointwise evaluation of the filtered signal, as shown in Figure 11 [7]. We then gradually generalize the sampling mechanism to address more general forms of linear continuous sampling, as described in "Generalized Sampling of Random Signals."

Shift-invariant sampling

The system of Figure 11 described the combined sampling and source coding system under a specific class of samplers. Each sampler in this class consists of a linear time-invariant filter applied to the analog source followed by a pointwise evaluation of the filter's output every $T_s = f_s^{-1}$ time units. Therefore, this sampler is characterized only by its sampling rate f_s and the frequency response $H(f)$ of the presampling operation. Samplers of this form are called *shift invariant,* (*SI*) since their operation is equivalent to taking $\lfloor Tf_s \rfloor$ inner products, with respect to the functions $h(t - nT_s)$ [7], for $n \in \mathbb{Z}$. When this sampler is used in the combined sampling and coding system of Figure 2, the resulting system model is shown in Figure 11. In this system, at each time T, the encoder observes the length $\lfloor Tf_s \rfloor$ vector of samples of the filtered

Generalized Sampling of Random Signals

Let $\mathcal{X}$ be a class of signals defined over the entire real line. We define the linear continuous sampling of $\mathcal{X}$ at the sampling rate f_s by the $\lfloor Tf_s \rfloor$ linear continuous functionals of $\mathcal{X}$. That is, when denoting the bilinear operation between $\mathcal{X}$ and its continuous dual $\mathcal{X}^*$ by an integral, the nth sample is given by

$$y_n = \int_{-\infty}^{\infty} x(t)\, g_n(t)\, dt, \tag{S6}$$

where $g_n \in \mathcal{X}^*$. To incorporate sampling techniques that arise in practice, the class of signals $\mathcal{X}$ is chosen such that pointwise evaluation is continuous, i.e., the Dirac distribution $\delta(t)$ belongs to $\mathcal{X}^*$.

When the source $X(t)$ is a random signal, the set of functionals is often associated with the statistics of the signal. To define the counterpart of (S6) when $X(t)$ is a stationary process with known statistics, we use the Fourier transform relation between the covariance of $X(t)$ and its power spectral density:

$$\mathbb{E}[X(t)\, X(s)] = \mathbb{E}[X(t-s)\, X(0)] = \int_{-\infty}^{\infty} e^{2\pi i (t-s) f}\, S_X(f)\, df. \tag{S7}$$

This equation defines an isomorphism between the Hilbert space generated by the closed linear span of the random source signal $X(t) = \{X(t), t \in \mathbb{R}\}$ with norm $\|X(t)\|^2 = \mathbb{E}[X^2(t)]$ and the Hilbert space $L_2(S_X)$ of complex-valued functions generated by the closed linear span (CLS) of the exponentials $\mathcal{E} = \{e^{2\pi i f t}, t \in \mathbb{R}\}$ with an L_2 norm weighted by $S_X(f)$ [45]. This isomorphism allows us to define sampling of the random signal $X(t)$ by describing its operation on the exponentials $\mathcal{E}$. Specifically, for any linear continuous functional h on the CLS of $\mathcal{E}$, denote

$$\phi_h(f) = \int_{-\infty}^{\infty} e^{2\pi i f t}\, h(t)\, dt. \tag{S8}$$

As long as ϕ_h is in $L_2(S_X)$, the sample of $X(t)$ by the functional h is defined by the inverse map of ϕ_h under the aforementioned isomorphism. For example, pointwise evaluation of $X(t)$ at time n/f_s is obtained when h is the Dirac distribution at $t = n/f_s$ and is well defined as long as the L_1 norm of $S_X(f)$ is finite. The last condition requires that $X(t)$ is bounded in energy, which is one of the few assumptions in our analog-to-digital compression setting.

The shift-invariant uniform sampler of Figure 11 corresponds to sampling with functionals $h(t - n/f_s), n \in \mathbb{Z}$, where h is an arbitrary linear continuous functional on the CLS of $\mathcal{E}$. Similarly, uniform multibranch sampling is obtained by sampling with respect to $h_1(t - nL/f_s), \ldots, h_L(t - nL/f_s)$, where $h_1, \ldots, h_L$ are L such functionals.

MMSE Under Sub-Nyquist Sampling

Consider the noncausal estimation of the process $X(t)$ from the discrete-time process Y_n at the output of the shift-invariant sampler of Figure 11. Since all signals are Gaussian, the optimal estimator and the resulting minimum mean squared error (MMSE) can be found using linear estimation techniques that generalize the Wiener filter [46], [47]. In our case, the optimal estimator $\tilde{X}(t) = \mathbb{E}[X(t) \mid Y]$ is given by

$$\tilde{X}(t) = \sum_{n\in\mathbb{Z}} Y_n w(t - nT_s), \quad t \in \mathbb{R}, \tag{S9}$$

where the Fourier transform of $w(t)$ is

$$W(f) = \frac{S_X(f)|H(f)|^2}{\sum_{k\in\mathbb{Z}} S_X(f - kT_s)|H(f - kT_s)|^2}.$$

The resulting MMSE is given by

$$\mathrm{mmse}_{\mathrm{SI}}(f_s) = \sum_{n\in\mathbb{Z}} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} [S_X(f - nf_s) - \tilde{S}_{X|Y}(f)]\,df, \tag{S10}$$

where

$$\tilde{S}_{X|Y}(f) \triangleq \frac{\sum_{n\in\mathbb{Z}} S_X^2(f - f_s n)|H(f - f_s n)|^2}{\sum_{n\in\mathbb{Z}} S_X(f - f_s n)|H(f - f_s n)|^2}. \tag{S11}$$

We interpret this fraction to be zero whenever both the numerator and denominator are zero.

When f_s is above the Nyquist rate of $X(t)$, the support of $S_X(f)$ is contained within the interval $(-f_s/2, f_s/2)$. It can be seen from (S9) that, in this case, provided that $H(f)$ is nonzero over the support of $S_X(f)$, we have that $\tilde{X}(t) = X(t)$, $\tilde{S}_{X|Y}(f)$ coincides with $S_X(f)$, and, therefore, $\mathrm{mmse}_{\mathrm{SI}}(f_s) = 0$. Hence, as the time horizon extends to infinity, it is possible to reconstruct $X(t)$ from its samples with the zero mean squared error. Alternatively, when f_s is below the Nyquist rate, (S10) shows how the MMSE in this estimation is affected by aliasing, i.e., interference of different frequency components of the signal due to sampling.

source at instances $\ldots, -T_s, 0, T_s, \ldots$ inside the interval $[-T/2, T/2]$. The decoder receives the length $\lfloor TR \rfloor$ binary sequence produced by the encoder from this vector. We denote the MMSE in recovering the source from this binary sequence as T extends to infinity by $D_{\mathrm{SI}}(f_s, R)$.

From the general decomposition (8), it follows that the minimal distortion for a SI sampler is obtained as the sum of the MMSE in estimating $X(t)$ from its filtered and uniform samples at rate f_s, plus Shannon's DRF of the noncausal estimator from these samples. As explained in "MMSE Under Sub-Nyquist Sampling," this MMSE vanishes whenever f_s exceeds the Nyquist rate of $X(t)$, provided that the presampling filter $H(f)$ does not block any part of the signal's spectrum $S_X(f)$. In this situation, the estimator $\hat{X}(t)$ coincides with the original signal $X(t)$ in the L_2 sense, and the decoder essentially encodes $X(t)$ directly, as in the previous section. Therefore, for bandlimited signals, we conclude that $D_{\mathrm{SI}}(f_s, R)$ equals Shannon's DRF of $X(t)$ when the sampling rate is above the Nyquist rate. Moreover, when $X(t)$ is not bandlimited, a similar equality holds as the sampling rate extends to infinity [48].

