Sub-Nyquist Sampling Achieves Optimal Rate-Distortion

Alon Kipnis
Department of Electrical Engineering
Stanford University

Andrea J. Goldsmith
Department of Electrical Engineering
Stanford University

Yonina C. Eldar
Department of Electrical Engineering
Technion, Israel Institute of Technology
Haifa, Israel

***Abstract*—The minimal sampling frequency required to achieve the rate-distortion function of a Gaussian stationary process is analyzed. Although the Nyquist rate is the minimal sampling frequency that allows perfect reconstruction of a bandlimited signal from its samples, relaxing perfect reconstruction to a prescribed distortion may allow a lower sampling frequency to achieve the optimal rate-distortion trade-off. We consider a combined sampling and source coding problem in which an analog Gaussian source is reconstructed from its rate-limited sub-Nyquist samples. We show that each point on the distortion-rate curve of the source corresponds to a sampling frequency f_{DR} smaller than the Nyquist rate, such that this point can be achieved by sampling at frequency f_{DR} or above. This can be seen as an extension of the sampling theorem in the sense that it describes the minimal amount of excess distortion in the reconstruction due to lossy compression of the samples, and provides the minimal sampling frequency required in order to achieve that distortion.**

I. INTRODUCTION

The minimal sampling frequency required for perfect reconstruction of an analog process from its samples is given by the celebrated works of Shannon, Nyquist, Whittaker, Kotelnikov and Landau. However, given a stationary Gaussian process with a known power spectral density (PSD), the sampling theorem does not address the error in reconstruction if the samples are quantized, or more generally, compressed in a lossy manner.

The error due to sampling and lossy compression can be analyzed by considering the combined sampling and source coding problem described in Fig. 1. In this model, if the analog source $X(\cdot)$ is sampled above its Nyquist rate, then the signal can be perfectly reconstructed from its samples. Hence, the optimal trade-off between the source coding rate and distortion is described by the distortion-rate function (DRF) $D_X(R)$ of the analog source. We ask the following question; given a source coding rate constraint R (for example, as a result of using an R bit quantizer), do we still need to sample at the Nyquist rate in order to achieve $D_X(R)$, or is a lower sampling frequency sufficient to describe $X(\cdot)$ up to an average distortion $D_X(R)$? In this work we establish a sampling frequency which is in general lower than the Nyquist frequency, such that sampling at this frequency achieves the distortion-rate bound. That is, for a Gaussian stationary process with known statistics, sampling below the Nyquist rate is possible without additional distortion over that given by Shannon's distortion-rate function associated with Nyquist rate sampling.

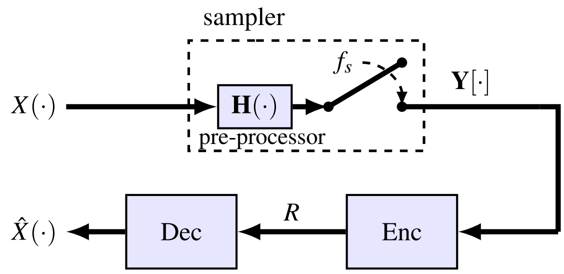

Fig. 1: Combined sampling and source coding model.

A. Background and Related Work

The celebrated Shannon-Nyquist-Kotelnikov-Wittaker sampling theorem asserts that a bandlimited signal can be perfectly reconstructed from its uniform samples at frequency $f_s > f_{Ny}$, where f_{Ny} is twice the bandwidth of the signal. Landau [1] showed that perfect reconstruction is possible iff the sampling frequency f_s exceeds the support of the spectrum of the signal, even if the sampling instants are non-uniformly placed (in which case we replace f_s with the *Beurling density* of the sampling set). For simplicity, in the remainder of this paper we will use the Nyquist rate to refer to the support of the spectrum, which is the minimum sampling frequency required for perfect reconstruction. In the case of a second order stationary random processes under a mean squared error (MSE) criterion, the Nyquist rate condition for perfect reconstruction is given in terms of the power spectral density (PSD) of the process [2]. Perfect reconstruction of signals sampled below their Nyquist rate has proven to be possible under further assumptions on the structure of the analog source, such as bandpass signals with sparse unknown spectral occupancy [3]. In this work we consider sub-Nyquist sampling of Gaussian stationary processes with a known PSD, and relax the requirement for perfect reconstruction to a given average quadratic distortion.

