Distortion-Rate Function for Undersampled Gaussian Processes

Iñaki Estella Aguerri[1], Deniz Gündüz[2], Andrea J. Goldsmith[3] and Yonina Eldar[4]

[1]Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain
[2]Department of Electrical Engineering, Imperial College London, London, UK
[3]Department of Electrical Engineering, Stanford University, CA, EEUU
[4]Department of EE Technion, Israel Institute of Technology, Haifa, Israel

***Abstract*—"THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD" We study the fundamental limits in acquisition and transmission of a stationary Gaussian continuous-time process corrupted by noise. The analog signal is digitized into discrete samples with a general analog to digital (A/D) converter that prefilters the continuous process followed by a pointwise sampler. We assume that the sampler, which is not necessarily uniform, is constrained to some fixed sampling rate. The samples are then compressed and transmitted at rate R such that the distortion between the original source and its reconstruction at the receiver is minimized. We first model this problem as a remote source coding problem and characterize the remote distortion-rate function for a fixed A/D converter. We then find the minimum distortion for some fixed sampling structures. We show that uniform sampling is suboptimal in general, and multibranch sampling achieves strictly lower distortion values. Finally, we show that if the sampling rate is sufficiently high, then multibranch sampling achieves the lower bound on the distortion obtained by assuming that the noisy process is directly available at the encoder.**

I. Introduction

Sensing devices take samples of analog signals and digitize them by an analog to digital (A/D) converter. While only discrete samples of the continuous process are available, the Nyquist-Shannon theorem claims that, in the absence of noise, a bandlimited analog process can be perfectly represented from its uniform samples as long as the sampling rate is twice the bandwidth of the process (i.e., the Nyquist rate).

Although most practical and theoretical approaches assume an analog signal acquired at a rate larger than or equal to the Nyquist rate, sampling at this rate might be excessive. In fact, Landau showed that the minimum sampling rate required to perfectly reconstruct a signal is the Lebesgue measure of its spectral support when this support is known, which is below the Nyquist rate in multiband signals [1]. Recently, sub-Nyquist sampling has received a lot of attention in the literature. Many works have shown the possibility of perfect recovery of certain class of signals with sub-Nyquist sampling by exploiting the structure of the signal, [2], [3], or finite rate of innovation [4].

In sensor networks, measurements are sent to a fusion center through a rate-limited channel. Before transmitting, lossy compression is applied to the sensor samples to adapt the information to the available rate. In such a scenario perfect recovery of the analog process at the destination is not possible no matter how high the sampling rate is since the compression process introduces distortion. Hence, relevant question is whether there are better ways to acquire less information through fewer samples rather than by having more samples and then compressing them in a lossy manner.

Moreover, sampling capabilities of an A/D converter might be limited by the technology or the available hardware. In such cases, it it important to evaluate the effect of sub-Nyquist sampling on the end-to-end distortion. The effects of over-sampling are studied in [5], while the capacity for sub-Nyquist sampled channels is considered in [6] and [7].

We study the acquisition of a time continuous Gaussian process corrupted by additive noise with an A/D converter consisting of an analog filter followed by a generalized point-wise sampler with a given sampling rate. We study the loss in performance for different A/D converters in terms of the distortion-rate performance. While more general A/D converter structures can be considered, the proposed scheme is very general and includes most practical A/D architectures in the literature, such as uniform and non-uniform samplers, periodic sampling or filterbank samplers.

For a given A/D converter, the problem under consideration is related to the remote source coding problem. In this class of problems, the source to be compressed is not directly observable, but is available at the encoder through a transformation. In our setting, this transformation corresponds to the combination of the additive noise and the A/D conversion, that map the analog signal to a digital sequence of samples. Unlike in usual source coding problems, the encoder is interested in compressing its noisy observations to minimize the distortion between the source, which is not available at the encoder, and the reconstruction. The remote source coding problem was first studied in a joint source-channel setting in [8] and [9] for finite alphabet sources and for Gaussian time-continuous processes under linear time invariant transforms. These results are extended to discrete transformations with memory in [10]. These works show that under quadratic distortion error, the distortion criterion can be decomposed into two terms using the minimum mean squared error (MMSE) estimator.

The main results of this paper are the following. First, we extend the results in [10] and characterize the remote distortion-rate function for stationary Gaussian processes for a given A/D structure and sampling rate constraint. In doing so, and show that MMSE estimation followed by standard optimal source compression on the estimated input signal is optimal. Then, we derive the optimal analog filters for multibranch uniform sampling, a generalization of single-branch uniform sampling. Finally, we show that if the sampling rate is sufficiently high (yet, still below the Nyquist rate), then multibranch sampling with a different sampling rate at each branch is optimal, that is, it achieves the minimum distortion achievable in the absence of an A/D converter.

