SAMPLING AND DISTORTION TRADEOFFS FOR INDIRECT SOURCE RETRIEVAL

Elaheh Mohammadi, Alireza Fallah, and Farokh Marvasti

Advanced Communication Research Institute (ACRI)

Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran

ABSTRACT

Consider a continuous signal that cannot be observed directly. Instead, one has access to multiple corrupted versions of the signal, which are correlated because they carry information about the common remote signal. The goal is to reconstruct the original signal from the data collected from its corrupted versions, and we are interested in the optimal number of samples and their locations for each corrupted signal to minimize the total reconstruction distortion of the remote signal. The correlation among the corrupted signals can be utilized to reduce the sampling rate. For a class of Gaussian signals, we propose a fundamental lower bound on the reconstruction distortion for any arbitrary non-uniform sampling strategy in general. In particular, we show that in the low sampling rate region, it is optimal to use a certain non-uniform sampling scheme on all the signals. On the other hand, in the high sampling rate region, it is optimal to uniformly sample all the signals. We also show that both of these sampling strategies are optimal if we are interested in recovering the set of corrupted signals, rather than the remote signal.

SYSTEM MODEL

Let $S(t)$ be the original stochastic signal, and show its Fourier coefficients by random variables A_l and B_l. We assume that the coefficients A_l and B_l are zero when $l > N_2$ or $l < N_1$ for some natural numbers $N_1 \le N_2$. Therefore,

$$S(t) = \sum_{l=N_1}^{N_2} [A_l \cos(l\omega t) + B_l \sin(l\omega t)], \qquad t \in [0,T]$$

where the coefficients A_l and B_l for $\mathrm{N}_1 \le l \le N_2$ are mutually independent identically distributed (i.i.d.) normal $\mathcal{N}(0,1)$variables.

We cannot observe $S(t)$ directly. Instead, we have $S_1(t), S_2(t), \dots, S_k(t)$, defined on the same interval, that are corrupted versions of $S(t)$. The corrupted versions of the signal can be expressed as

$$S_i(t) = \sum_{l=N_1}^{N_2} [A_{il} \cos(l\omega t) + B_{il} \sin(l\omega t)], \qquad t \in [0,T], i \in \{1,2,..,k\}$$

Where $A_{il} = A_l + W_{il}$ and $B_{il} = B_l + V_{il}$; here W_{il} and V_{il} are independent perturbations that are added to the original signal. It is assumed that the perturbations W_{il} and V_{il} for $i \in \{1,2,..,k\}$ and l∈ $\{N_1, N_1+1, \dots, N_2\}$ are i.i.d. variables according to $\mathcal{N}(0,\eta)$. The perturbations are also mutually independent of the signal coefficients A_l and B_l for $\mathrm{N}_1 \le l \le N_2$.

We are allowed to take m_i samples from the ith corrupted signal $S_i(.)$ at time instances $t_{i1}, t_{i2}, \dots, t_{im_i} \in [0,T]$ of our choice, for $i = 1,2,..,k$. Therefore$^{m_i}/_T$ can be viewed as the sampling rate of the $S_i(.)$

SYSTEM MODEL (Cont.)

We assume that the samples are noisy. The sampling noise can model quantization noise of an A/D convertor, or the noise incurred by transmitting the samples to a fusion center over a communication channel. The sampling noise of each signal $S_i(t)$ is modeled by an independent normal $\mathcal{N}(0, \sigma_i^2)$ variable. We use the samples to reconstruct either the remote signal $S(t)$, or the collection of corrupted signals $S_1(t), S_2(t), \dots, S_k(t)$. We are interested in the latter since these individual signals may contain some other information of interest beside $S(t)$, e.g., the differences $S(t) - S_i(t)$ might be correlated with some other signal of interest.

The reconstruction of the remote signal and the corrupted signal are denoted by $\hat{S}(t)$ and $\hat{S}_i(t)$, respectively. The goal is to optimize over the sampling times t_{ij} to minimize the average Minimum Mean Square Error (MMSE) distance between the signals and their reconstructions. More specifically, we consider the minimization

$$D_{a\,min} = \min_{\{t_{i1}, t_{i2}, \dots, t_{im_i}\}_{i=1}^k} \frac{1}{T} \int_{t=0}^{T} \mathbb{E}\{|\hat{S}(t) - S(t)|^2\} dt \quad (1)$$

for the remote signal, or the minimization

$$D_{b\,min} = \min_{\{t_{i1}, t_{i2}, \dots, t_{im_i}\}_{i=1}^k} \frac{1}{T} \sum_{i=1}^{k} \int_{t=0}^{T} \mathbb{E}\{|\hat{S}_i(t) - S_i(t)|^2\} dt \quad (2)$$

for reconstruction of the k corrupted signals $S_i(t), i = 1,2,\dots,k$. Here t_{ij} is the jth sampling time of the ith signal.

