Abstract—In this paper, the optimal sampling strategies (uniform or nonuniform) and distortion tradeoffs for stationary bandlimited periodic signals are studied. We justify and use both the average and variance of distortion as the performance criteria. To compute the optimal distortion, one needs to find the optimal sampling locations, as well as the optimal pre-sampling filter. A complete characterization of the optimal distortion for the rates lower than half the Landau rate is provided. It is shown that nonuniform sampling outperforms uniform sampling. In addition, this nonuniform sampling is robust with respect to missing samples. Next, for the rates higher than half the Landau rate, we find bounds that are shown to be tight for some special cases. An extension of the results for random discrete periodic signals is discussed, with simulation results indicating that the intuitions from the continuous domain carry over to the discrete domain.

I. INTRODUCTION

Consider the problem of finding the best locations for sampling a continuous signal (nonuniform sampling) to minimize the reconstruction distortion. To find the optimal sampling points, one should utilize any available prior information about the structure of the signal. Furthermore, sampling noise should be taken into account. Due to its importance, nonuniform sampling and its stability analysis has been the subject of numerous works (see [1] for an overview of nonuniform sampling). However, there has been relatively less work for a fully Bayesian model of the signal and sampling noise, where one is interested in the statistical average distortion of signal reconstruction over a class of signals. In this paper, we study this problem for stationary periodic bandlimited signals.

The best locations of nonuniform sampling via an adaptive sampling method is considered in [2] for a first order Markov source. The first order Markov assumption allows the authors to compute the distortions in terms of the length of consecutive sampling times. A tradeoff among sampling rate, information rate and distortion of uniform sampling is derived in [3] for Gaussian sources corrupted by noise. The authors in [4] look at choosing the sampling points for a stationary Gaussian signal with auto-correlation function $R(\tau) = \rho^{\left|\tau\right|}$. It is shown that uniform sampling is optimal in this case. Matthews in [5] considers a wide-sense stationary process, and computes the linear MMSE estimate from sub-Nyquist uniform samples. The authors in [6] compute the optimal pre-and post-filters for a signal corrupted by noise, when uniform sampling is used. The authors in [7] derive a rate distortion function for lossy reconstruction of a multi-component signal, when only a subset of the signal components are sampled. They only consider uniform and random sampling strategies.

Moreover, there has been some work in compressed sensing on the tradeoffs between sampling rate and the reconstruction error. In [8], a sampling rate-distortion region is found; it is shown that arbitrarily small but constant distortion is achievable with a constant measurement rate and a per-sample signal-to-noise ratio. The authors in [9] define a sampling-rate distortion function for an i.i.d. source, and derive lower bounds on the achievable performance.

In this paper, we consider a continuous stationary periodic signal. The signal is assumed to be bandlimited, with at most $2(N_2 - N_1 + 1)$ non-zero Fourier series coefficients from frequency $N_1 \omega_0$ to $N_2 \omega_0$ where $\omega_0 = 2\pi/T$ is the fundamental frequency. Since there are $2N \triangleq 2(N_2 - N_1 + 1)$ free variables, the signal can be uniquely reconstructed from $M = 2N$ noiseless samples (Landau rate). But as we consider taking $M$ noisy samples (at time instances that we choose), unique reconstruction is not feasible, and distortion is inevitable. We also consider a pre-sampling filter in our model. See Fig. 1 for a description of our problem.

The Fourier series coefficients and the sampling noises are all random variables. Therefore, both the input signal and its reconstruction from its samples are random signals, and so is the reconstruction distortion. We are interested in minimizing the expected value and variance of this distortion, by choosing the best sampling points and pre-sampling filter. We denote the minimum of the expected value and variance of the distortion by $D_{\min}$ and $V_{\min}$, respectively. A formal definition of $D_{\min}$ and $V_{\min}$ is given in Section II. A small $D_{\min}$ guarantees a good average performance over all instances of the random signal, while a small $V_{\min}$ guarantees that with high probability, we are close to the promised average distortion for a given random signal (see [12, Appendix A] for further justification). Therefore, our goal is to find the tradeoffs between the number of samples ($M$), the pre-sampling filter, the variance of noise ($\sigma^2$) and the optimal expected value and variance of distortion.