II. Problem Statement

A. Sampling and time-preserving transforms

We consider sampling schemes that include both uniform and non-uniform sampling. The notion of sampling rate is generalized for non-uniform sampling schemes using the Beurling density, that quantifies the concentration of samples taken per time unit. For a sampling set $\Lambda \triangleq \{t_n | n \in \mathbb{Z}\}$, such that if $y(t)$ is sampled with Λ, the k-th sample is given by $y(t_k)$, we define the upper and lower Beurling density as

$$B^+(\Lambda) = \lim_{r\to\infty} \sup_{z\in\mathbb{R}} r^{-1}\text{Card}(\Lambda \bigcap [z, z+r]), \quad (1)$$

$$B^-(\Lambda) = \lim_{r\to\infty} \inf_{z\in\mathbb{R}} r^{-1}\text{Card}(\Lambda \bigcap [z, z+r]). \quad (2)$$

Here $\text{Card}(\cdot)$ is the cardinality of the set in the argument. If $B^+(\Lambda) = B^-(\Lambda)$, we say that the sampling set has Beurling density $B(\Lambda) \triangleq B^-(\Lambda)$. Note that for a uniform sampling scheme of sampling period T_s, the Beurling density coincides with the usual definition of sampling rate, i.e. $B(\Lambda) = T_s^{-1}$. In this paper, we define sampling rate as Beurling density.

In order to provide proper sampling rate definition we use the time-preserving transforms notion from [7]. These transforms correspond to the class of functions that do not modify the scale of the input. This class includes most of the blocks in communication systems such as linear time invariant (LTI) systems, filter-banks as well as modulation operations.

B. System model

The system model, whereby a Gaussian process $x(t)$ corrupted by noise is acquired through an A/D converter consisting of a time-preserving analog processor followed by generalized pointwise sampling at a given sampling rate, is illustrated in Fig. IV.

We assume $x(t)$ to be a zero-mean stationary Gaussian process with integrable spectral density $S_x(f)$ for which $\sigma^2 \triangleq \int S_x(f)df$. The process is corrupted by an analog additive noise signal as $y(t) = x(t) + n(t)$, where $n(t)$ is modeled as a zero-mean Gaussian noise with spectral density $S_n(f)$, which we allow to be colored or white.

Before sampling, the noisy process $y(t)$ undergoes a time-preserving transform, that we model by a time-varying linear filter $\mathcal{T}\{\cdot\}$ characterized by its impulse response $q(s,t)$ which outputs $\nu(t) = \mathcal{T}\{y(t)\} = \int_{-\infty}^{\infty} y(s)q(s,t)ds$. Then, $\nu(t)$ is sampled with a pointwise sampler Λ of fixed Beurling density $B(\Lambda) = f_s$. We define $\mathcal{P}$ as the A/D converter $(\Lambda, q(s,t))$ and we denote the samples acquired by $\mathcal{P}$ as $\boldsymbol{\nu}^{\mathcal{P}} \triangleq \{\nu(t_k) : t_k \in \mathbb{Z}\}$, where $\nu(t_k)$ is the k-th sample. Samples $\boldsymbol{\nu}^{\mathcal{P}}$ are available at the input of an encoder f, that transmits them over an error-free channel at a rate R bits per second (bps). The receiver reconstructs the original process as

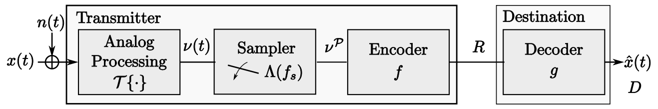

Fig. 1. Source Coding problem of analog process with Constrained Sampling

$\hat{x}(t)$ with a mapping g. We are interested in characterizing the minimum squared error distortion of the reconstruction. For time-continuous processes, the distortion is commonly found as a limiting process of reconstructing the time-truncated versions of the original signal, $x_T(t) \triangleq \{x(t) : |t| \leq T\}$ [10]. Using $||x(t)||_T^2 \triangleq E[\frac{1}{2T}\int_{-T}^{T} x^2(t)dt]$, we have

$$D^{\mathcal{P}} \triangleq \lim_{T\to\infty} D_T^{\mathcal{P}} = \lim_{T\to\infty} ||x_T(t) - \hat{x}_T(t)||_T^2. \quad (3)$$

This system model includes most of the practical acquisition schemes: multibranch sampling, non-uniform sampling, periodic sampling and random sampling, among others. In fact, any time-preserving multibranch system, with any number of branches and non-uniform sampling, is equivalent to a single branch system with some possibly non-uniform sampling sequence that preserves information (see [7, Fact 1]).