The optimal trade-off between quadratic average distortion and the source coding rate in any description of a Gaussian stationary process $X(\cdot)$ is described by its quadratic DRF which was derived by Pinsker [4]. This was extended by Dubroshin and Tsybakov [5] to the case where the information on the source at the encoder is available only through a different process jointly stationary with $X(\cdot)$, rather than $X(\cdot)$ itself. Such a problem belongs to the regime of *indirect* or *remote* source coding [6, Sec. 4.5.4].

The problems of indirect source coding and sub-Nyquist reconstruction were recently combined in [7], in which the combined sampling and source coding problem of Fig. 1 was considered. The quantities of merit in this problem are the source coding rate R, the average sampling frequency f_s, and the average distortion D. The exact relation among D, R and f_s under uniform filter-bank sampling is the main result of [7]. This relation is given by the indirect distortion-rate function (iDRF) of a Gaussian process given its samples, denoted by $D_{X|Y}(f_s,R)$. One conclusion from this characterization is that for a given sampling frequency f_s, a bound $\underline{D}_X(f_s,R)$ on the iDRF is obtained by reverse water-filling over the part of the spectrum of Lebesgue measure f_s with maximum energy.

In [8] it was shown that the bound $\underline{D}_X(f_s,R)$ cannot be improved by nonuniform sampling or by a wide class of pre-sampling operations, when we replace f_s with the density of the non-uniform sampling set. These results establish $\underline{D}_X(f_s,R)$ as a fundamental quantity in information theory and signal processing, which describes the amount of excess distortion incurred due to encoding based on the information in any time preserving sampling scheme.

B. Contribution

In this work we prove an important property of $\underline{D}_X(f_s,R)$ which asserts that for a given point (R,D) on the DRF of $X(\cdot)$, there exists a frequency f_{DR} such that the equality

$$D_X(R) = \underline{D}_X(f_s,R)$$

is attained for all $f_s \geq f_{DR}$ which is typically less than the Nyquist rate f_{Ny}. This critical frequency f_{DR} depends on R and the PSD $S_X(f)$, and it is strictly smaller than the Nyquist rate if $S_X(f)$ is not constant over its entire bandwidth. In addition, f_{DR} increases monotonically with R and reduces to the Nyquist rate of $X(\cdot)$ as $R\to\infty$. We note that from the definition of $\underline{D}_X(f_s,R)$ it follows that $D_X(R) = \underline{D}_X(f_s,R)$ for any $f_s \geq f_{Ny}$, since then the sampling theorem asserts that there is no information loss in observing $Y[\cdot]$ over $X(\cdot)$. As illustrated in Fig. 2, our main result says that the condition $f_s \geq f_{Ny}$ can be relaxed to $f_s \geq f_{DR}$ if we work under a source coding rate constraint R. In other words, there is no further information loss due to sub-Nyquist sampling compared to the distortion-rate trade-off. This result can be seen as an extension of the Shannon-Nyquist sampling theorem and the characterization of the minimal MSE $mmse(f_s)$ in estimating a signal from its samples at frequency f_s, in the sense that it describes the lowest sampling frequency f_{DR} that achieves $D(R)$.

The rest of this paper is organized as follows: in Section II we review results from [7] and define the function $\underline{D}_X(f_s,R)$. Our main result is described in Section III. Concluding remarks are provided in Section IV.

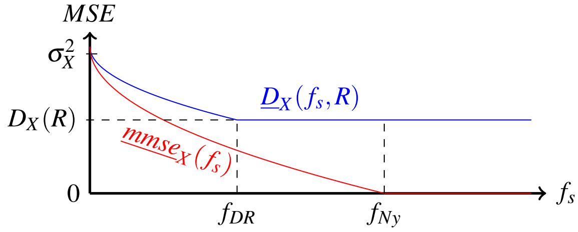

Fig. 2: Minimal distortion versus sampling frequency. If $R<\infty$, then $D_X(R)$ is attained for $f_s \geq f_{DR}$. Here $\underline{mmse}_X(f_s) = \lim_{R\to\infty}\underline{D}(f_s,R)$.