When the sampling rate is below the Nyquist rate, the expression for the optimal estimator and the resulting MMSE are obtained by standard linear estimation techniques, as explained in "MMSE Under Sub-Nyquist Sampling." In this case, the estimator $\tilde{X}(t)$ has the form of a stationary process modulated by a deterministic pulse and is therefore a block-stationary or a cyclostationary process [49]. It is shown in [50] that Shannon's DRF for this class of processes can be described by a generalization of the orthogonal transformation and rate allocation that leads to the water-filling expression (6), in a way analogous to the description in "The Water-Filling Scheme." By evaluating the resulting expression for the DRF of the cyclostationary process $\tilde{X}(t)$ and using the decomposition (8), we obtain the following closed-form formula for $D_{\mathrm{SI}}(f_s, R)$, initially derived in [51]:

$$D_{\mathrm{SI}}(f_s, R_\theta) = \mathrm{mmse}_{\mathrm{SI}}(f_s) + \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \min\{\tilde{S}_{X|Y}(f), \theta\}\,df \tag{9a}$$

$$R_\theta = \frac{1}{2}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \log_2^+[\tilde{S}_{X|Y}(f)/\theta]\,df, \tag{9b}$$

where $\mathrm{mmse}(X \mid Y)$ and $\tilde{S}_{X|Y}(f)$ are given by (S10) and (S11), respectively. The parametric expression (9) combines the MMSE (S10), which depends only on f_s and $H(f)$, with the reverse water-filling expression (6), which also depends on the bit rate R. The function $\tilde{S}_{X|Y}(f)$ arises in the MMSE estimation of $X(t)$ from its samples. As explained in [50], this function is the average over the PSD of each polyphase component of the cyclostationary process $\tilde{X}(t)$. To summarize, (9) provides the MMSE distortion in encoding a Gaussian stationary signal at rate R from its uniform samples taken at rate f_s. Moreover, according to Figure 10, the coding scheme that attains this minimal distortion can be described by the composition of the noncausal MMSE estimate of $X(t)$ as in (S9), followed by an optimal encoding of the estimated process to attain its Shannon's DRF.

It is possible to extend the system model of Figure 2 to include a noisy input signal before the sampler. In this extended model, the excess distortion is a result of lossy compression, sampling,

and independent noise. Therefore, the problem of estimating the source signal from the digital output of the encoder combines a linear filtering problem, an interpolation problem, and a lossy compression problem. The only adjustment to the description of the minimal distortion under this extension is to replace the function $\tilde{S}_{X|Y}(f)$ in (9) and (S10) with [51]

$$\tilde{S}_{X|Y}(f) = \frac{\sum_{n\in\mathbb{Z}} S_X^2(f-f_s n)\,|H(f-f_s n)|^2}{\sum_{n\in\mathbb{Z}} (S_X(f-f_s n)+S_\eta(f-f_s n))\,|H(f-f_s n)|^2}. \quad (10)$$

Equations (S10), (9), and (10) describe the MMSE in noncausal filtering, the MSE due to uniform sampling, and the distortion under optimal lossy compression. That is, these equations determine the combined effect of three of the most fundamental operations in signal processing: quantization, sampling, and interference by noise. Most importantly, these equations provide a unified representation for the distortion in these three fundamental operations, allowing us to explore the interaction among them. In this article, we consider a less general case; we explore the interaction between sampling and lossy compression and assume that the noise is zero. Hence, the simplified form (S11) for $\tilde{S}_{X|Y}(f)$ is used.

As a simple example for using (9), we consider $X(t)$ to be a stationary Gaussian signal with a flat, bandlimited PSD, i.e.,

$$S_\Pi(f) = \begin{cases} \frac{1}{2W}, & |f| < W, \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

As long as the presampling filter passes all frequencies $f \in (-W, W)$, the relation between the distortion in (9a) and the bit rate in (9b) is given by

$$D_{\mathrm{SI}}(f_s, R) = \begin{cases} \mathrm{mmse}_{\mathrm{SI}}(f_s) + \frac{f_s}{2W} 2^{-\frac{2R}{f_s}}, & \frac{f_s}{2W} < 1 \\ 2^{-\frac{R}{W}}, & \frac{f_s}{2W} \geq 1, \end{cases} \quad (12)$$

where $\mathrm{mmse}_{\mathrm{SI}}(f_s) = 1 - f_s/2W$. Expression (12) is shown in Figure 12 for two fixed values of the bit rate R. It has a very intuitive structure—for frequencies below the signal's Nyquist rate $2W$, the distortion as a function of the rate increases by a constant factor because of the error as a result of nonoptimal sampling. This factor completely vanishes once the sampling rate exceeds the Nyquist frequency, in which case $D_{\mathrm{SI}}(f_s, R)$ coincides with Shannon's DRF of $X(t)$.

In the previous example with PSD $S_\Pi(f)$, the filter $H(f)$ has no effect on the distortion as long as its passband contains the support of $S_\Pi(f)$. However, when the spectrum is nonflat over its support, there is a precise way to choose the passband of the presampling filter to minimize the function $D_{\mathrm{SI}}(f_s, R)$.

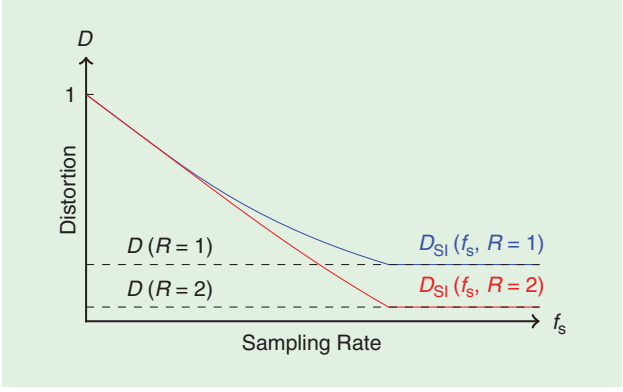


FIGURE 12. The distortion as a function of the sampling rate for the source with PSD $S_\Pi(f)$ of (11) and source coding rates $R = 1$ and $R = 2$ bits per time unit.

Optimal sampling rate under bit-rate constraint

We now consider the expression $D_{\mathrm{SI}}(f_s, R)$ of (9) for the unimodal PSD shown in Figure 13, where the presampling filter $H(f)$ is an ideal LPF with a cutoff frequency of $f_s/2$. This LPF operates as an antialiasing filter, and, therefore, the part

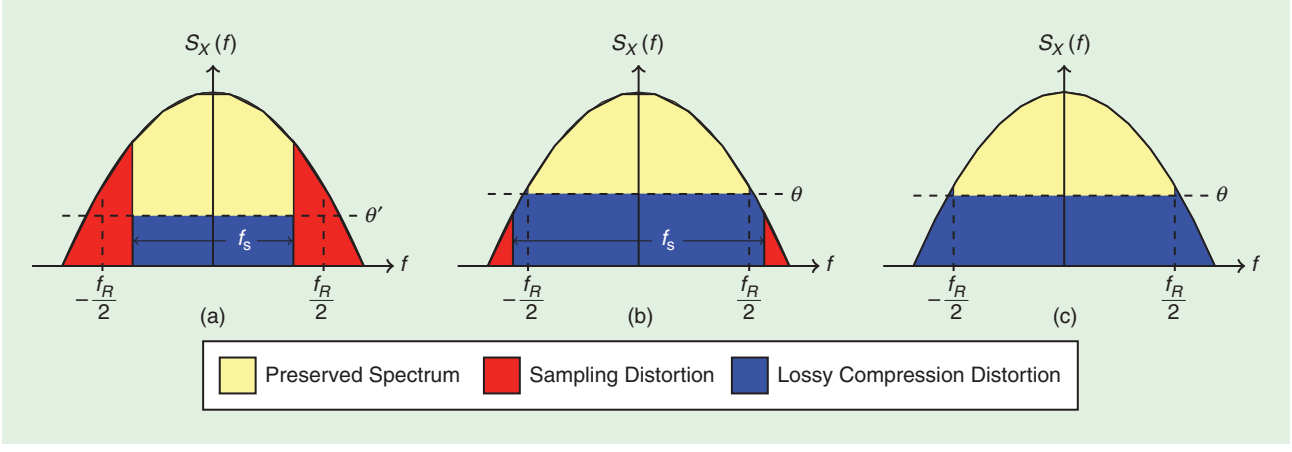


FIGURE 13. A water-filling interpretation of (9) with $H(f)$ as an LPF of cutoff frequency $f_s/2$. The distortion is the sum of the sampling and the lossy compression distortions. All figures correspond to the same bit rate R and different sampling rates: (a) $f_s < f_R$, (b) $f_s \geq f_R$ and (c) $f_s > f_{\mathrm{Nyq}}$. The DRF of $X(t)$ is attained for all f_s greater than $f_R < f_{\mathrm{Nyq}}$.

of $D_{\mathrm{SI}}(f_s,R)$ associated with the sampling distortion is due only to those energy bands blocked by the filter. As a result, the function $D_{\mathrm{SI}}(f_s,R)$ can be described by the sum of the red and the blue parts in Figure 13(a). Figure 13(b) describes the function $D_{\mathrm{SI}}(f_s,R)$ under the same bit rate R and a higher sampling rate, while the cutoff frequency of the LPF is adjusted to this higher sampling rate. As can be seen from the figure, at this higher sampling rate, $D_{\mathrm{SI}}(f_s,R)$ equals the DRF of $X(t)$ in Figure 13(c), although this sampling rate is still below the Nyquist rate of $X(t)$. In fact, it follows from Figure 13 that the DRF of $X(t)$ is attained at some critical sampling rate f_R that equals the spectral occupancy of the preserved part in the Pinsker–Kolmogorov water-filling expression (6). The existence of this critical sampling rate can also be seen in Figure 14, which illustrates $D_{\mathrm{SI}}(f_s,R)$ as a function of f_s with $H(f)$ an LPF.