We investigate three fundamental problems: first, the characterization of the minimum distortion achievable for a given A/D converter $\mathcal{P}$ with sampling Beurling density f_s. Second, the optimal $\mathcal{P}$ minimizing the distortion over the class of A/D converters $\mathcal{P}$ with Beurling density f_s, and, third, determining similar to the Nyquist theorem, conditions sufficient to reconstruct the analog signal from its samples as if the analog process would be directly accessible without discretization.

III. Remote Gaussian Sources

A lower bound on the minimum distortion achievable for any A/D converter $\mathcal{P}$ of sampling rate f_s is given by assuming that $y(t)$ is directly available at the encoder, and hence, that no A/D sampling is applied. The remote distortion-rate function of a Gaussian process with noise in the absence of A/D conversion is considered in [10]. The minimum achievable distortion D_{lb} for a given rate source encoder R is given in the next theorem.

Theorem 1 ([10], p.129). *The remote distortion-rate function when $y(t)$ is accessible without sampling has the parametric representation in terms of $0 \leq \theta \leq \sup S_x(f)$ as*

$$D_{lb}(R) = \int_{-\infty}^{\infty} S_x(f) - S_{x|y}(f) + \min\{\theta, S_{x|y}(f)\}df, \quad (4)$$

$$\text{s.t. } R = \frac{1}{2}\int_{-\infty}^{\infty} \log^+(S_{x|y}(f)\theta^{-1})df,$$

where $S_{x|y}(f) \triangleq \frac{S_x^2(f)}{S_x(f)+S_n(f)}$ is the spectral density of the $E[x(t)|y(t)]$ MMSE estimator of $x(t)$ given $y(t)$.

The optimal source coding scheme estimates $x(t)$ from $y(t)$ with an MMSE estimator, and applies reverse water-filling over the spectrum of the reconstruction, $S_{x|y}(f)$, i.e., for θ satisfying the constraint, each frequency component satisfying $S_{x|y}(f) \geq \theta$ is transmitted while the other frequency components are discarded.

IV. Rate Distortion Function with Sampling

In this section we return to the system model of Fig. to characterize the minimum achievable distortion for a given A/D scheme $\mathcal{P}$ with a sampling rate f_s in terms of the remote distortion-rate function, i.e., we find f, g minimizing $D^{\mathcal{P}}$ for given $\mathcal{P}$. Results in [8] and [9] cannot be applied to solve this problem since $\boldsymbol{\nu}^{\mathcal{P}}$ is not a time-finite input nor is it jointly stationary with $x(t)$. We show that for stationary Gaussian processes, there is no loss in optimality in separating the acquisition and transmission into three steps: sample the noisy source, reconstruct the original process with an MMSE estimator, and apply standard optimal source coding.

For a fixed T, distortion $D_T^{\mathcal{P}}$ in (3) can be decomposed into two terms using the MMSE estimator of $x_T(t)$ given the encoder input $\boldsymbol{\nu}^{\mathcal{P}}$, $u_T^{\mathcal{P}}(t) = E[x_T(t)|\boldsymbol{\nu}^{\mathcal{P}}]$. Following [9] and using the orthogonality principle of the MMSE estimator, that is $E[(x_T(t) - u_T^{\mathcal{P}}(t))(u_T^{\mathcal{P}}(t) - \hat{x}_T(t))] = 0$, we have,

$$D_T^{\mathcal{P}} = ||x_T(t) - u_T^{\mathcal{P}}(t)||_T^2 + ||u_T^{\mathcal{P}}(t) - \hat{x}_T(t)||_T^2. \quad (5)$$

The first term, which we denote by D_{T1}, corresponds to the distortion in reconstructing $x_T(t)$ from the samples $\boldsymbol{\nu}^{\mathcal{P}}$ using MMSE estimation, and depends only on $\mathcal{P}$, while the second term in (5) denoted by D_{T2} corresponds to the error between $u_T^{\mathcal{P}}(t)$ and $\hat{x}_T(t)$, and hence, depends on $\mathcal{P}$ and on f and g.