II. Minimal Distortion under a Combined Sampling and Source Coding Model

In this section we describe the combined sampling and source coding problem depicted in Fig. 1, which was considered in [7]. The source $X(\cdot)=\{X(t), t\in\mathbb{R}\}$ is a real Gaussian stationary process with a *known* PSD

$$S_X(f) \triangleq \int_{-\infty}^{\infty} \mathbb{E}\left[X(t+\tau)X(t)\right] e^{-2\pi i\tau f} d\tau,$$

and variance $\sigma_X^2 \triangleq \int_{-\infty}^{\infty} S_X(f)\,df < \infty$. The sampler receives the process $X(\cdot)$ as an input and produces a discrete-time process $Y[\cdot]=\{Y[n], n\in\mathbb{Z}\}$. The specific structure of the sampler will be described in the sequel. The distortion between a source realization $x(\cdot)$ and its reconstruction $\hat{x}(\cdot)$ is given by

$$d(x(\cdot),\hat{x}(\cdot)) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}(x(t)-\hat{x}(t))^2 dt.$$

The indirect distortion-rate function (iDRF) of $X(\cdot)$ given $Y[\cdot]$, denoted by $D_{X|Y}(f_s,R)$, is defined to be the minimal average quadratic distortion:

$$D_{X|Y}(f_s,R) = \inf_{Y\xrightarrow{R}\hat{X}} \mathbb{E}d\left(X(\cdot),\hat{X}\right)$$

where the infimum is taken over all mappings from $Y[\cdot]$ to $\hat{X}(\cdot)$ such that the mutual information rate $I\left(Y[\cdot];\hat{X}(\cdot)\right)$ is limited to R/f_s bits per sample, where f_s is the average sampling frequency defined by the sampler. Although implicit in our notation, $D_{X|Y}(f_s,R)$ depends on the sampling mechanism through the samples $Y[\cdot]$.

In the case where the sampler in Fig.1 is uniform without a pre-processor, i.e., $Y[n]=X(n/f_s)$ for all $n\in\mathbb{Z}$, the iDRF of $X(\cdot)$ is given by ([7])

$$R(\theta) = \frac{1}{2}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}\left[\left(\log\widetilde{S_X}(f)/\theta\right)\right]^+ df, \tag{1a}$$

$$D_{X|Y}(\theta) = \sigma_X^2 - \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}}\left[\widetilde{S_X}(f)-\theta\right]^+ df, \tag{1b}$$

where $[x]^+ = \max\{0,x\}$ and

$$\widetilde{S_X}(f) = \frac{\sum_{k\in\mathbb{Z}} S_X^2(f-f_sk)}{\sum_{k\in\mathbb{Z}} S_X(f-f_sk)}. \tag{2}$$

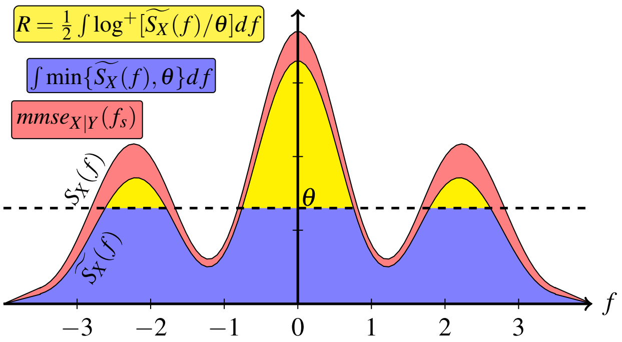

Fig. 3: Water-filling interpretation of (1). The overall distortion is the sum of the red and the blue areas.

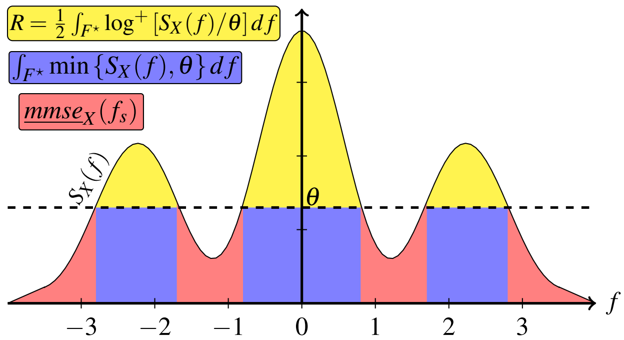

Fig. 5: Water-filling interpretation of (5). The overall distortion is the sum of the red and the blue areas. The set $F^\star(f_s)$ is the section below the blue/yellow area.