In the "Shift-Invariant Sampling" section, we concluded that the DRF of $X(t)$ can be attained by sampling at or above the Nyquist rate, since then the MMSE term in (6) vanishes. Now we see that by using the LPF with a cutoff frequency of $f_s/2$, the equality between $D_{\mathrm{SI}}(f_s,R)$ and the DRF, which is the minimal distortion subject to a bit-rate constraint, occurs at a sampling rate smaller than the Nyquist rate.

An intriguing way to explain this phenomena is as an alignment of the degrees of freedom in the signal after the presampling operation with the degrees of freedom that the lossy compression with bit rate R can capture in this sampled signal. For stationary Gaussian signals, the degrees of freedom in the signal representation are those spectral bands where the PSD is nonzero. When the signal energy is not uniformly distributed over these bands—unlike in the example of the PSD in (11)—the optimal lossy compression scheme calls for discarding those bands with the lowest energy, i.e., the parts of the signal with the lowest uncertainty. The presampling operation removes these low-energy signal components such that the resulting signal has the same degrees of freedom as those that can be captured by the lossy compressed signal representation that follows the sampler. Thus, the presampling operation in a sense aligns the degrees of freedom of the presampled signal with those of the postsampled lossy compression operation.

The degree to which the new critical rate f_R is smaller than the Nyquist rate depends on the energy distribution of $X(t)$ along its spectral occupancy. The more uniform it is, the more degrees of freedom are required to represent the lossy compressed signal, and therefore f_R is closer to the Nyquist rate. Figure 15 shows the dependency of f_R on R for various PSD functions. Note that, whenever the energy distribution is not uniform and the signal is bandlimited, the critical rate f_R converges to the Nyquist rate as R extends to infinity and to zero as R reaches zero.

In the prior discussion, we considered only signals with unimodal PSDs (for example, the PSD in Figure 9 is not unimodal). The main challenge in extending the previously mentioned conclusions to signals with nonunimodal PSDs is the design of a sub-Nyquist system that samples the signal components containing the most information about the signal (i.e., the signal's high-energy bands) to obtain the optimal lossy compressed signal representation when these samples are encoded at the fixed bit rate R. Before describing this extension, we consider the general structure of a presampling transformation that minimizes the distortion in the ADX setting.

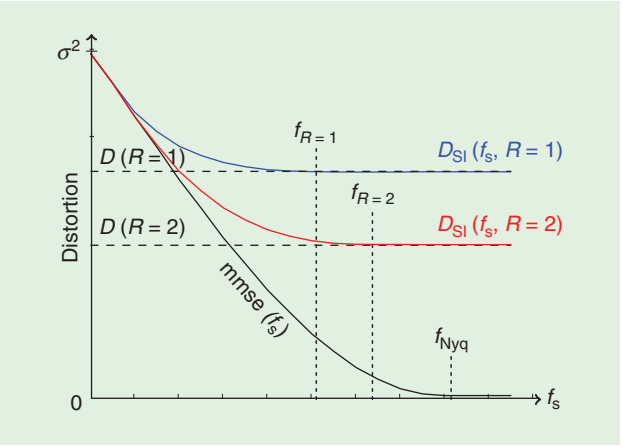


FIGURE 14. The function $D_{\mathrm{SI}}(f_s, R)$ for the PSD of Figure 13 with an LPF with a cutoff frequency of $f_s/2$ and two values of the bit rate R. This function describes the minimal distortion in recovering a Gaussian signal with this PSD from a bit rate-R encoded version of its uniform samples. This minimal distortion is bounded from below by Shannon's DRF of $X(t)$, where the latter is attained at the sub-Nyquist sampling rate f_R.

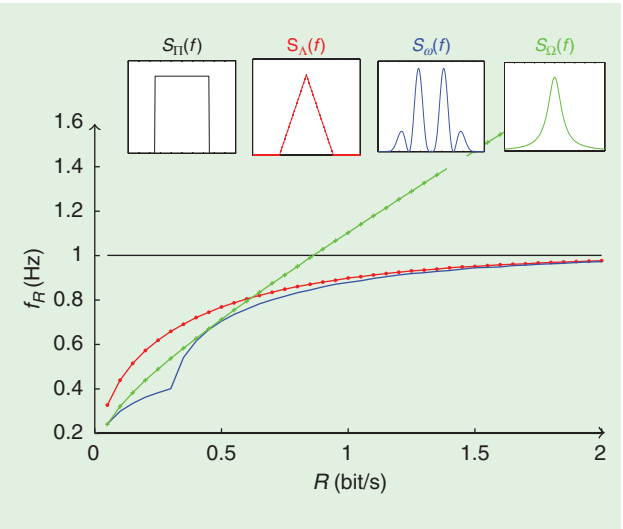


FIGURE 15. The critical sampling rate f_R as a function of the bit rate R for the PSDs given in the small frames. For the bandlimited PSDs $S_{\Pi}(f)$, $S_{\Lambda}(f)$ and $S_{\omega}(f)$, the critical sampling rate is always below the Nyquist rate. The critical sampling rate is finite for any R, even for the nonbandlimited PSD $S_{\Omega}(f)$.

Optimal presampling transformation

We now consider the presampling filter $H(f)$ that minimizes the function $D_{\mathrm{SI}}(f_s,R)$ subject to a fixed bit rate R and sampling rate f_s. By examining expressions (9) and (S10), we conclude that this minimization is equivalent to the maximization

Optimal Presampling Transformation

Properties 1 and 2 of the optimal presampling filter imply that to minimize the mean squared error (MSE) and, hence, the overall distortion, it is preferred to eliminate all information on lower-energy subbands where they interfere with higher-energy bands. To provide an intuitive explanation for this phenomenon, we consider two independent Gaussian random variables X_1 and X_2 with the zero mean and variances σ_1^2 and σ_2^2, respectively. These random variables can be seen as two different spectral lines in the spectrum of $X(t)$ that interfere with each other because of aliasing in uniform sampling. Assume that we are given the linear combination $U = h_1 X_1 + h_2 X_2$, and are interested in the joint estimation of X_1 and X_2 subject to an MSE criterion. That is, we want to minimize

$$\text{mmse}(X_1, X_2 \mid U) \triangleq \mathbb{E}(X_1 - \hat{X}_1)^2 + \mathbb{E}(X_2 - \hat{X}_2)^2.$$

The optimal estimator of each variable as well as the corresponding estimation error can be easily found, since the optimal estimator is linear. We further ask how to choose the coefficients h_1 and h_2 in the linear combination such that the MSE is minimized. A simple optimization over the expression for $\text{mmse}(X_1, X_2 \mid U)$ shows that $h_1 \neq 0, h_2 = 0$ is the answer whenever $\sigma_1^2 > \sigma_2^2$, and $h_1 = 0, h_2 \neq 0$ whenever $\sigma_1^2 < \sigma_2^2$. That is, the optimal linear combination eliminates all the information on the part of the signal with the lowest variance and passes only the part with the highest uncertainty. Going back to spectral components, the MSE is minimized by a presampling filter $H(f)$ that eliminates all spectral components of low energy whenever they interfere with high-energy spectral components because of the aliasing that results from uniform sampling.