Process $x_T(t)$ can be represented by the Karhunen-Loève (KL) expansion as $x_T(t) = \sum_{i=1}^{\infty} \alpha_i \varphi_i(t)$, $|t| \le T$, where $\varphi_i(t)$ are the orthonormal eigenfunctions of the integral equation $\int_{-T}^{T} \phi_x(s,t)\varphi(s)ds = \lambda\varphi(t)$, $|t| \le T$, and where $\phi_x(s,t) \triangleq E[x(s)x(t)]$ is the covariance function of $x(t)$. The expansion coefficients are given by $\alpha_i \triangleq \int_{-T}^{T} x(t)\varphi_i(t)dt$, and the α_i are uncorrelated with each other. If $x(t)$ is Gaussian, the expansion coefficients α_i, $i = 1, 2, ...$, are independent random variables given by $\alpha_i \sim \mathcal{N}(0, \lambda_i)$ where λ_i is the eigenvalue associated to $\varphi_i(t)$. Similarly, $u_T^{\mathcal{P}}(t)$ can be represented in terms of the KL expansion over the eigenfunctions $\varphi_i'(t)$ of its covariance function $\phi_u^{\mathcal{P}}(s,t)$, as $u_T^{\mathcal{P}}(t) = \sum_{i=1}^{\infty} u_i \varphi_i'(t)$, $|t| \le T$, with independent zero mean Gaussian coefficients $u_i \sim \mathcal{N}(0, \lambda_{ui}^{\mathcal{P}})$, where $\lambda_{ui}^{\mathcal{P}}$ is the i-th eigenvalue of the KL expansion. Then, the first term in (5) can be expressed as

$$D_{T1} \stackrel{(a)}{=} ||x_T(t)||_T^2 - ||u_T^{\mathcal{P}}(t)||_T^2 \stackrel{(b)}{=} \frac{1}{2T}\sum_{i=1}^{\infty}\left(\lambda_{\alpha i} - \lambda_{ui}^{\mathcal{P}}\right), (6)$$

where (a) is due to the orthogonality principle; (b) is obtained by the orthonormality of $\phi_i(t)$ and $\phi_i'(t)$ in the KL expansion of $x_T(t)$ and $u_T^{\mathcal{P}}(t)$ respectively.

Since $x_T(t)$ and $n(t)$ are Gaussian processes and $\nu(t)$ is a linear transform of $y(t)$, $\nu(t)$ is a Gaussian process and each sample $\nu(t_k)$ is a Gaussian random variable. Then, it can be shown that $u_T^{\mathcal{P}}(t)$ is a sufficient statistics for $x_T(t)$ given the samples at the encoder input $\boldsymbol{\nu}^{\mathcal{P}}$, i.e., that $p(x_T(t)|\boldsymbol{\nu}^{\mathcal{P}})$ is a function only of $u_T^{\mathcal{P}}(t)$. As a consequence, encoder f can be decomposed without loss of optimality into two blocks: a block that computes the MMSE estimation $u_T^{\mathcal{P}}(t)$ concatenated with a second encoder f', that only has $u_T^{\mathcal{P}}(t)$ as the input.

Now, for given Λ and $q(s,t)$, D_{T1} is fixed and independent of f' and g. Hence, $D_T^{\mathcal{P}}$ is minimized by finding the encoder decoder pair (f', g) minimizing D_{T2}. The minimum distortion when the encoder input $u_T^{\mathcal{P}}(t)$ is a continuous process in a Hilbert space is found in [11]. Expanding $u_T^{\mathcal{P}}(t)$ with the KL expansion the minimum $D_T^{\mathcal{P}}$ is found in terms of θ satisfying

$$D_T^{\mathcal{P}} = D_{T1} + \frac{1}{2T}\sum_{i=1}^{\infty}\min\{\theta, \lambda_{ui}^{\mathcal{P}}\}, \quad (7)$$

$$R_{an}(D_T) = \frac{1}{2T}\sum_{i=1}^{\infty}\frac{1}{2}\log^+(\lambda_{ui}^{\mathcal{P}}\theta^{-1}). \quad (8)$$

Substituting (6) in (7), we have the following lemma.

Lemma 1. *For a given $\mathcal{P}$ of sampling rate f_s, the distortion-rate function with sampling is achieved by MMSE estimation of $x_T(t)$ given the encoder input $\boldsymbol{\nu}^{\mathcal{P}}$, $u_T^{\mathcal{P}}(t) \triangleq E[x_T(t)|\boldsymbol{\nu}^{\mathcal{P}}]$, concatenated optimal source coding for $u_T^{\mathcal{P}}(t)$ at rate R; and is given by $D^{\mathcal{P}}(R, f_s) = \lim_{T\to\infty} D_T^{\mathcal{P}}(R_T, f_s)$, such that*