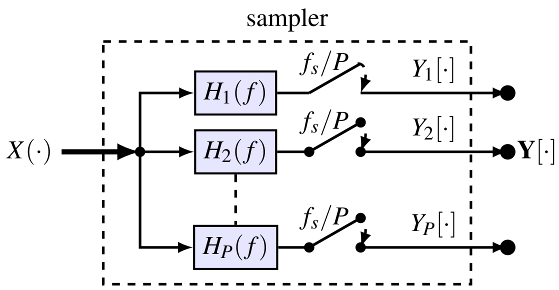

Fig. 4: Uniform filter-bank sampler.

The solution (1) has the water-filling interpretation illustrated in Fig. 3. If the support of $S_X(f)$ is contained in the interval $(-f_s/2, f_s/2)$, then[1] $\widetilde{S_X}(f) = S_X(f)$, and (1) reduces to

$$R_X(\theta) = \frac{1}{2}\int_{-\infty}^{\infty} \log^{+}[S_X(f)/\theta]\, df, \tag{3a}$$

$$D_X(\theta) = \int_{-\infty}^{\infty} \min\{S_X(f),\theta\}\, df, \tag{3b}$$

which is the reverse water-filling solution of Pinsker [4]. Another extreme case of (1) is obtained when $R \to \infty$. In this case $D_{X|Y}(f_s,R)$ reduces to the minimal mean squared error (MMSE) achievable in estimating the source from its samples:

$$mmse_{X|Y}(f_s) = \sigma_X^2 - \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \widetilde{S_X}(f)df, \tag{4}$$

as derived in [9] and [10].

Assume now that the sampler in Fig. 1 is the filter-bank sampler of Fig. 4, which consists of P sampling branches each with a pre-sampling filter $H_p(f)$ and a uniform pointwise sampler of sampling frequency f_s/P. That means that the processes $Y_1[\cdot],\ldots,Y_P[\cdot]$, defined by

$$Y_p[n] \triangleq \int_{-\infty}^{\infty} h_i(nP/f_s - \tau)X(\tau)d\tau,$$

are available at the encoder. For a given PSD $S_X(f)$, a source coding rate R and an average sampling frequency f_s, the distortion can be reduced by an optimal choice of the pre-sampling filters $H_1(f),\ldots,H_P(f)$. The optimal filters are chosen to eliminate aliasing on each sampling branch and at the same time to pass the part of the spectrum with the highest energy, as described in [7]. Denote by $D_X^\star(P,f_s,R)$ the iDRF of $X(\cdot)$ given $Y[\cdot]$ obtained under this optimal choice of $H_1(f),\ldots,H_P(f)$. As the number of sampling branches P goes to infinity, the part of the spectrum that is not blocked by the filters is a set $F^\star(f_s)$ of Lebesgue measure f_s with highest energy. That is, denote

$$\underline{D}_X(f_s,R) \triangleq \lim_{P\to\infty} D_X^\star(P,f_s,R).$$

An expression for $\underline{D}_X(f_s,R)$ is given by

$$R(\theta) = \frac{1}{2}\int_{F^\star(f_s)} \log^{+}[S_X(f)/\theta]\, df, \tag{5a}$$

$$\underline{D}_X(f_s,\theta(R)) = \underline{mmse}_X(f_s) + \int_{F^\star(f_s)} \min\{S_X(f),\theta\}\, df \tag{5b}$$

$$= \sigma_X^2 - \int_{F^\star(f_s)} [S_X(f)-\theta]^{+}\, df,$$

where the set $F^\star(f_s)$ maximizes the integral

$$\int_F S_X(f)df$$

over all measurable sets $F \subset \mathbb{R}$ of Lebesgue measure not exceeding f_s. In (5) we define

$$\underline{mmse}_X(f_s) \triangleq \sigma_X^2 - \int_{F^\star(f_s)} S_X(f)df = \int_{\mathbb{R}\setminus F^\star(f_s)} S_X(f)df,$$

which gives a bound on the minimal MSE in estimating $X(\cdot)$ from it samples under filter-bank sampling with average sampling frequency f_s. Fig. 5 illustrates a water-filling interpretation of (5). The functions $D_X^\star(P,f_s,R)$ and $\underline{D}_X(f_s,R)$ are depicted in Fig. 6.

As discussed in [7], an operative scheme that achieves the distortion-rate bound $\underline{D}_X(f_s,R)$ is as follows:

(i) Filter-bank sampling at average frequency f_s with optimized pre-sampling filters and a sufficient number of sampling branches P.