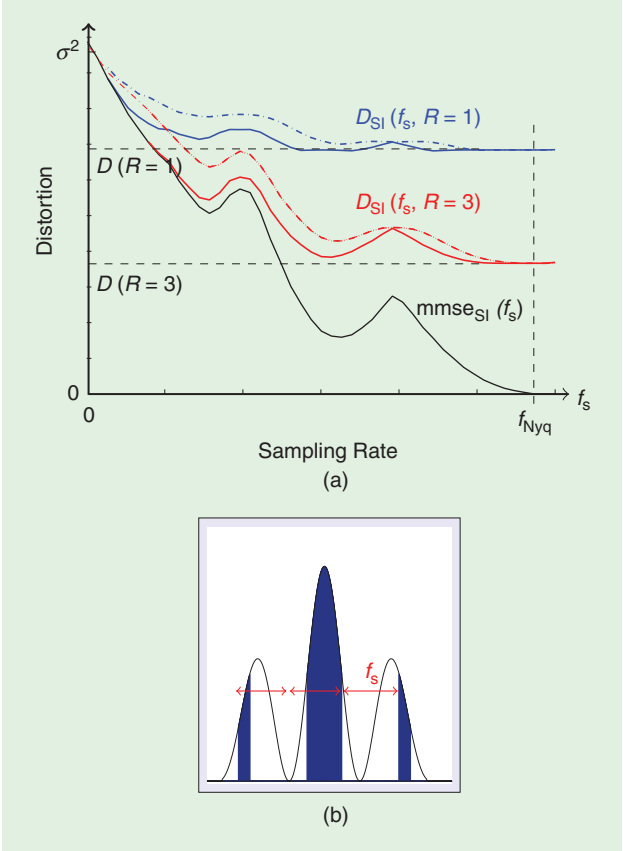


FIGURE 16. (a) The minimal distortion $D_{\text{SI}}(f_s, R)$ using an optimal presampling filter as a function of the sampling rate for two values of the bit rate R. The dashed lines represent the distortion with an all-pass presampling filter that allows aliasing. (b) The support of the optimal presampling filter over the source PSD for a particular sub-Nyquist sampling rate f_s. The difference between any two bands in the support is not an integer multiple of f_s.

of $\tilde{S}_{X|Y}(f)$ for any f in the interval $(-f_s/2, f_s/2)$. This fact is not surprising, since we have seen in (S10) that $\tilde{S}_{X|Y}(f)$ represents the part of the source available to the encoder. Because the function $\tilde{S}_{X|Y}(f)$ is independent of R, the optimal filter $H(f)$ that minimizes $D_{\text{SI}}(f_s, R)$ is only a function of the sampling rate, and it is, therefore, identical to the presampling filter that minimizes $\text{mmse}(X \mid Y)$, i.e., the MMSE without the bit-rate constraint. Note that, since $\tilde{S}_{X|Y}(f)$ is indifferent to scaling in $H(f)$, the only effect of the presampling filter on the distortion is through its passband, i.e., the support of $H(f)$. We explain in "Optimal Presampling Transformation" that the passband of the presampling filter that minimizes $\text{mmse}(X \mid Y)$ can be completely characterized by the following two properties:

1) *Aliasing-free*: The passband is such that the filter eliminates aliasing in sampling at frequency f_s. That is, all integer shifts of the support of the filtered signal by f_s are disjoint.
2) *Energy maximization*: The passband is chosen to maximize the energy of $X(t)$ at the output of the filter, subject to the aliasing-free property 1).

In the case of a unimodal PSD, an LPF with a cutoff frequency of $f_s/2$ satisfies both the aliasing-free and energy maximization properties and is therefore the optimal presampling filter that minimizes $D_{\text{SI}}(f_s, R)$. For this reason, Figure 13 describes the minimal value of $D_{\text{SI}}(f_s, R)$ for the PSD considered there. In general, however, the set that maximizes the passband energy is not aliasing-free. As an example, consider the PSD shown in Figure 16(b). The colored area represents the support of the optimal presampling filter. This support is aliasing-free, since the difference between any two bands in the support is not an integer multiple of f_s. The example in Figure 16(a) also shows that, although $D_{\text{SI}}(f_s, R)$ is guaranteed to coincide with $D(R)$ for $f_s > f_{\text{Nyq}}$, the convergence to this

value may not be monotonic in f_s. That is, some sub-Nyquist sampling rates may introduce more aliasing than sampling rates that are lower. This phenomenon does not occur in sampling signals with a unimodal PSD.

The dependency of the passband of $H(f)$ on the sampling frequency f_s comes from the aliasing-free property. In particular, this property restricts the Lebesgue measure of the passband of any aliasing-free filter to be smaller than f_s [52, Prop. 2]. It follows from this that a lower bound on the function $D_{\mathrm{SI}}(f_s,R)$ is obtained by taking the part of the spectrum of highest energy and overall Lebesgue measure not exceeding f_s. That is, we denote by $F^*(f_s)$ the part of the spectrum that maximizes $\int_F S_X(f)\,df$ over all sets F of Lebesgue measure not exceeding f_s. The following expression bounds the function $D_{\mathrm{SI}}(f_s,R)$ from below:

$$D(f_s,R) = \mathrm{mmse}(f_s) + \int_{F^*(f_s)} \min\{S_X(f),\theta\} \tag{13a}$$

$$R(\theta) = \frac{1}{2}\int_{F^*(f_s)} \log_2^+[S_X(f)/\theta]\,df, \tag{13b}$$

where

$$\mathrm{mmse}(f_s) = \int_{-\infty}^{\infty} S_X(f)\,df - \int_{F^*(f_s)} S_X(f)\,df. \tag{14}$$

A graphical water-filling interpretation of the prior expression is given in Figure 17. In the next section, we describe how to attain this lower bound by extending SI samplers to an array of such samplers.

Multibranch sampling

In contrast to the case of a unimodal PSD, it is, in general, impossible to attain the function $D(f_s,R)$ of (13) using a single SI sampler. Indeed, once we fix a band, no other bands located at integer multiples of the sampling rate are included in the support of the optimal presampling filter because of the aliasing-free property. This limitation implies that the support of the optimal presampling filter does not necessarily consist of a set of measured f_s with the largest signal energy, as in the definition of $D(f_s,R)$. By using more sampling branches, the global aliasing-free property is relaxed to a local aliasing-free property at each sampling branch. Therefore, while each branch has constraints on the position of the bands in the support of its filter to avoid aliasing, the increment in sampling branches allows for more freedom in selecting the overall part of the spectrum passed by all filters. As a result, the union of the supports of an optimal set of L filters that are aliasing-free with respect to f_s/L approximates the set of maximal energy of measure f_s better than is possible with a single filter that is aliasing-free with respect to f_s. This situation is shown in Figure 16. In particular, components that needed to be eliminated in the single-branch case because of aliasing with higher-energy components can now be retained, as these two components can be preserved on separate branches without interference with each other after sampling. In other words, multibranch sampling reduces part of the constraint on retaining desired signal components that arises as a result of the aliasing-free requirement in a single SI filter, leading to higher-energy frequency components in the resulting signal representation before encoding and therefore lower distortion after encoding.

This intuition motivates replacing the SI sampler in Figure 11 with an array of such samplers, as shown in Figure 18. Within each branch, the presampling filter may pass only a narrow part of the signal's spectrum and apply passband sampling [52]. This multibranch uniform sampler covers a wide class of sampling systems used in practice, including single-branch SI sampling, nonuniform periodic sampling, and multicoset sampling [7], [53].

The analysis of the system is greatly simplified if all of the sampling branches have the same sampling rate. Thus, we assume that the sampling rate at each branch equals f_s/L, so that the overall effective sampling rate is f_s. Similar to the case of a single SI sampler, the optimal selection of the presampling filters across all branches leads to a collection of filters with the aliasing-free property at each branch, such that the net energy passed by these filters is maximal [51]. Since the measure of the passband of each aliasing-free filter for sampling at rate f_s/L is

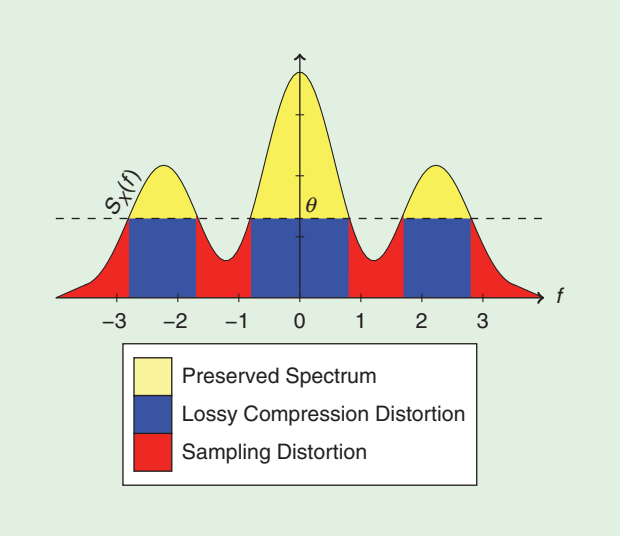


FIGURE 17. A water-filling interpretation of the fundamental minimal distortion $D(f_s,R)$ in ADX. The overall distortion is the sum of the sampling distortion and the lossy compression distortion. The set $F^*(f_s)$ defining $D(f_s,R)$ is the support of the preserved spectrum.