$$D_T^{\mathcal{P}}(R_T, f_s) = \frac{1}{2T}\sum_{i=1}^{\infty}\lambda_{\alpha i} - \frac{1}{2T}\sum_{i=1}^{\infty}[\lambda_{ui}^{\mathcal{P}} - \theta]^+, \quad (9)$$

$$\text{s.t. } R_T = \frac{1}{2T}\sum_{i=1}^{\infty}\frac{1}{2}\log^+(\lambda_{ui}^{\mathcal{P}}\theta^{-1}),$$

where $\lambda_{\alpha i}$, $\lambda_{ui}^{\mathcal{P}}$, $i = 1, 2, ...$, are the eigenvalues of $q_x(t,s) \triangleq E[x_T(t)x_T(s)]$ and $q_u^{\mathcal{P}}(t,s) \triangleq E[u_T^{\mathcal{P}}(t)u_T^{\mathcal{P}}(s)]$ respectively.

When $u_T^{\mathcal{P}}(t)$ is stationary with spectral density $S_u^{\mathcal{P}}(t)$, the modified Szergo's theorem in [10, Theorem 4.5.4] can be used to show that when $T \to \infty$, the distortion-rate function with sampling converges to

$$D^{\mathcal{P}}(R, f_s) = \int_{\infty}^{\infty} S_x(f)df - \int_{\infty}^{\infty}[S_u^{\mathcal{P}}(f) - \theta]^+ df, \quad (10)$$

$$\text{s.t. } R = \int_{\infty}^{\infty}\frac{1}{2}\log^+(S_u^{\mathcal{P}}(f)\theta^{-1}).$$

V. Achievability

In this section, we use the remote distortion-rate function derived in Section IV to characterize the minimum achievable distortion for two given multibranch sampling schemes.

A. Multibranch LTI filtering and uniform sampling: MUS

Commonly, $y(t)$ is filtered with a LTI filter to reduce aliasing and noise, uniformly sampled at the Nyquist rate and the samples are compressed and transmitted to the destination. We consider a generalization of this scheme in which $y(t)$ is processed by a filter bank of M branches and uniformly sampled at each branch. This multi-branch uniform sampling (MUS) is illustrated in Fig.2.

Before sampling, the noisy signal is processed with a bank of M LTI filters given by $\mathbf{q(t)} \triangleq [q_1(t), ..., q_M(t)]$. Denote by $\nu_k(t)$, $k = 1, ..., M$, the output of the filter at each branch. Let us define the vector of signal outputs as $\boldsymbol{\nu}(t) \triangleq [\nu_1(t), ..., \nu_M(t)]$. Then, each filter output is uniformly sampled every $T_s^M = MT_s$ at each branch, i.e. at a sampling rate $f_s^M = \frac{f_s}{M}$ (such that $\sum_{i=1}^{M} f_s^M = f_s$). Let $\boldsymbol{\nu}_k = \{\nu_k(nT_s^M)\}_{n\in\mathbb{Z}}$ for $k = 1, ..., M$ be the output samples at each branch. The set of samples available at the encoder,

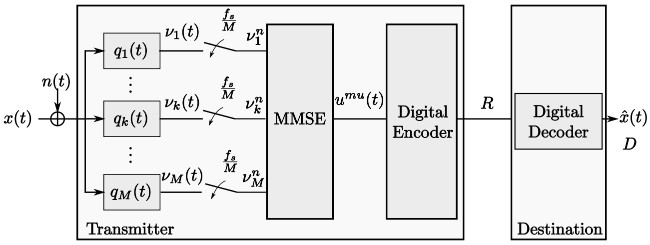

Fig. 2. Multibranch uniform sampling+compressor scheme.

$\boldsymbol{\nu}^{mu} \triangleq \{\boldsymbol{\nu}_1, ..., \boldsymbol{\nu}_M\}$ are used to estimates $x(t)$ with an MMSE estimator, $u^{mu}(t) = E[x(t)|\boldsymbol{\nu}^{mu}]$, which is obtained similarly to [6], and has a spectral density $S_u^{mu}(f)$ given by

$$S_u^{mu}(f) = S_x(f)\mathbf{Q}^H(f)\mathbf{K}^{-1}(f)\mathbf{Q}(f)S_x(f), \qquad (11)$$

where $\mathbf{Q}(f) \triangleq \mathcal{F}\{\mathbf{q}(t)\}$ is the Fourier transform of the filters, and $\mathbf{K}(f) \triangleq \sum_{l=-\infty}^{\infty} \mathbf{S}_\nu(f - l\tilde{f}_s)$ is the aliased expansion matrix of the spectrum of the samples given by

$$\mathbf{S}_\nu(f) = (S_x(f) + S_n(f))\mathbf{Q}(f)\mathbf{Q}^H(f). \qquad (12)$$

Then, $u^{mu}(t)$ is compressed using the optimal source code. Since $u^{mu}(t)$ is stationary the distortion-rate for multibranch sampling for any given bank of filters $\mathbf{q}(t)$, which we denote by $D_{mu}(R, \mathbf{q}(t))$, is obtained by substituting $S_u^{mu}(f)$ in (10).

Since $D_{mu}(R, \mathbf{q}(t))$ characterizes the minimum distortion for fixed $\mathcal{P}$, filters $\mathbf{q}(\mathbf{t})$ can be optimized to jointly minimize the distortion in both the MMSE step and in the compression step. In general, the solution to the filters depends on the water level θ in (10). Fortunately, the set of optimal filters are independent of the value of θ and a general solution can be given, as shown in the next lemma.

Lemma 2. *The minimum distortion achievable with MUS with M branches and sampling rate $f_s^M = \frac{f_s}{M}$ is given by*

$$D_M^{mu}(R, f_s) = \sigma^2 - \int_{-\frac{f_s}{2}}^{-\frac{f_s}{2}} \sum_{i=1}^{M} [\lambda_i(\mathbf{S}_{x|y}(f)) - \theta]^+ df, \qquad (13)$$

$$\text{s.t. } R = \int_{-\frac{f_s}{2}}^{-\frac{f_s}{2}} \frac{1}{2} \sum_{i=1}^{M} \log^+ \left(\lambda_i(\mathbf{S}_{x|y}(f))\theta^{-1}\right) df, \qquad (14)$$

where we define the infinite diagonal matrix $[\mathbf{S}_{x|y}(f)]_{l,l} = S_{x|y}(f - lf_s^M)$, $l \in \mathbb{Z}$; and $S_{x|y}(f)$ is the spectral density in Theorem 1. The distortion-rate function is minimized by the filters $q_k(t) = \mathcal{F}^{-1}\{Q_k(f)\}$, $k = 1, ..., M$, satisfying

$$Q_k(f - lf_s^M) = \begin{cases} 1, & \text{if } S_{x|y}(f - lf_s^M) = \lambda_k(\mathbf{S}_{x|y}(f)), \\ 0 & \text{otherwise}. \end{cases} \qquad (15)$$

This is equivalent to performing reverse water-filling over the M largest components of the aliased spectrum $S_{x|y}(f)$. Interestingly, these optimal filters coincide with the filters minimizing the distortion between $x(t)$ and $u^{mb}(t)$ [6].

B. Multibranch LTI filtering and nonuniform sampling: MNUS

We now propose a multibranch nonuniform sampling (MNUS) scheme that uses uniform sampling at different sampling rates at each branch, determines the spectrum support with Lebesgue measure f_s containing the largest spectral components of $\mathcal{S}_{x|y}(f)$, recovers $S_y(f)$ for this set, and obtains $S_{x|y}(f)$ with MMSE estimation, and finally applies reverse water-filling over it.

Let us define the set of frequencies $\mathcal{S}(f_s)$ as the frequency set containing the larger components of $\mathcal{S}_{x|y}(f)$ of measure $\mu(\mathcal{S}(f_s)) = f_s$, given by

$$\mathcal{S}(f_s) \triangleq \left\{ f : \int_{f\in\mathcal{S}(f)} S_{x|y}(f)df = \sup_{B:\mu(B)=f_s} \int_{f\in B} S_{x|y}(f)df \right\}.$$

For a given spectrum $S_{x|y}(f)$, the encoder calculates the reverse-water filling threshold $\gamma(f_s) \triangleq \inf \mathcal{S}(f_s)$. Then, set $\mathcal{S}(f_s)$ can be expressed as a union of non-overlapping frequency bands $\mathcal{S}(f_s) = \bigcup_i \mathcal{S}_{i\in\mathcal{X}}$ such that $S_i \triangleq \{f : |f| \in [a_i, b_i]\}$ is each of the non-overlapping frequency sets for which $S_{x|y}(f) \geq \gamma(f_s)$ and $\mathcal{X}$ is a countable set indexing S_i. Note that the number of intervals S_i depends on $\gamma(f_s)$.