[1] We interpret expression (2) and similar expressions henceforth as zero if both nominator and denominator are zero.

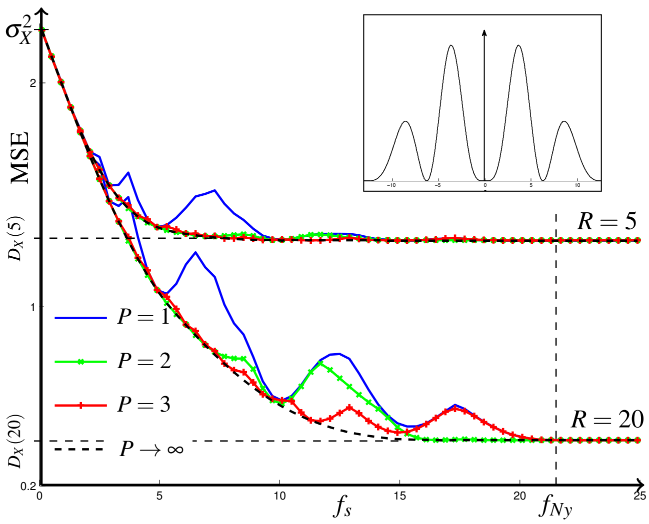

Fig. 6: $D^\star_X(P,f_s,R)$ as a function of f_s for two fixed values of the source coding rate R and source with a PSD given in the small frame. As f_s increases, $D^\star_X(P,f_s,R)$ converges to $D_X(R)$ (with equality guaranteed for $f_s \geq f_{Ny}$). As P increases, $D^\star_X(P,f_s,R)$ converges to $\underline{D}_X(f_s,R)$.

(ii) Minimal MSE estimation of $X(\cdot)$ from $Y[\cdot]$.

(iii) Vector quantizer of source coding rate R that implements the optimal test channel in the Gaussian distortion-rate function [11] of the estimator of $X(\cdot)$ from $Y[\cdot]$.

The function $\underline{D}_X(f_s,R)$ describes an achievable lower bound for the combined sampling and source coding problem of Fig. 1. It was shown in [8] that the bound $\underline{D}_X(f_s,R)$ still holds even if we consider nonuniform sampling where we replace f_s with the Beurling density [12] of the sampling set. As a result of the characterization of time preserving sampling systems in [13], the function $\underline{D}_X(f_s,R)$ is said to describe the amount of information lost under any meaningful practical sampling system.

It follows from (5) that for $f_s > f_{Ny}$, $\underline{mmse}_X(f_s) = 0$ and $\underline{D}_X(f_s,R) = D_X(R)$, where $D_X(R)$ is given by (3). In the next section we will see that this equality also holds for $f_s \geq f_{DR}$, where f_{DR} is usually strictly smaller than the Nyquist rate of $X(\cdot)$.

III. Main Result

Our main result is summarized in the following theorem:

Theorem 3.1: Let $X(\cdot)$ be a Gaussian stationary process with PSD $S_X(f)$. For each point $(R,D) \in [0,\infty) \times (0,\sigma_X^2)$ on the distortion-rate curve of $X(\cdot)$ associated with a water-level θ, let F_θ be the set of frequencies $f \in \mathbb{R}$ such that $S_X(f) > \theta$. Denote by f_{DR} the Lebesgue measure of F_θ. Then for all $f_s \geq f_{DR}$,

$$D_X(R) = \underline{D}_X(f_s,R),$$

where $D_X(R)$ is the distortion-rate function of $X(\cdot)$ and $\underline{D}_X(f_s,R)$ is the optimal indirect distortion-rate function of

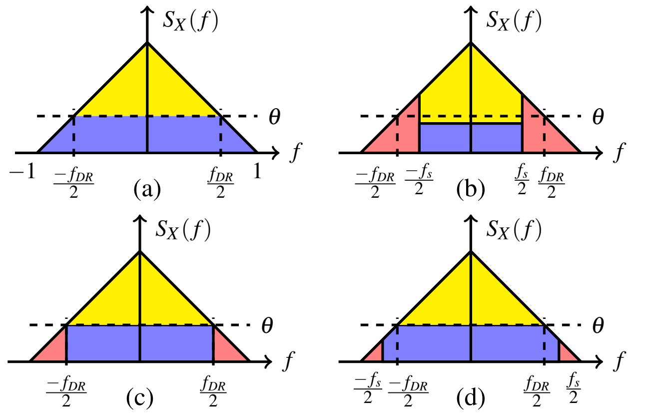

Fig. 7: An illustration of the proof of Theorem 3.1: the distortion is the sum of the red and the blue areas in each figure. (a) Fix R and find $D_X(R)$ from (3). (b) $\underline{D}_X(f_s,R) > D_X(R)$ for $f_s < f_{DR}$. (c) $f_s = f_{DR}$. (d) $\underline{D}_X(f_s,R) = D(R)$ for all $f_s \geq f_{DR}$.

$X(\cdot)$ given its samples at frequency f_s.

Proof sketch: We only provide here a sketch of the proof from which a rigorous proof; a rigorous proof based on this sketch can be found in [14]. In Fig. 7-(a) we fix a point (R,D) on the distortion rate curve of $X(\cdot)$, which is given by (3). The set $F_\theta = \{f \in \mathbb{R} : S_X(f) > \theta\}$ is the support of the yellow area in Fig. 7-(a). The sampling frequency f_{DR} is given by the Lebesgue measure of F_θ. Fig. 7-(b) shows the function $\underline{D}_X(f_s,R)$ for $f_s < f_{DR}$, where the overall distortion is the sum of the $\underline{mmse}_X(f_s)$ term given by the red area, and the water-filling term given by the blue area. Figs. 7 (c) and (d) show the function $\underline{D}_X(f_s,R)$ for $f_s = f_{DR}$ and $f_s > f_{DR}$, respectively. The assertion of Theorem 3.1 is that the sum of the red area and the blue area stays the same for any $f_s \geq f_{DR}$.

From the definition of the function $\underline{D}_X(f_s,R)$ it directly follows that $\underline{D}_X(f_s,R) = D_X(R)$ for all $f_s \geq f_{Ny}$. Theorem 3.1 implies that this equality can be further extended to frequencies below f_{Ny} if the support of F_θ is strictly smaller than the support of $S_X(f)$. If $S_X(f)$ is the rectangular PSD in Fig. 8, then F_θ coincides with the support of $S_X(f)$ and $f_{DR} = f_{Ny}$ for all source coding rates R. If $S_X(f)$ varies over its support, then there exists a region of values of R such that $f_{DR} < f_{Ny}$. Fig. 7-(d) shows that within this region, the error as a result of sampling (red) can be traded with error as a result of the lossy representation of the samples (blue) such that the overall distortion is unaffected. Since optimal lossy compression is achieved by a vector quantizer, Theorem 3.1 implies a trade-off between the vector quantizer resolution and the sampling frequency in schemes that approaches the DRF. A similar trade-off is studied in the case of a scalar quantizer in [14].

As R goes to infinity, $\underline{D}_X(f_s,R)$ converges to $\underline{mmse}_X(f_s)$, the water-level θ goes to zero, the set F_θ coincides with the support of $S_X(f)$ and f_{DR} converges to f_{Ny}. Theorem 3.1 then implies that $\underline{mmse}_X(f_s) = 0$ for all $f_s \geq f_{Ny}$. In summary,

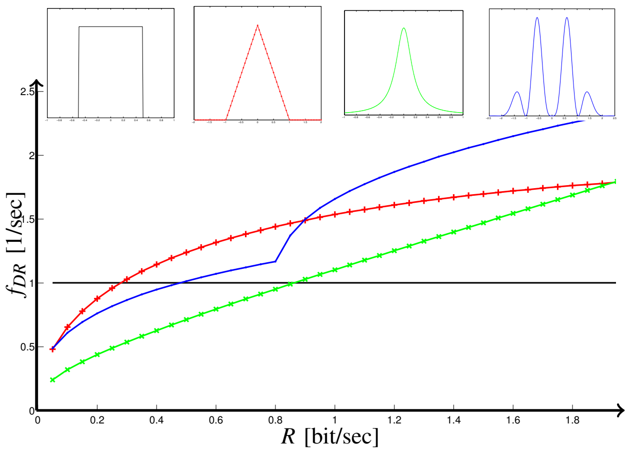

Fig. 8: The frequency f_{DR} as a function of the source coding rate R for the PSDs given in the small frames.

Theorem 3.1 extends the sampling theorem in the sense that it describes the necessary sampling frequency to achieve a prescribed point on the distortion-rate curve.