Here, we provide tight results or bounds on the tradeoffs among various parameters such as distortion, sampling rate and sampling noise. When we are below half the Landau rate, $M \leq N$, we find the optimal average distortion ($D_{\min}$) and variance ($V_{\min}$). Interestingly, we show that the minimum
of both $D_{\min}$ and $V_{\min}$ are obtained at the same sampling locations and the same pre-sampling filter. Furthermore, these two minima are obtained when we do not use a pre-sampling filter on the signal bandwidth, and use a nonuniform sampling procedure in which the $M$ samples are arbitrary chosen from the set $\{0, T/N, 2T/N, 3T/N, \ldots , (N-1)T/N\}$. Observe that this set contains uniform sampling points at half the Landau rate. Thus, for $M < N$, the optimal samples are non-uniform. It is worth to note that the sampling locations only depend on the bandwidth of the signal; they are optimal for all values of the noise variance ($\sigma^2$) and $N_1$. Moreover, the sampling points are robust with respect to missing samples. Note that the optimal sampling points are any arbitrary $M$ points from the set $\{0, T/N, 2T/N, 3T/N, \ldots , (N-1)T/N\}$. Thus, if we sample at these positions and we miss some samples, i.e., getting $M' < M$ samples instead of $M$ samples, the set of $M'$ sample points is still a subset of $\{0, T/N, 2T/N, 3T/N, \ldots , (N-1)T/N\}$, and hence optimal. Therefore, if we knew in advance that we get $M'$ samples, this knowledge would not have helped us.

For above half the Landau rate, we provide some lower and upper bounds that are shown to be tight in some cases. When $N < M \leq 2N$, i.e., between half the Landau rate and the Landau rate, we find optimal average distortion when $N/2N_1 - 1$, using a non-uniform set of sampling locations. In addition, when $M > 2N$, uniform sampling is shown to be optimal under certain constraints. Whenever we find $D_{\min}$ and $V_{\min}$ explicitly, the minima are achieved simultaneously at the same optimal sampling points and the pre-sampling filter.

This paper is organized as follows: Section II formally defines the problem. Section III presents the main results of the paper, with the proofs given in [12]. Section IV provides simulation results.

II. PROBLEM DEFINITION

We consider a continuous stationary bandlimited periodic signal defined as follows:

$$ S(t) = \sum_{\ell=-N_1}^{N_2} [A_\ell \cos(\ell \omega_0 t) + B_\ell \sin(\ell \omega_0 t)] $$

where $\omega_0 = 2\pi/T$ is the fundamental frequency. The summation is from $\ell = N_1$ to $N_2$, indicating that the signal is bandlimited. We assume that $A_\ell$ and $B_\ell$ for $N_1 \leq \ell \leq N_2$ are mutually independent Gaussian r.v’s, distributed according to $\mathcal{N}(0, p)$ for $p > 0$. Thus, the signal power is $Np/2$ where $N = N_2 - N_1 + 1$.

The discrete version of the signal is

$$ S[n] = \sum_{\ell=-N_1}^{N_2} [A_\ell \cos(\ell \omega_0 n) + B_\ell \sin(\ell \omega_0 n)] $$

where $\omega_0 = 2\pi/T$ for some integer $T$.