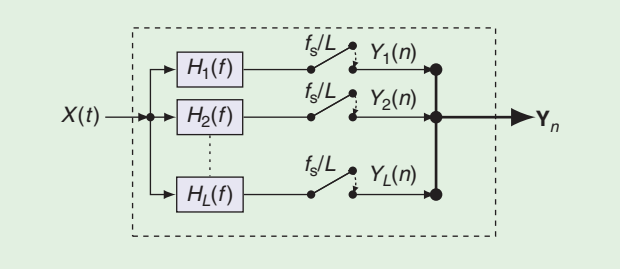


FIGURE 18. A multibranch filter-bank uniform sampler.

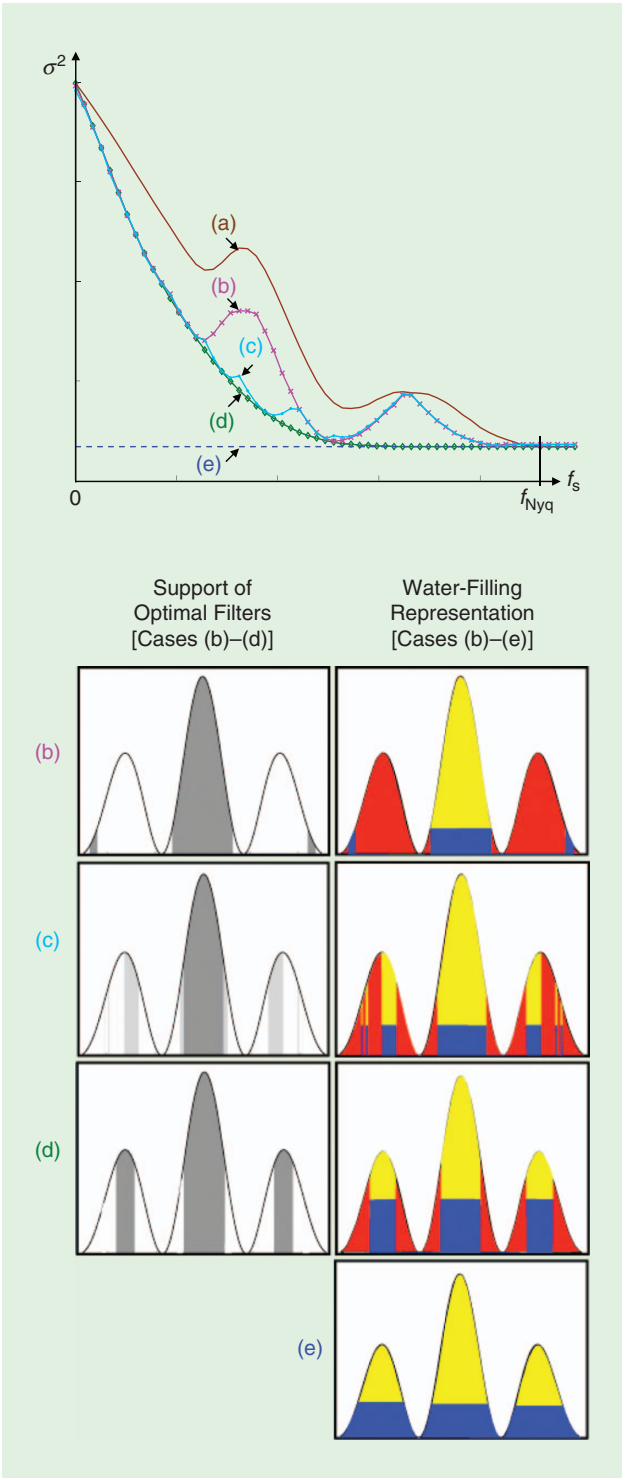


FIGURE 19. The minimal distortion versus the sampling rate f_s for a fixed value of R. The case of no sampling prefilter is given in case (a), and the cases of one, two, and five sampling branches with optimal branch prefiltering are considered in cases (b), (c), and (d), respectively. For each of these cases and a fixed f_s, the union of support for the optimal filters, which equals f_s, is shown in the gray-scale images, and how these bands are identified through water-filling and the sampling distortion that results is shown in the color images. Case (d), of five SI sampling branches, preserves the part of the spectrum of measure f_s with the highest energy, and therefore achieves $D(f_s, R)$. (e) Shannon's DRF with its water-filling representation.

at most f_s/L, the overall part of the spectrum passed by the L filters is at most of size f_s. This property implies that the lower bound $D(f_s,R)$ of (13) is kept under this form of sampling.

The next question is whether this lower bound is attainable, provided that we are allowed to increase the number of sampling branches L and the presampling filters $H_1(f),\ldots,H_L(f)$. A positive answer to this question was given in [51], where it was shown that for any PSD, the distortion level $D(f_s,R)$ can be attained using some finite number L^* of sampling branches and a particular set of filters, each of which is antialiasing for sampling at a rate of f_s/L^*. The reduction of the distortion in ADX using the optimal filter-bank sampler as the number of branches increases is shown in Figure 19. Also shown are the supports of the optimal presampling filters at a specific sampling rate f_s.

We conclude that the function $D(f_s,R)$ describes an achievable lower bound for the distortion in the ADX setting with a multibranch uniform sampler. In the next section, we extend this result to nonuniform and generalized linear sampling procedures.

Nonuniform and generalized sampling

We now extend the ADX setting to include a nonuniform sampling system with time-varying preprocessing. We show that under some mild assumptions on the sampling set, it is impossible to achieve a distortion lower than $D(f_s,R)$, where, here, f_s equals the density of the sampling set. The definition of the density of a sampling set and more detailed background on nonuniform sampling can be found in "Nonuniform Sampling." This extension includes all cases of linear continuous sampling, as given in "Generalized Sampling of Random Signals."

A nonuniform time-varying sampler is shown in Figure 20. It is characterized by a discrete and ordered sampling set of sampling times $\Lambda=\{\ldots,t_{-1},t_0,\ldots,t_n,\ldots\}\subset\mathbb{R}$ and a time-varying impulse response $g(t,\tau)$. The sampling set is assumed to be uniformly discrete, in the sense that there exists a universal constant $\varepsilon>0$ such that each two elements of Λ are at least ε apart. The nth output of the sampler is the convolution of $g(t_n,t)$ with $X(t)$, where $t_n\in\Lambda$. For every finite time lag $[-T/2,T/2]$, the vector $\mathbf{Y}$ is the sampler output at times $[-T/2,T/2]\cap\Lambda$. Our goal is to map this vector to one of $2^{\lfloor TR\rfloor}$ elements and, by observing this element, recover $X(t)$ over this time interval under MSE distortion. We note that although the sampler in Figure 20 has only a single sampling branch, the multibranch sampling system of Figure 18 may be realized by this filter using a particular choice of the time-varying operation [54].

As in the case of uniform sampling, it is instructive to begin our discussion with the lower bound on the minimal distortion obtained by the MMSE in estimating $X(t)$ from its nonuniform sampled version Y_n. A classical result in functional analysis and signal processing due to Landau asserts that a signal can be perfectly recovered from its nonuniform samples if, and only if, the density of Λ exceeds its spectral occupancy [55]. See "Nonuniform Sampling" for an overview of this result. In our setting, the spectral occupancy takes the form of the

Consider a sampling set Λ for which there exists an $\varepsilon > 0$ such that $|t_k - t_n| > \varepsilon$ for every $t_n \neq t_k \in \Lambda$. The density of Λ is defined as the number of elements of Λ contained in a single interval of length r divided by r, in the limit as r extends to infinity and provided this limit exists. For example, the density of a uniform sampling set $\Lambda = f_s\mathbb{Z}$ is f_s.

The isomorphism described by (15) establishes an equivalence between the problem of estimating a Gaussian stationary process from its samples at times Λ under the mean squared error (MSE) criterion, and the problem of orthogonal projection onto the space spanned by $\mathcal{E}(\Lambda) \triangleq \{e^{2\pi i f t_n}, t_n \in \Lambda\}$. The conditions for this MSE to vanish are related to the fact that every element of $L_2(S_X)$ can be approximated by a linear combination of exponentials in $\mathcal{E}(\Lambda)$, [56], [57]. This property, however, turns out to be too weak for practical sampling systems, since it does not guarantee stability: the approximation may not be robust to small perturbations in the time instances that inevitably are present in practice [3], [58], [59]. As a result, only stable sampling schemes [60] should be considered in applications. A necessary and sufficient condition for stable sampling was given by Landau [55], who showed that it can be obtained if, and only if, the density of Λ exceeds the spectral occupancy of $X(t)$.

support of the PSD. Therefore, the function $D(f_s,R)$ of (13) agrees with Landau's characterization, since it implies that as R extends to infinity, zero MSE is attained if, and only if, the sampling rate exceeds the spectral occupancy.