Let $z_i(t)$ be a Gaussian process with spectrum $S_{zi}(f) = S_y(f)$ if $f \in S_i$ and $S_{zi}(f) = 0$ otherwise. By the Nyquist theorem, each $z_i(t)$ can be perfectly reconstructed from $y(t)$ by extracting the frequency components S_i with an ideal bandlimited filter and then uniformly sampling the output at rate $f_{s,i} = 2(b_i - a_i)$ at each branch. The number of branches is given by the cardinality of $\mathcal{X}$ and a countable number of branches might be required if $\mathcal{X}$ is not finite. From the definition of $\gamma(f_s)$, we have $\sum_{\mathcal{X}} f_{s,i} = f_s$. Then, MMSE estimation is applied as $u_T^{mn}(t) = E[x(t)|z_i(t), i \in \mathcal{X}]$. The spectral density of the reconstruction is given by $S_u^{mn}(f) = S_{x|y}(f)$ if $f \in \mathcal{S}(f_s)$ and $S_u^{mn}(f) = 0$ if $f \notin \mathcal{S}(f_s)$. Then $u_T^{mn}(t)$ is transmitted using the optimal source code. The distortion-rate function for MNUS is given next.

Lemma 3. *The distortion-rate function for MNUS is given by*

$$D^{mn}(R, f_s) = \int_{-\infty}^{\infty} S_x(f)df - \int_{\mathcal{S}(f_s)} [S_x(f) - \theta]^+ df,$$

$$\text{s.t. } R = \frac{1}{2} \int_{\mathcal{S}(f_s)} \log^+ \left(S_{x|y}(f)\theta^{-1}\right) df.$$

VI. NUMERICAL RESULTS

In this section we compare the performance of the proposed MUS and MNUS schemes with the lower bound on the distortion. We consider the acquisition and transmission of a stationary Gaussian process with spectrum $S_x(f)$ with $\sigma^2 = 1$, bandlimited to $|f| \leq$ 20MHz and corrupted by colored Gaussian noise with spectrum $S_n(f)$, as shown in Fig. 3. The spectrum of $y(t)$, $S_y(f) = S_x(f) + S_n(f)$ is also shown.

In Fig. 3 we show the minimum achievable distortion as a function of the available sampling rate f_s for a fixed compression rate of $R =$ 30bps. The minimum distortion for single branch uniform sampling is given by Lemma 2 with $M = 1$. MNUS for single-branch sampling and the lower bound $D_{lb}(R)$ from Theorem 1 are also shown in this figure. It can be observed that for small sampling rates f_s, both schemes exhibit distortion far from the lower bound and they perform closer to it as f_s increases. Moreover, we see that single branch uniform sampling is outperformed by MNUS although their performance gets very close for some f_s. It is

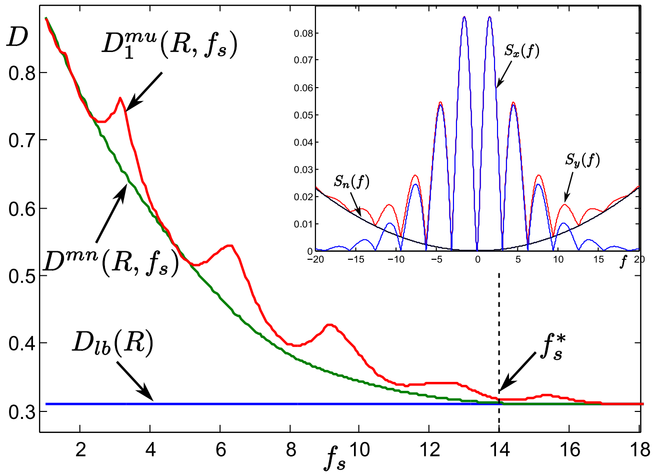

Fig. 3. Upper and lower bounds on the distortion versus sampling rate f_s and rate $R = 30$bps. The upper right corner shows the spectrum of the transmit signal and the noise.

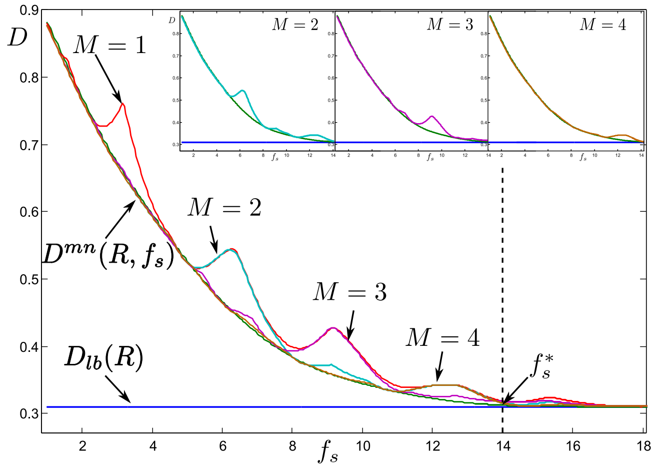

Fig. 4. Upper and lower bounds on the distortion for MU with $M = 1, 2, 3, 4$ for a given f_s and rate $R = 30$. The upper right shows the minimum distortion for MU for each M.

shown in Lemma 2 that for each frequency, the optimal filter selects the best spectral component possible over the aliased MMSE reconstruction. While intuitively a larger sampling rate should imply less distortion, interestingly, due to the aliasing, the distortion is not monotonically decreasing in f_s since the aliased $S_{x|y}(f)$ might have worse spectral components.