Our model is depicted in Figure 1. The signal $S(t)$, given in (1), is passed through a pre-sampling filter, $H(\omega)$, to produce $\hat{S}(t)$. The signal $\hat{S}(t)$ is sampled at time instances $t = t_1, t_2, \cdots , t_M$, where $t_i \in [0, T]$. These samples are then corrupted by $Z(t)$, an i.i.d. zero-mean Gaussian noise denoted as $\mathcal{N}(0, \sigma^2)$. Thus, our observations are $Y_i = \hat{S}(t_i) + Z_i$, $i = 1, 2, \cdots , M$. An estimator uses the noisy samples $(Y_1, Y_2, \cdots , Y_M)$ to reconstruct the original signal, denoted by $S(t)$. The incurred sampling distortion given by MMSE (minimum mean square error) criterion is equal to

$$ \frac{1}{T} \int_{t=0}^{T} |\hat{S}(t) - S(t)|^2 dt. $$

This distortion is a random variable. Our goal is to compute the minima of the expected value and variance of this random variable, i.e., to minimize

$$ D = \mathbb{E}\left\{ \frac{1}{T} \int_{t=0}^{T} |\hat{S}(t) - S(t)|^2 dt \right\}, $$

$$ V = \text{Var}\left\{ \frac{1}{T} \int_{t=0}^{T} |\hat{S}(t) - S(t)|^2 dt \right\}. $$

We are free to choose $H(\omega)$ and the sampling locations $t_i$’s, $i \in \{1, 2, \cdots , M\}$. Therefore, the optimal distortion $D_{\min}$ is defined as follows:

$$ D_{\min} = \min_{H(\cdot); t_1, \cdots , t_M} \mathbb{E}\left\{ \frac{1}{T} \int_{t=0}^{T} |\hat{S}(t) - S(t)|^2 dt \right\}, $$

$$ V_{\min} = \min_{H(\cdot); t_1, \cdots , t_M} \text{Var}\left\{ \frac{1}{T} \int_{t=0}^{T} |\hat{S}(t) - S(t)|^2 dt \right\}, $$

where the minimization is taken over the sampling locations and the pre-sampling filter.

We assume that $H(\omega)$ is a real linear time invariant (LTI) filter such that $|H(\omega)|^2 \leq 1$ for all $N_1 \leq \ell \leq N_2$, meaning that the frequency gain of the signal is at most one; in other words, we assume that the filter is passive and hence cannot increase the signal energy in each frequency. In particular, all pass filters, $|H(\omega)| = 1$, satisfy this assumption. The reason for introducing this assumption is that we can always normalize the power gain of the filter.

III. MAIN RESULTS

In this section, first two general lower bounds on the average and variance of distortion are given. Then the main results of the paper are stated. Note that $\text{SNR} = Np/\sigma^2$.

A. Two General Lower Bounds on the Average and Variance of Distortion

Lemma 1. For any noise variance $\sigma > 0$ and any choice of the pre-sampling filter, the following lower bounds hold for all values of $M$:

$$ \frac{D_{\min}}{p} \geq \frac{1}{2} \left( 2N - M + \frac{M}{1 + \text{SNR}} \right), $$

$$ \frac{V_{\min}}{p^2} \geq 2N - M + \frac{M}{(1 + \text{SNR})^2}. $$

When we do not use a pre-sampling filter on the signal bandwidth (i.e., $H(\omega_0) = 1$ for $\ell = N_1, \cdots , N_2$), equality in
the above equations holds if and only if for any time instances \( t_i \) and \( t_j \), one of the following two equations holds:

\[
|t_i - t_j| = T \frac{m_1}{N}, \quad \text{for some integer } m_1, \quad (5)
\]

\[
|t_i - t_j| = T \frac{2m_2 + 1}{2(N_1 + N_2)}, \quad \text{for some integer } m_2. \quad (6)
\]

**Lemma 2.** For any \( \sigma > 0 \) and any choice of the pre-sampling filter, the following lower bounds hold for all values of \( M \):

\[
D_{\text{min}} \geq \frac{Np}{1 + \frac{M}{2} SNR}, \quad (7)
\]

\[
V_{\text{min}} \geq \frac{2Np^2}{(1 + \frac{M}{2} SNR)^2}. \quad (8)
\]

Furthermore, when we do not use a pre-sampling filter on the signal bandwidth, equality in the above equation holds if the sampling time instances satisfy \( \sum_{i=1}^{M} e^{j2\pi k/2} = 0 \) for \( k \in \{1, 2, \cdots, N-1\} \cup \{2N_1, 2N_1+1, \cdots, 2N_2\} \).

Note that for \( M < 2N \), Lemma 1 gives better lower bounds on \( D_{\text{min}} \) and \( V_{\text{min}} \) whereas for \( M > 2N \), Lemma 2 gives tighter ones.

**B. Sampling Distortion Tradeoffs for \( M \leq N \)**

The sampling-rate distortion tradeoffs are studied when the sampling rate is lower than the Landau rate.

**Theorem 1.** For \( M \leq N \), the optimal average and variance of distortion for any given pre-sampling filter are given by

\[
D_{\text{min}} = \frac{1}{2} \left( 2N - M + \frac{M}{1 + SNR} \right), \quad (9)
\]

\[
V_{\text{min}} = 2N - M + \frac{M}{(1 + SNR)^2}. \quad (10)
\]

Both the minimal average distortion and its variance are obtained with the pre-sampling filter \( H(\ell \omega_0) = 1 \) for \( \ell = N_1, \cdots, N_2 \) and choosing \( M \) distinct time instances arbitrarily from the following set of \( N \) samples:

\[
\{0, \frac{1}{N}, \frac{2}{N}, \frac{3}{N}, \cdots, \frac{N-1}{N}\}.
\]

The optimal interpolation formula for this set of sampling points is given by

\[
\hat{S}(t) = \frac{p}{Np + \sigma^2} \sum_{i=1}^{M} \sum_{\ell = N_1}^{N_2} \cos(\ell \omega_0(t - t_i)) Y_i,
\]

\[
= \frac{p}{Np + \sigma^2} \sum_{i=1}^{M} \left( \frac{\cos(\frac{T}{2}(t - t_i))}{2(1 + SNR)} \sin(\frac{N\omega_0}{2}(t - t_i)) \cdot Y_i \right), \quad (11)
\]

where \( Y_i \) is the noisy sample of the signal at \( t = t_i \).

**Remark 1.** Consider (9). Observe that for fixed values of \( p \) and \( \sigma \), the minimum distortion is linear in \( M \). But it is not linear in \( N \) (as \( SNR \) in the denominator is \( Np/\sigma^2 \), except when \( SNR \) goes to infinity. In the case of noiseless samples, \( \sigma = 0 \), the \( SNR \) will be infinity and the minimal distortion will be \( p(2N-M)/2 \). To intuitively understand this equation, observe that there are \( 2N \) free variables and we can recover \( M \) of them using the samples. Therefore, we will have \( 2N-M \) free variables, giving a total distortion of \( (2N-M)p/2 \) as the power of each sinusoidal function is \( p/2 \). Moreover, the maximum distortion is \( 2N(p/2) \), which is obtained when \( \sigma = \infty \) (\( SNR = 0 \)) or \( M = 0 \). Observe that when \( SNR \) is large, \( D_{\text{min}} \approx \frac{p}{2} \left( 2N - M + \frac{M}{SNR} \right) = \frac{p}{2} (2N - M) + \frac{M \sigma^2}{2N} \)

which is linear in both \( p \) and \( \sigma^2 \). Observe that when \( M \leq N \), \( D_{\text{min}} \) is increasing in both \( p \) and \( \sigma^2 \).

**C. Sampling-Distortion Tradeoffs for \( N < M \leq 2N \)**

For the sampling rates \( N < M \leq 2N \), we have:

**Theorem 2.** For \( N < M \leq 2N \) and any choice of pre-sampling filter, the optimal average distortion can be bounded as follows

\[
D_{\text{min}} \geq \frac{1}{2} \left( 2N - M + \frac{M}{1 + SNR} \right), \quad (12)
\]

\[
\frac{D_{\text{min}}}{p} \leq \frac{1}{2} \left( 2N - M + \frac{2N - M}{1 + SNR} \right) + \text{Num} \cdot \frac{1}{1 + 2 SNR} + (M - N - \text{Num}) \cdot \frac{1}{1 + SNR}, \quad (13)
\]

in which

\[
\text{Num} = \min \left( f(N_1, N), M - N \right)
\]

and \( f(a, b) \) is equal to

\[
f(a, b) = \begin{cases} 
    b - 1 & \text{if } r = 0 \\
    2b - 2r + 1 & \text{if } 2r > b \\
    2r - 1 & \text{if } 0 < 2r \leq b 
\end{cases}
\]
where $r$ is the remainder of dividing $a$ by $b$.