The ADX with the nonuniform sampler extends the prior result, since it considers the case of a limited finite bit rate and linear preprocessing of the samples. For this setting, it is shown in [18] that the lower bound on the distortion $D(f_s,R)$ still holds, provided f_s is replaced by the density of Λ. That is, for any time-varying system $g(t,\tau)$ and any sampling set Λ for which a density exists, the minimal distortion in the ADX setting with a time-varying nonuniform sampler is lower-bounded by $D(f_s,R)$, where f_s equals the density of Λ.

It follows that minimal distortion in the ADX setting under the class of linear pointwise samplers at rate f_s is fully characterized by the function $D(f_s,R)$. In general and according to Landau's condition for stable sampling, an equality between $D(f_s,R)$ and Shannon's DRF of the analog source is expected for sampling rates higher than the spectral occupancy of $X(t)$. We have seen, however, that this equality usually already occurs as the sampling rate f_s exceeds the support of the preserved part of the spectrum in the Pinsker–Kolmogorov water-filling expression (6). In other words, the sampling structure that attains $D(f_s,R)$ utilizes the special structure associated with the optimal lossy compression of analog signals given by the Pinsker–Kolmogorov result. It, in effect, aligns the degrees of freedom of the presampled signal with those of the postsampled lossy compressed signal so that the part of the signal removed prior to the sampling stage matches the part of the signal removed under the optimal lossy compression of the signal subject to the bit-rate constraint.

As a final remark, we note that any linear continuous sampler as defined in "Generalized Sampling of Random Signals" can be expressed as the time-varying nonuniform sampler of Figure 20. Indeed, the kernel of the time-varying operation $g(t,\tau)$ defines a set of linear continuous functionals $g_n(t) = g(t_n,t), t_n \in \Lambda$.

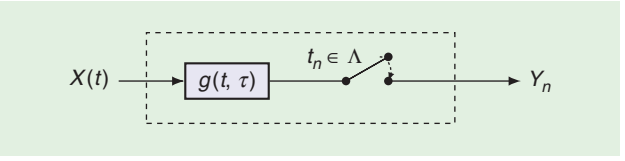


FIGURE 20. A nonuniform sampler with time-varying preprocessing.

Summary of ADX

We have shown that the optimal tradeoff among distortion, bit rate, and sampling rate under the class of linear samplers with pointwise operations is fully described by the function $D(f_s,R)$ of (13). Moreover, the procedure for attaining an optimal point in this tradeoff is summarized in the following steps.

1) Given the bit-rate constraint R, use the Pinsker–Kolmogorov water-filling (6) over the PSD $S_X(f)$. The critical sampling rate f_R is the support of the frequency components associated with the preserved part of the spectrum in this expression.
2) Use a multibranch uniform sampler with a sufficient number of sampling branches optimized such that the combined passband of all of the samplers is the support of the preserved part of the spectrum [52, Sec. IV].
3) Recover the part of the signal associated with the preserved part of the spectrum from all branches, as in standard MSE interpolation [61].
4) Fix a large time lag T and use a vector quantizer with $\lfloor TR \rfloor$ bits to encode the estimate in step 3 over this lag.

The previously described procedure calls for a few comments and extensions. First, we note that, although our description determines the minimal distortion and sampling rate as a function of the bit rate, this dependency can be inverted. That is, given a target distortion D, the Pinsker–Kolmogorov expression (6) leads to a minimal bit rate R and a corresponding sampling rate required to attain this target. Second, steps 1–4 can be easily adjusted to consider a different distortion criterion according to a spectral importance masking, as described in the "Minimal Distortion Subject to

a Bit-Rate Constraint" section. In addition, steps 3 and 4 may be replaced by different techniques to attain the optimal lossy compression performance [6]. For example, the output of each sampling branch can be encoded independently of the other outputs using a separate bitstream. The bit rate of each bitstream is determined by the water-filling principle of (6b), with the PSD replaced by the PSD of the filtered signal at each sampling branch. Finally, we note that the multibranch uniform sampler can be replaced by a nonuniform sampler with a single branch and possibly time-varying operation [54], or fewer uniform sampling branches of different sampling rates. That is, although uniform multibranch sampling attains the minimal distortion $D(f_s,R)$, it may not achieve it using the most compact system implementation. In addition to these extensions, we note that the characterization of the minimal distortion in ADX has also been derived for the Wiener process and for sparse source signals [62], [63].

Applications

The most straightforward application of sampling according to the optimal ADX scheme is the possibility to reduce the sampling rates in systems operating under bit-rate restrictions. Examples are listed in "System Constraints on Bit Rate." These systems process information that originated in an analog signal under a bit-rate constraint. Therefore, in these cases, the rate at which the analog input is sampled can be reduced to be as low as the critical sampling rate f_R, without increasing the overall distortion. How low this f_R is, compared to the Nyquist rate or the spectral occupancy of the signal, depends on our assumptions on the source statistics through its PSD. Examples for the dependency between the two are shown in Figure 15. Evidently, reducing the sampling rate allows the saving of other system parameters, such as power and thermal noise resulting from lower clock cycles. Alternatively, this reduction provides a way to sample wide-band signals that cannot be sampled at their Nyquist rate without introducing additional distortion due to sampling, on top of the distortion due to a bit-rate constraint. Next, we explore additional theoretical and practical implications of our ADX scheme.

The lack of computational resources for the extraction of useful information from large data sets is one of the most pressing issues of the digital age.

Sampling infinite bandwidth signals

While a common assumption in signal processing is that for all practical purposes the bandwidth of the source signal is bounded, there are many important cases where this assumption does not hold. These cases include Markov processes, autoregressive processes, and the Wiener process or other semimartingales. An important contribution of the ADX paradigm is in describing the optimal tradeoff among distortion, sampling rate, and bit rate, even if the source signal is bandlimited. This tradeoff is best explained by an example.

Consider a Gaussian stationary process $X_\Omega(t)$ with a PSD of

$$S_\Omega(f) = \frac{1/f_0}{(\pi f/f_0)^2 + 1}, \quad f_0 > 0. \tag{15}$$

The signal $X_\Omega(t)$ is also a Markov process, and it is, in fact, the unique Gaussian stationary process that is also Markovian (also known as the *Ornstein–Uhlenbeck process*). The PSD $S_\Omega(f)$ is shown in Figure 15, along with the relation between the bit rate R and the minimal sampling frequency f_R required to achieve Shannon's DRF of $X_\Omega(t)$. This relation is obtained by evaluating $D(f_s,R)$ for the PSD $S_\Omega(f)$. In fact, the exact equation describing the green curve in Figure 15 can be evaluated in closed form, from which it follows [18] that

$$R = \frac{1}{\ln 2}\left(f_R - f_0 \frac{\arctan(\pi f_R/f_0)}{\pi/2}\right). \tag{16}$$

Notice that, although the Nyquist frequency of the signal in this example is infinite, for any finite R, there exists a critical sampling frequency f_R, satisfying (16), such that Shannon's DRF of $X_\Omega(t)$ can be attained by sampling at or above f_R.

The asymptotic behavior of (16) as R extends to infinity is given by $R \sim (f_R/\ln 2)$. Thus, for R sufficiently large, the optimal sampling rate is linearly proportional to R and, in particular, in the limit of zero distortion when R grows to infinity. The ratio R/f_s is the average number of bits per sample used in the resulting digital representation. It follows from (16) that, asymptotically, the right number of bits per sample converges to $1/\ln 2 \approx 1.45$. If the number of bits per sample is below this value, then the distortion in ADX is dominated by Shannon's DRF of $X_\Omega(t)$, as there are not enough bits to represent the information acquired by the sampler. If the number of bits per sample is greater than this value, then the distortion in ADX is dominated by the sampling distortion, as there are not enough samples for describing the signal up to a distortion equals to its Shannon's DRF.

As a numerical example, assume that we encode $X_\Omega(t)$ using two bits per sample, i.e., $f_s = 2R$. As $R \to \infty$, the ratio between the minimal distortion $D(f_s,R)$ and Shannon's DRF of the signal converges to approximately 1.08, whereas the ratio between $D(f_s,R)$ and $\mathrm{mmse}(f_s)$ converges to approximately 1.48. In other words, it is possible to attain the optimal encoding performance within a gap of approximately 8% by providing one sample per each two bits per unit time used in this encoding. Alternatively, it is possible to attain the optimal sampling performance within a gap of approximately 48% by providing two bits per each sample taken.

Theoretical limits on estimation from sampled and quantized information

The limitation on bit rate in the scenarios mentioned in "System Constraints on Bit Rate" are the result of engineering limitations. However, sampling and quantization

constraints may also be inherent in the system model and the estimation problem. As an example, consider the estimation of an analog signal describing the behavior of the price of a financial asset. Although we assume that the price follows some continuous-time behavior, the value of the asset is only observed whenever a transaction is reported. This limitation on the observation can be described by a sampling constraint. If the transactions occur at nonuniform time lags, then this sampling is nonuniform. Moreover, it is often assumed that the instantaneous change of the price is given by a deterministic signal representing the drift plus an additive infinite bandwidth and stationary noise [9]. Therefore, the signal in question is of infinite bandwidth, and sampling occurs below the Nyquist rate.

In addition to the sampling constraint, it may be the case that the values of the transactions are hidden from us. The only information we receive is through a sequence of actions taken by the agent controlling this asset. Assuming the set of possible actions is finite, this last limitation corresponds to a quantization constraint. Therefore, the MMSE in estimating the continuous-time price based on the sequence of actions is described by the minimal distortion in the ADX.

While, in this case, we have no control over the actual way the samples are encoded (into actions), the minimal distortion in the ADX setting provides a lower bound on the distortion in estimating the continuous-time price. This distortion can be expressed by an additional noise in a model that makes decisions based on the estimated price.

Removing redundancy at the sensing stage

At the end of the section "ADX Via Pulse-Code Modulation," we concluded that, under an optimal encoder, oversampling does not affect the fundamental distortion limit, since the introduced redundancy is removed in the encoding. However, oversampling may still be undesirable to the overall system performance, since it results in redundant data that must be removed by additional processing. In fact, since analog signal processing is not constrained by memory or bit rate, when information originating in an analog signal is converted to a digital form, it may bloat the system's memory with a large amount of redundant data. The processing of these data requires additional resources that are proportional to their size and may severely restrict the system's ability to extract useful information. Indeed, the lack of computational resources for the extraction of useful information from large data sets is one of the most pressing issues of the digital age [64].

One way to address this big data challenge is by collecting only relevant information from the analog world, i.e., attaining a nonredundant digital representation of the analog signal. For example, oversampling in the PCM as in "ADX Via Pulse-Code Modulation" section leads to a redundant digital representation of the quantized samples, since these become more correlated with one another as the sampling rate increases. Indeed, the properties of PCM imply that the optimal sampling rate that minimizes the distortion also maximizes the entropy rate of its digital output.

The counterpart of the PCM redundancy phenomena in the more general setting of ADX is the representation attained by optimal sampling at the critical rate. This optimal sampling can be seen as a mechanism to remove redundancy at the sampling stage. It guarantees that the signal postsampling does not contain any parts that would be removed under an optimal lossy compression.

As an example of a system that benefits from operating according to the previously discussed principle, we envision a real-time voice-to-text transcriber based on an artificial neural network [65]. Such a system consists of an artificial neural network that maps a sequence of bits to words, where this sequence is obtained by an ADX unit, as shown in Figure 21. Since the rate of information per unit time that can be processed by the neural net is limited, an optimal design of the ADX would provide bits into the neural network consistent with this rate. The challenge is, therefore, to sample and encode the audio signal at the rate of the neural network processing so as to provide the most relevant information subject to that rate constraint for the network to perform its classification task. If we assume that the most relevant information is described by a spectral psychoacoustic distortion function, then the optimal ADX scheme with a signal PSD weighted by this distortion function provides the most relevant information for classification, subject to the processing constraint.

Conclusions

The processing, communication, and/or digital storage of an analog signal is achieved by first representing it as a bit sequence. The restriction on the bit rate of this sequence is the result of

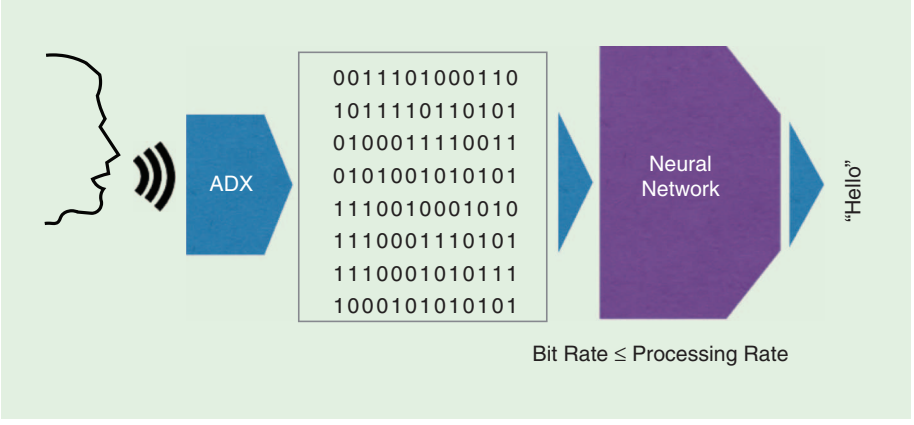


FIGURE 21. The bit rate of a digital representation of the sound of the word hello should not exceed the processing rate of the neural network. Sampling and lossy compression according to ADX preserves the most relevant part of the analog signal with respect to the distortion criterion and subject to the bit-rate constraint.

constraints on power, memory, communication, and computation. In addition, hardware and modeling constraints in processing analog information imply that the digital representation is obtained by first sampling the analog waveform and then quantizing or encoding its samples. That is, the transformation from analog signals to bits involves the composition of sampling and quantization or, more generally, lossy compression operations.

In this article, we explored the minimal sampling rate required to attain the fundamental distortion limit subject to a strict constraint on the bit rate of the system. We concluded that when the energy of the signal is not uniformly distributed over its spectral occupancy, the optimal signal representation can be attained by sampling at a rate lower than the Nyquist rate, which depends on the actual bit-rate constraint. This reduction in the optimal sampling rate under finite bit precision is made possible by designing the sampling mechanism to sample only those parts of the signals that are not discarded because of optimal lossy compression.

The characterization of the fundamental distortion limit and the sampling rate required to attain it has several important implications. Most importantly, it provides an extension of the classical sampling theory of Whittaker, Kotelnikov, Shannon, and Landau, as it describes the minimal sampling rate required for attaining the minimal distortion in sampling an analog signal. It also leads to a theory of representing signals of infinite bandwidth with a vanishing distortion. In particular, it provides the average number of bits per sample, i.e., the ratio of the bit rate (bits per unit of time) and the sampling rate (samples per unit of time) so that, as the number of bits and samples per unit of time extend to infinity, the ratio between the distortion under optimal sampling and encoding and the DRF decreases to one.

Our results also indicate that sampling at the Nyquist rate is not necessary when working under a bit-rate constraint for signals of either finite or infinite bandwidth. Such a constraint may be due to hardware power, cost, or memory limitations. Moreover, sampling a signal at its critical sampling rate associated with a given bit-rate constraint results in the most compact digital representation of the analog signal and thus provides a mechanism to remove redundant information at the sensing stage.

Acknowledgments

This work was supported in part by the National Science Foundation (NSF) under grant CCF-1320628, by the NSF's Center for Science of Information grant CCF-0939370, and by the U.S.-Israel Binational Science Foundation (BSF) under the BSF Transformative Science grant 2010505.

Authors

Alon Kipnis (kipnisal@stanford.edu) received his B.Sc. degree in mathematics (summa cum laude) and his B.Sc. degree in electrical engineering (summa cum laude), both in in 2010, and his M.Sc. degree in mathematics in 2012, all from Ben-Gurion University of the Negev, Israel. He recently received his Ph.D. degree in electrical engineering from Stanford University, California, where he is now a postdoctoral scholar in the Department of Statistics. His research focuses on the intersection of signal processing, machine learning, and statistics with data compression.

Yonina C. Eldar (yonina@ee.technion.ac.il) is a professor in the Department of Electrical Engineering at the Technion–Israel Institute of Technology, Haifa, where she holds the Edwards Chair in engineering. She is also an adjunct professor at Duke University, Durham, North Carolina, and a research affiliate with the Research Laboratory of Electronics at the Massachusetts Institute of Technology, Cambridge, and she was a visiting professor at Stanford University, California. She is a member of the Israel Academy of Sciences and Humanities and of the European Association for Signal Processing. She has received many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award, the IEEE/Aerospace and Electronic Systems Society Fred Nathanson Memorial Radar Award, the IEEE Kiyo Tomiyasu Award, the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, and the Wolf Foundation Krill Prize for Excellence in Scientific Research. She is the editor-in-chief of *Foundations and Trends in Signal Processing* and serves the IEEE on several technical and award committees. She is a Fellow of the IEEE.