In Fig.4 the minimum distortion for the MUS scheme is shown for $M = 1, 2, 3, 4$ branches. While in general using an increasing number of branches achieves lower distortion, it is observed that in some regimes, using less branches and sampling at a higher rate at each branch achieves lower distortions due to the aliasing suffered by sampling at a lower rate at each branch. We also observe that when the sampling rate f_s is larger that a certain threshold, the lower bound (4) is achieved by MNUS, achieving the lowest distortion achievable by any general A/D scheme, as given in the next lemma.

Lemma 4. *Let θ^* be the parameter θ satisfying equation (4), $\gamma(f_s) \triangleq \inf\{\mathcal{S}(f_s)\}$ and f_s^* be the frequency satisfying $\gamma(f_s^*) = \theta^*$. If $f_s \geq f_s^*$, MNUS achieves the lower bound on the distortion-rate function, i.e. $D^{mn}(R, f_s) = D_{lb}(R)$.*

When $y(t)$ is directly available at the encoder, only the set of spectral components satisfying $S_{x|y}(f) \geq \theta^*$ are transmitted. Hence, to achieve the minimum distortion it is sufficient to reconstruct these bands of $S_{x|y}(f)$. Since, f_s^* is the measure of these bands by definition, when $f_s \geq f_s^*$, MNUS is able to perfectly reconstruct the required frequency set of $S_y(f)$ of measure f_s^* and achieve the lower bound. However, when $f_s < f_s^*$, only a set of measure f_s is recovered with MNUS. The acquisition problem becomes that of recovering an unknown spectral support of $S_y(f)$ to minimize distortion. In a way, f_s^* is a general Nyquist condition in the sense that, when $f_s \geq f_s^*$, the digital samples are sufficient to characterize the continuous process at the lowest possible distortion as if the A/D scheme were not present. In general, f_s^* is much lower the Nyquist rate f_{Ny}, required to reconstruct the process perfectly (in the absence of noise), for example $f_s^* = 14$MHz and $f_{Ny} = 40$MHz in Fig. 3, and the sampling rate required for optimal reconstruction is significantly reduced.

VII. Conclusions

We have considered the acquisition and transmission of a stationary Gaussian time-continuous process corrupted by noise with a general A/D structure consisting of analog processing and a pointwise sampler of constrained sampling rate. We characterized the minimum distortion achievable by any A/D by means of the remote distortion-rate function. We have shown that MMSE estimation followed by standard optimal source compression on the input signal is optimal for stationary Gaussian processes. Then, we have characterized the minimum achievable distortion by multibranch uniform sampling and found the optimal filters minimizing distortion. Finally, we showed that for sufficiently large sampling rates (below the Nyquist rate) multibranch sampling with different sampling rates at each branch is optimal since it achieves the minimum distortion achievable in the absence of A/D conversion.

References

[1] H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, 1967.

[2] M. Mishali and Y. Eldar, "From theory to practice: Sub-nyquist sampling of sparse wideband analog signals," *IEEE Jour. Sel. Topics Sig. Proc.*, vol. 4, no. 2, pp. 375 –391, Apr. 2010.

[3] ——, "Sub-nyquist sampling," *IEEE Sig. Proc. Mag.*, vol. 28, no. 6, pp. 98 –124, Nov. 2011.

[4] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *Signal Processing, IEEE Transactions on*, vol. 50, no. 6, pp. 1417 –1428, Jun. 2002.

[5] 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.

[6] Y. Chen, Y. C. Eldar, and A. J. Goldsmith, "Shannon meets nyquist: Capacity of gaussian sampled channels," *CoRR*, vol. abs/1109.5415, 2011.

[7] Y. Chen, A. J. Goldsmith, and Y. C. Eldar, "Channel capacity under general nonuniform sampling," *CoRR*, vol. abs/1202.0871, 2012.

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

[9] 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, Jul. 1970.

[10] T. Berger, *Rate Distortion Theory*. Prentice Hall, 1971.

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