The lower bound given in (12) is tight when $N|2N_1 - 1$ for any arbitrary $M \leq 2N$. To achieve the optimal distortion, no pre-sampling filter is necessary and the optimal time instances can be chosen from the set
\[ \left\{ 0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N - 1}{N} \right\}. \]
Moreover, these lower bounds are tight with uniform sampling, where $M$ can be chosen from the set $\{ 0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N - 1}{N} \}$. Hence, the lower bounds given in Lemma 1 are optimal in general.

\[ \frac{D_{\min}}{p} \geq \frac{N}{1 + \frac{2N^2}{SNR}}, \quad (14) \]
\[ \frac{V_{\min}}{p^2} \geq \frac{2N}{(1 + \frac{2N^2}{SNR})^2}. \quad (15) \]

Moreover, these lower bounds are tight with uniform sampling, i.e., $t_i = iT/M$, $i = 1, 2, \ldots, M$, and without filtering on the signal bandwidth, if for each $k$ in the interval $2N_1 \leq k \leq 2N_2$, we have $M \leq k$. In particular, uniform sampling is optimal when $M > 2N_2$. Also the reconstruction formula for the optimal set of sampling points is given by
\[ \hat{S}(t) = \frac{p}{Np + \sigma^2} \sum_{i=1}^{M} \sum_{\ell=N_1}^{N_2} \cos(\ell \omega_0 (t - t_i)) Y_i, \]
\[ = \frac{p}{Np + \sigma^2} \sum_{i=1}^{M} \left( \cos \left( \frac{\omega_0}{2} (t - t_i) \right) \right. \]
\[ \times \left. \frac{\sin \left( \frac{N \omega_0}{2} (t - t_i) \right)}{\sin \left( \frac{\omega_0}{2} (t - t_i) \right)} \cdot Y_i \right), \quad (17) \]
where $Y_i$ is the noisy sample of the signal at $t = t_i$.

**Remark 2.** When $M$ goes to infinity, the minimum distortion $D_{\min} = O(1/M)$ goes to zero, regardless of the sampling noise value. To see this, observe that when $M$ is very large ($M > 2N_2$), the lower bounds given in (14) and (15) are tight.

**Remark 3.** Unlike the case of $M \leq N$ where optimal sampling points could be found without any need to know $N_1$, we show in [12] that this is no longer the case for $M > 2N$. Besides, the uniform sampling strategy is not optimal in general.

**E. Discrete Signals**

Consider a real periodic discrete signal of the form
\[ S[n] = \sum_{\ell = N_1}^{N_2} \left[ A_\ell \cos(\ell \omega_0 n) + B_\ell \sin(\ell \omega_0 n) \right], \quad (18) \]
where $\omega_0 = \frac{2\pi}{T}$ for some integer $T$, and $1 \leq N_1 \leq N_2 < T/2$ are natural numbers. Observe that $T$ is the period of the discrete signal and $A_\ell - jB_\ell$ is the $\ell$-th DFT coefficient of the discrete signal $S[n]$.

Suppose that we have $M$ noisy samples at time instances $\{ t_1, t_2, \ldots, t_M \}$. We want to use these samples to reconstruct the discrete signal $S[n]$. If the reconstruction is $\hat{S}[n] = \sum_{\ell = N_1}^{N_2} \left[ A_\ell \cos(\ell \omega_0 n) + B_\ell \sin(\ell \omega_0 n) \right]$, the distortion $D = \sum_{n=1}^{N} |S[n] - \hat{S}[n]|^2$ is proportional to $\sum_{\ell=1}^{M} |A_\ell - \hat{A}_\ell|^2 + |B_\ell - \hat{B}_\ell|^2$. Thus, the formulation for minimizing the distortion is the same as that of the continuous signals. The only additional restriction is that $t_i$’s should be integers.