In the following example the exact dependency of f_{DR} on R and D is found for the triangle PSD of Fig. 7.

Example 1 (triangle PSD): Consider a Gaussian stationary source with PSD

$$S_X(f) = 1 - \min\{|f|, 0\}.$$

Let $(R,D) \in [0,\infty) \times [0,1]$ be a point on the distortion-rate curve of $X(\cdot)$. From Fig. 7 we see that $F_\theta = [-1+\theta, 1-\theta]$ and $f_{DR} = 2-2\theta$. The relation between R to $f_{DR} \leq 2$ is given by

$$R = \frac{1}{2}\int_{-\frac{f_{DR}}{2}}^{\frac{f_{DR}}{2}} \log\left(\frac{1-|f|}{1-\frac{f_{DR}}{2}}\right) df = \frac{1}{\log 2}\left(\log\frac{1}{1-\frac{f_{DR}}{2}} - \frac{f_{DR}}{2}\right). \tag{6}$$

Expressing f_{DR} as a function of D leads to $f_{DR} = \sqrt{1-D}$. Fig. 8 shows f_{DR} as a function of R according to (6) and similar relations for various other PSDs.

IV. Conclusions

We considered a combined sampling and source coding model in which an analog source is described from its rate-limited samples. Under this model, sampling above the Nyquist rate achieves the regular distortion-rate trade-off given by the distortion-rate function of the analog source. We have shown that for a source with a PSD which varies over its support, the same distortion-rate trade-off can be achieved by sampling above a critical frequency which is strictly smaller than the Nyquist rate. This critical frequency depends on the particular point on the distortion-rate curve and the PSD of the source. It increases as the source coding rate increases and converges to the Nyquist rate as the source coding rate goes to infinity.

This implies that with an optimized filter-bank sampler, sampling below the Nyquist rate does not degrade performance in the case where lossy compression of the samples is introduced. Since lossy compression due to quantization is an inherent part of any analog to digital conversion scheme, our work suggests that sampling below the Nyqusit rate is optimal in terms of minimizing distortion in many practical scenarios.

There is still an interesting difference between practical analog to digital schemes and the combined sampling and source coding problems considered here and in [7] which relates to oversampling: While oversampling above the Nyquist rate does not affect the indirect distortion-rate function of a source given its samples, practical analog to digital schemes use oversampling to reduce quantization error [15]. An attempt to bridge this gap is the subject of a future work [14].

Acknowledgment

This work was supported in part by the NSF Center for Science of Information (CSoI) under grant CCF-0939370 and BSF Transformative Science Grant 2010505.

References

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

[2] A. Balakrishnan, "A note on the sampling principle for continuous signals," *IRE Trans. Inform. Theory*, vol. 3, no. 2, pp. 143–146, June 1957.

[3] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 375–391, April 2010.

[4] M. S. Pinsker, "Computation of the message rate of a stationary random process and the capacity of a stationary channel," *Dokl. Akad. Nauk. USSR*, vol. 111, no. 4, 1956.

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

[6] T. Berger, *Rate-Distortion Theory*. Wiley Online Library, 1971.

[7] A. Kipnis, A. J. Goldsmith, T. Weissman, and Y. C. Eldar, "Distortion-rate function of sub-Nyquist sampled gaussian sources," 2014, submitted for publication. [Online]. Available: http://arxiv.org/abs/1405.5329

[8] A. Kipnis, A. Goldsmith, and Y. C. Eldar, "Distortion-rate function under sub-Nyquist nonuniform sampling," in *52st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2014.

[9] D. Chan and R. Donaldson, "Optimum pre-and postfiltering of sampled signals with application to pulse modulation and data compression systems," *IEEE Trans. Commun. Technol.*, vol. 19, no. 2, pp. 141–157, April 1971.

[10] 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.

[11] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*, ser. Kluwer international series in engineering and computer science: Communications and information theory. Springer US, 1992.

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

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

[14] A. Kipnis, A. J. Goldsmith, and Y. C. Eldar, "Optimal trade-off between sampling frequency and quantization precision in A/D conversion," in preparation.

[15] J. de la Rosa, "Sigma-delta modulators: Tutorial overview, design guide, and state-of-the-art survey," *IEEE Trans. Circuits Syst.*, vol. 58, no. 1, pp. 1–21, Jan 2011.