Andrea J. Goldsmith (andrea@wsl.stanford.edu) received her B.S. degree in 1986, her M.S. degree in 1989, and her Ph.D. degree in 1994, all in electrical engineering, from the University of California, Berkeley. She is the Stephen Harris professor of electrical engineering at Stanford University, California. She also cofounded and served as chief technology officer of Plume WiFi (formerly Accelera Inc.) and of Quantenna Communications. She is a member of the National Academy of Engineering and the American Academy of Arts and Sciences. She has received several awards for her work, including the IEEE Communications Society Edwin Howard Armstrong Achievement Award and the Silicon Valley Business Journal's Women of Influence Award. She has authored three books on wireless communications and is an inventor with 28 patents. Her research interests are in information and communication theory and their application to wireless communications and related fields.

References

[1] E. T. Whittaker, "On the functions which are represented by the expansions of the interpolation theory," *Proc. Roy. Soc. Edinburgh*, vol. 35, pp. 181–194, July 1915.

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

[3] H. Landau, "Sampling, data transmission, and the Nyquist rate," *Proc. IEEE*, vol. 55, no. 10, pp. 1701–1706, Oct. 1967.

[4] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 379–423, 1948.

[5] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec.*, vol. 4, pp. 142–163, 1959.

[6] T. Berger, *Rate-Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice Hall, 1971.

[7] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[9] I. Karatzas and S. E. Shreve, *Methods of Mathematical Finance*. New York: Springer-Verlag, 1998.

[10] M. W. Mahoney, "Randomized algorithms for matrices and data," *Found. Trends Mach. Learning*, vol. 3, no. 2, pp. 123–224, 2011.

[11] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, 2008.

[12] B. P. Ginsburg, "Energy-efficient analog-to-digital conversion for ultra-wideband radio," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2007.

[13] R. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[14] T. S. Rappaport, R. W. Heath, Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Westford, MA: Pearson Education, 2014.

[15] H. S. Black and J. Edson, "Pulse code modulation," *Trans. Amer. Inst. Elect. Engineers*, vol. 66, no. 1, pp. 895–899, 1947.

[16] B. Oliver, J. Pierce, and C. Shannon, "The philosophy of PCM," *IRE Trans. Inform. Theory*, vol. 36, no. 11, pp. 1324–1331, Nov. 1948.

[17] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, 1998.

[18] A. Kipnis, Y. C. Eldar, and A. J. Goldsmith, "Fundamental distortion limits of analog-to-digital compression," arXiv. [Online]. https://arxiv.org/abs/1601.06421, 2016.

[19] R. Gray, "Quantization noise spectra," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1220–1244, Nov. 1990.

[20] H. Viswanathan and R. Zamir, "On the whiteness of high-resolution quantization errors," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 2029–2038, July 2001.

[21] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 446–472, 1948.

[22] L. Schuchman, "Dither signals and their effect on quantization noise," *IEEE Trans. Commun. Technol.*, vol. 12, no. 4, pp. 162–165, 1964.

[23] S. Shamai, "Information rates by oversampling the sign of a bandlimited process," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1230–1236, 1994.

[24] R. Zamir and M. Feder, "Rate-distortion performance in coding bandlimited sources by sampling and dithered quantization," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 141–154, Jan. 1995.

[25] J. Candy, "A use of limit cycle oscillations to obtain robust analog-to-digital converters," *IEEE Trans. Commun.*, vol. 22, no. 3, pp. 298–305, Mar. 1974.

[26] A. Kipnis, A. J. Goldsmith, and Y. C. Eldar, "Optimal trade-off between sampling rate and quantization precision in sigma-delta A/D conversion," in *Proc. Int. Conf. Sampling Theory and Applications (SampTA)*, May 2015, pp. 627–631.

[27] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Trans. Inf. Theory*, vol. 23, no. 3, pp. 337–343, May 1977.

[28] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Trans. Inf. Theory*, vol. 24, no. 5, pp. 530–536, 1978.

[29] F. M. Willems, Y. M. Shtarkov, and T. J. Tjalkens, "The context-tree weighting method: Basic properties," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 653–664, 1995.

[30] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[31] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.

[32] H. Gish and J. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inf. Theory*, vol. 14, no. 5, pp. 676–683, 1968.

[33] A. Perez, "Extensions of Shannon-McMillan's limit theorem to more general stochastic processes," in *Proc. Trans. 3rd Prague Conf. Information Theory, Statistical Decision Functions and Random Processes*, 1964, pp. 545–574.

[34] T. Berger, "Rate distortion theory for sources with abstract alphabets and memory," *Inform. and Control*, vol. 13, no. 3, pp. 254–273, 1968.

[35] A. Kolmogorov, "On the Shannon theory of information transmission in the case of continuous signals," *IRE Trans. Inform. Theory*, vol. 2, no. 4, pp. 102–108, Dec. 1956.

[36] T. Berger and J. Gibson, "Lossy source coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2693–2723, 1998.

[37] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*. Berlin: Springer-Verlag, 2013.

[38] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Trans. Inform. Theory*, vol. 8, no. 5, pp. 293–304, 1962.

[39] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic, 1997.

[40] B. Bharucha and T. Kadota, "On the representation of continuous parameter processes by a sequence of random variables," *IEEE Trans. Inf. Theory*, vol. 16, no. 2, pp. 139–141, 1970.

[41] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[42] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[43] J. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 406–411, 1970.

[44] H. Witsenhausen, "Indirect rate distortion problems," *IEEE Trans. Inf. Theory*, vol. 26, no. 5, pp. 518–521, 1980.

[45] H. Dym and H. McKean, *Gaussian Processes, Function Theory, and the Inverse Spectral Problem*. New York: Academic, 1976.

[46] M. Matthews, "On the linear minimum-mean-squared-error estimation of an undersampled wide-sense stationary random process," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 272–275, 2000.

[47] T. Michaeli and Y. C. Eldar, "High-rate interpolation of random signals from nonideal samples," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 977–992, 2009.

[48] D. L. Neuhoff and S. S. Pradhan, "Information rates of densely sampled data: Distributed vector quantization and scalar quantization with transforms for Gaussian sources," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5641–5664, 2013.

[49] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, no. 4, pp. 639–697, Apr. 2006.

[50] A. Kipnis, A. J. Goldsmith, and Y. C. Eldar, "The distortion rate function of cyclostationary Gaussian processes," *IEEE Trans. Inf. Theory*, Aug. 2017. doi: 10.1109/TIT.2017.2741978.

[51] A. Kipnis, A. J. Goldsmith, Y. C. Eldar, and T. Weissman, "Distortion rate function of sub-Nyquist sampled Gaussian sources," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 401–429, Jan. 2016.

[52] R. G. Vaughan, N. L. Scott, and D. R. White, "The theory of bandpass sampling," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 1973–1984, 1991.

[53] P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proc. IEEE*, vol. 78, no. 1, pp. 56–93. 1990.

[54] Y. Chen, A. J. Goldsmith, and Y. C. Eldar, "Channel capacity under sub-Nyquist nonuniform sampling," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4739–4756, Aug. 2014.

[55] H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, no. 1, pp. 37–52, 1967.

[56] F. J. Beutler, "Error-free recovery of signals from irregularly spaced samples," *SIAM Rev.*, vol. 8, no. 3, pp. 328–335, 1966.

[57] H. Landau, "A sparse regular sequence of exponentials closed on large sets," *Bull. Amer. Math. Soc.*, vol. 70, no. 4, pp. 566–569, 1964.

[58] H. G. Feichtinger and K. Gröchenig, "Irregular sampling theorems and series expansions of band-limited functions," *J. Math. Anal. Applicat.*, vol. 167, no. 2, pp. 530–556, 1992.

[59] K. Yao and J. Thomas, "On some stability and interpolatory properties of nonuniform sampling expansions," *IEEE Trans. Circuit Theory*, vol. 14, no. 4, pp. 404–408, Dec. 1967.

[60] R. Young, *An Introduction to Nonharmonic Fourier Series*. New York: Academic, 2001.

[61] P. P. Vaidyanathan, "Theory of optimal orthonormal subband coders," *IEEE Trans. Signal Process.*, vol. 4, no. 6, pp. 1528–1543, June 1998.

[62] A. Kipnis, A. J. Goldsmith, and Y. C. Eldar. (2016). The distortion-rate function of sampled Wiener processes. [Online]. Available: http://arxiv.org/abs/1608.04679

[63] A. Kipnis, G. Reeves, Y. C. Eldar, and A. J. Goldsmith, "Compressed sensing under optimal quantization," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Aachen, Germany, 2017, pp. 2148–2152. doi: 10.1109/ISIT.2017.8006909.

[64] M. Hilbert and P. López, "The world's technological capacity to store, communicate, and compute information," *Science*, vol. 332, no. 6025, pp. 60–65, 2011.

[65] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 6645–6649.

SP