Whenever the optimal sampling points in the continuous formulation turn out to be integer values, they are the optimal points in the discrete case. And when the optimal sampling points are not integers, simulation results with exhaustive search show that their closest integer values are either optimal or nearly optimal. For example, suppose that the signal period is $T = 15$ and $(M, N_1, N) = (3, 1, 4)$. In this case $M \leq N$ and the optimal sampling points in the continuous case are $\{ t_1, t_2, t_3 \} \subset \{ 0, 3.75, 7.5, 11.25 \} + \tau$, for some $\tau \in [0, T]$. The rounds of these point for $\tau = 1.1$ are optimal in the discrete case, i.e., any choice of time instances from the set $\{ 1, 5, 8, 12 \}$ is optimal. On the other hand, when the signal period is $T = 15$ and $(M, N_1, N) = (4, 1, 6)$. Here again $M \leq N$ and $\{ t_1, t_2, t_3, t_4 \} \subset F = \{ 0, 2.5, 5, 7.5, 10, 12.5 \} + \tau$ are the only continuous optimal points. Choosing the time instances from their closest integers yields $D = 8.0069$. However, the optimal points are $\{ 1, 2, 8, 9 \}$ resulting in $D_{\min} = 8.0068$. Note that the distance between the first two optimal points is 1, which cannot be achieved by choosing any arbitrary value of $\tau$, since the distance between any of the points in $F$ is 2.5.

**IV. Simulation Results**

In Fig. 3, two random signals are generated using parameters $(N, N_1, p, \sigma) = (7, 4, 0.1, 1)$. Their reconstruction with $M = 13$ samples are also depicted, employing the optimal MMSE method and the iterative method of [11]. The sampling points are optimally chosen. The average reconstruction distortion (MMSE) over all random signals with the given parameters is $D = 0.5093$. In the top subfigure, the distortion of the optimal MMSE method is $D = 0.0200$ and the distortion of the iterative method is $D_{\text{iterative}} = 0.0379$. In the bottom subfigure, the distortion of the optimal MMSE method is $D = 1.6038$ and the distortion of the iterative method is $D_{\text{iterative}} = 1.6042$. We see that the performance of the iterative method is near optimal in the high SNR scenario. Fig. 4 uses the same parameters for generating the signal, but the variance of the noise is 4 (low SNR scenario). In this case, the average distortion over all random signals is $D = 1.9444$. The distortions of the top subfigure are $D = 0.2414$ for the optimal MMSE method, and $D_{\text{iterative}} = 0.3079$ for the iterative method. For subfigure (b), they are $D = 1.8347$ and $D_{\text{iterative}} = 1.9700$, respectively.

The maximum of the lower bounds given in Lemma 1 and Lemma 2 is depicted in Fig. 2. In this figure, the distortion of the uniform sampling without any pre-sampling filter on
the signal bandwidth is also depicted. The performance of the uniform sampling is close to optimal for $M \leq N$ (its curve almost matches that of the lower bound in Lemma 1 which is optimal for $M \leq N$). We observe that near the Landau rate, increasing the number of sampling points does not decrease the distortion. Here the curve for uniform sampling reaches its maximum at $M = 26$, whereas the Landau rate is 18. The curve reaches the second lower bound at $M = 35$. Thus, uniform sampling is optimal for the rates above the Nyquist rate ($M > 2N_2 = 34$). Finally, the explicit upper bound on the performance of the sampling points given in Theorem 2 is also depicted. As shown in the figure, this upper bound can be below the distortion of the uniform sampling.

ACKNOWLEDGMENT

The authors would like to thank Amin Gohari and Salman Beigi for helpful discussions.

REFERENCES