MAIN RESULTS

Let $N = N_2 - N_1 + 1$ and $f_0 = \frac{1}{T}$. Also we call Nf_0 the signal bandwidth, and $2N_2 f_0$ the Nyquist rate (twice the maximum frequency of the signal).

To state the main result, we need a definition. For any real p, let $\phi_i = \frac{m_i}{2\sigma_i^2} + \frac{1}{\eta}$, and

$$\Phi_p = \sum_{i=1}^{k} \phi_i^p$$

Theorem 1 (Reconstruction of the original signal) The following general lower bound on the optimal distortion (given in (1)) holds:

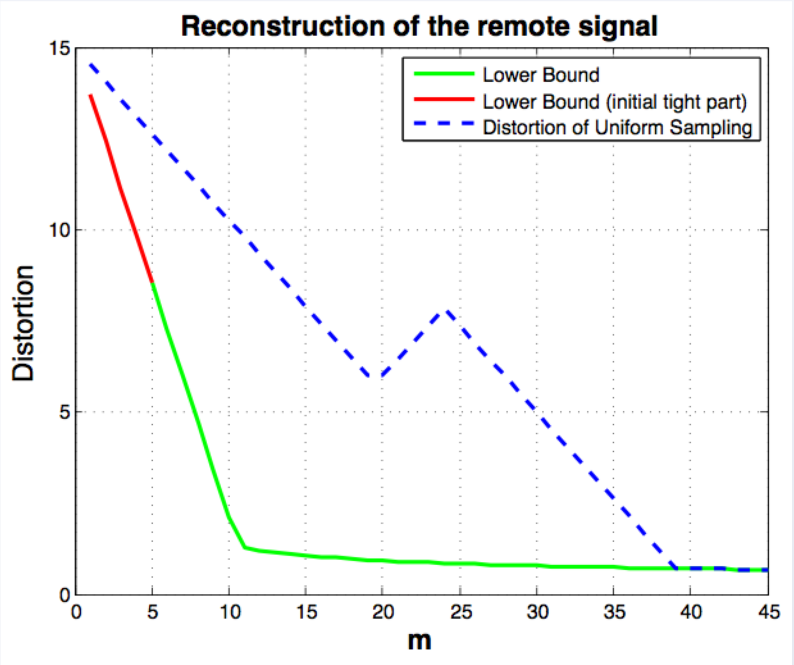
$$D_{a\,min} \ge \max\left(N - N \sum_{i=1}^{k} \frac{m_i}{2(1+\eta)N + 2\sigma_i^2}, \frac{N}{\left(1 + \frac{k}{\eta}\right) - \frac{k^2}{\eta^2}(\Phi_1)^{-1}}\right)$$

And the equality holds when

1. $\sum_{i=1}^{k} m_i \le N$: in this case, the optimal sampling points, t_{ij}, are all distinct for $1 \le i \le k, 1 \le j \le m_i$, and belong to the set $\{0, \frac{T}{N}, \dots, \frac{(N-1)T}{N}\}$.
2. $m_i > 2N_2$ for $1 \le i \le k$: in this case uniform sampling of each signal $S_i(t)$ is optimal.

MAIN RESULTS(Cont.)

For $(k, N_1, N, N_2, \sigma_i^2, \eta) = (3,5,15,19,1,0.1)$, and assuming $m_i = m$ we have:



Theorem 2 (Reconstruction of the set of corrupted signals): The following general lower bound on the optimal distortion (given in (2)) holds:

$$D_{b\,min} \ge \max\left(Nk(1+\eta) - N((1+\eta)^2 + (k-1)) \sum_{i=1}^{k} \frac{m_i}{2(1+\eta)N + 2\sigma_i^2}, N\Phi_{-1} + \frac{N}{\eta(\eta+k) - \Phi_{-1}} \Phi_{-2}\right)$$

And the equality holds when

1. $\mathrm{m} = \sum_{i=1}^{k} m_i \le N$: in this case, the optimal sampling points, t_{ij}, are all distinct for $1 \le i \le k, 1 \le j \le m_i$, and belong to the set $\{0, \frac{T}{N}, \dots, \frac{(N-1)T}{N}\}$.
2. $m_i > 2N_2$ for $1 \le i \le k$: in this case uniform sampling of each signal $S_i(t)$ is optimal.

For same parameters we have: