TIME-STAMPLESS ADAPTIVE NONUNIFORM SAMPLING FOR STOCHASTIC SIGNALS

Soheil Feizi, Vivek K Goyal, and Muriel Médard

RLE at MIT

ABSTRACT

In this paper, we introduce a *time-stampless adaptive nonuniform sampling* (TANS) framework, in which time increments between samples are determined by a function of the m most recent increments and sample values. Since only past samples are used in computing time increments, it is not necessary to save sampling times (time stamps) for use in the reconstruction process. We focus on two TANS schemes for discrete-time stochastic signals: a greedy method, and a method based on dynamic programming. We analyze the performances of these schemes by computing (or bounding) their trade-offs between sampling rate and expected reconstruction distortion for Markovian signals. Simulation results support the analysis of the sampling schemes. We show that by opportunistically adapting to local signal characteristics TANS may lead to improved power efficiency in some applications.

I. INTRODUCTION

Sampling is essential in any digital system that interfaces with the analog world. All else being equal, it is desirable to minimize the number of samples while maintaining an acceptable reconstruction distortion. In some applications, minimizing the number of samples can be translated into having a power-efficient sampling, since the power consumption at an analog-to-digital converter (ADC) is approximately proportional to its sampling rate [1]. Also, having fewer samples can increase the efficiency of other processing of these measurements. For example, if these samples should be transmitted to another place via a communication channel, having fewer samples will improve power and bandwidth efficiencies.

A uniform sampling at the Nyquist rate of the signal may cause some redundant samples, since the global signal bandwidth may not be a good measure of local variations of the signal. Although traditional nonuniform sampling schemes (e.g., [2]) deal with this problem, they have certain limitations. Firstly, they are mostly designed to operate under specific conditions for restrictive signal models (e.g., [3]) and, secondly, sampling times (i.e., time stamps) must be stored/transmitted to be used in the reconstruction process. This may cause power/bandwidth inefficiencies in sampling/communication procedures.

In this paper, we introduce a new framework for an adaptive nonuniform sampling scheme (see Figure 1). The key idea of this framework is that *time increments between samples are computed by using a function of previously taken samples*. Therefore, keeping sampling times (time stamps), except initialization times, is not necessary. The function by which sampling time intervals is computed is called the *sampling function*. The aim of this sampling framework is to have a balance between the reconstruction distortion and the average sampling rate. We refer to this sampling framework as *Time-stampless Adaptive Nonuniform Sampling* (TANS). The TANS concept can be applied on continuous- or discrete-time signals, and the design and analysis can be based on deterministic or stochastic models.

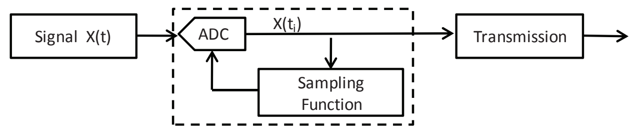

Fig. 1. A schematic view of the TANS framework: sampling times are determined by a function of m most recently taken samples. Hence, it is not necessary to save sampling times (time stamps) for use in the reconstruction process.

II. TANS FRAMEWORK

In this section, we introduce the TANS framework. Fix some nonnegative integer m and suppose the ith sample of signal $X(t)$ is taken at time t_i. We take the $(i+1)$st sample after a time increment of $T_i = f\left(\{(t_j, X(t_j)) : i-m+1 \leq j \leq i\}\right)$, where f is called the sampling function. This makes the sampling rate adapt to local characteristics of the signal. Since the time increment is a function of the m most recently taken samples, we say the *order* of the sampling function f is m. The sampling is nonuniform except in the trivial cases when f is a constant-valued function (e.g., $m = 0$). Some initialization of the first m sampling times is necessary, but the effect of this initialization on the rate is amortized.

The sampling function is known at the reconstruction side. Assuming that the *state* $S_{t_i} = \{(t_j, X(t_j)) : i - m + 1 \leq j \leq i\}$ is also known at the reconstruction side when reconstructing $X(t)$ on $[t_i, t_{i+1}]$, there is *no need for the sampling times (time stamps) to be transmitted*. These times can be computed by using the sampling function and previously taken samples: $t_{i+1} = t_i + f(S_{t_i})$. This type of synchronization in an adaptive system without explicit communication is often called backward adaptation [4]. In a practical setting involving both sampling and quantization, backward adaptivity requires using the quantized values to drive the adaptation [5]. Here, to maintain focus on sampling rate and adaptation of sampling increments, we do not explicitly include quantization effects. Note that while the sampling time selection is causal, the reconstruction method can be causal or non-causal.

Suppose $\widehat{X}(t)$ is the reconstructed signal computed by some reconstruction method. For the case of discrete time and a stochastic signal model, define $d(S_{t_i}, T_i)$ as the expected reconstruction distortion over samples from time $t_i + 1$ until time $t_{i+1} - 1$. That is,

$$d(S_{t_i}, T_i) = \mathbb{E}_{\mathcal{X}}\left[\sum_{t=t_i+1}^{t_{i+1}-1} D\big(X(t), \widehat{X}(t)\big)\right],$$

where $\mathcal{X}$ is the known probabilistic model of the signal $X(t)$ and $D(X(t), \widehat{X}(t))$ represents the distortion at time t.Note that at times t_i and t_{i+1} the reconstruction distortion is zero since exact sample values are known at these times. In realistic cases and for a given state S_{t_i}, $d(S_{t_i}, T_i)$ is an increasing function with respect to T_i, because the greater the next sampling step, the greater the reconstruction distortion. On the other hand, the greater the next sampling step, the larger the rate benefit. Hence, a rate penalty can be defined as $a(S_{t_i}, T_i) = \rho/f(S_{t_i}) = \rho/T_i$, where ρ is a rate award parameter. We define the cost of each sampling state as the sum of the expected reconstruction distortion and the rate penalty, that is, $c(S_{t_i}, T_i) = d(S_{t_i}, T_i) + a(S_{t_i}, T_i)$. The overall cost of the sampling process is the sum of different sampling state costs, that is, $\sum_i c(S_{t_i}, T_i)$.

In this paper, we consider Markovian signals:

$$X(t+1) = \alpha_{\theta_t} X(t) + Z_{\theta_t}(t+1), \tag{1}$$

where θ_t represents the state of a hidden Markov chain (MC) with state transition probabilities depicted in Figure 2. At time t, if the MC is at state 0, $\theta_t = 0$; otherwise, $\theta_t = 1$. Depending on the value of θ_t, the signal is generated by a first-order AR model with parameter α_{θ_t} and the noise variance $1 - \alpha_{\theta_t}^2$.

In the TANS problem setup, we are interested in designing sampling functions which minimize the overall sampling cost. In this paper, we propose two systematic ways to design sampling functions. In Section III, we describe greedy methods and in Section IV, dynamic programming based sampling functions are investigated.

III. GREEDY TANS

In this section, we investigate greedy sampling functions. In all of these sampling schemes, the reconstruction function is assumed to be a generalized linear prediction filter (a linear prediction filter which uses a set of nonuniform samples) introduced in reference [6]. Note that it is a causal reconstruction function.

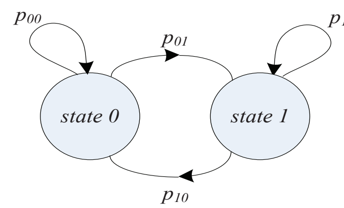

Fig. 2. A hidden Markov chain considered in Markovian signal model of equation (1).

In greedy methods, a sampling function is computed as follows:

$$T_i = \arg\min_{T} \quad c(S_{t_i}, T) \tag{2}$$

where $f(S_{t_i}) = T_i$.

Intuitively, at each sampling state, the next sample is taken to minimize the cost of that state. Consider a Markovian signal described by (1), where θ_t represents the state of a hidden underlying Markov chain depicted in Figure 2. In this section, for simplicity we assume the MC is symmetric (i.e., $p_{01} = p_{10}$). However, all arguments can be extended for a general MC. We also assume that α_0 and α_1 are known. However, the state of the Markov chain (i.e., θ_t) is unknown and needed to be estimated by using the taken samples. We use a generalized linear prediction filter for the reconstruction.

Extending the previous notation, define $\theta_{S_{t_i}}$ as the state of the MC during the sampling state S_{t_i}. If during S_{t_i} the MC state stays at zero, $\theta_{S_{t_i}} = 0$. Similarly, if the MC state stays at one, $\theta_{S_{t_i}} = 1$. Otherwise, if there is an MC transition within this sampling state, $\theta_{S_{t_i}} = 2$. We assume that $\theta_{S_{t_i}}$ is unknown and needs to be estimated by using the taken samples. The estimated value of $\theta_{S_{t_i}}$ is referred by $\hat{\theta}_{S_{t_i}}$. The error probability of this estimation is referred by $P_e(S_{t_i}) = Pr(\hat{\theta}_{S_{t_i}} \neq \theta_{S_{t_i}})$.

The following described a greedy TANS method for Markovian signals:

Algorithm 1: A greedy sampling function for the considered Markovian signal has the following steps:

- Step i,0: Compute $\hat{\theta}_{S_{t_i}}$ and $P_e(S_{t_i})$.
- Step i,1: Compute $T_i = \arg\min_T c(S_{t_i}, T|\hat{\theta}_{S_{t_i}})$, where $c(S_{t_i}, T|\hat{\theta}_{S_{t_i}})$ is the sampling state cost given $\hat{\theta}_{S_{t_i}}$ (see (3) and (4)).
- Step i,2: Take a sample at time $t_i + T_i$.
- Step i,3: Compute S_{t_i+1}. Repeat.

For simplicity, we assume that the sampling increment T is small enough that the probability of having more than one MC transition is negligible. In other words, we assume

$\max_i T_i \leq T_{up}$, $p_{00}^{(T_{up}-1)} \gg \frac{1}{2}$ and $p_{11}^{(T_{up}-1)} \gg \frac{1}{2}$. If MC transition probabilities p_{01} and p_{10} are small enough, this assumption is reasonable. Under these assumptions, the sampling state cost, $c(S_{t_i}, T)$, can be conditioned on the value of $\hat{\theta}_{S_{t_i}}$ as follows:

$$\begin{aligned} c(S_{t_i}, T \,|\, \hat{\theta}_{S_{t_i}} = 0) & \\ \approx \quad (1 - P_e(S_{t_i})) \sum_{\ell=1}^{T-1} (1 - \alpha^{2\ell}) & \\ + P_e(S_{t_i})(T-1)\sigma_{\max}^2 + \frac{\rho}{T}. & \quad (3) \end{aligned}$$

It says, if the estimation is correct (with probability $1 - P_e(S_{t_i})$), the sampling state cost is $\sum_{l=1}^{T} 1 - \alpha_0^{2l}$. If the estimation process fails (with probability $P_e(S_{t_i})$), a maximum prediction error variance $\sigma_{\max}^2$ occurs where $\sigma_{\max}^2$ is the maximum prediction error variance (in this example, $\sigma_{\max}^2 = 1$). The sampling state cost function conditioned on $\hat{\theta}_{S_{t_i}} = 1$ (i.e., $c(S_{t_i}, T \,|\, \hat{\theta}_{S_{t_i}} = 1)$) can be written similarly. Finally, for the case $\hat{\theta}_{S_{t_i}} = 2$, we assume that the prediction variance is the maximum prediction error variance $\sigma_{\max}^2$:

$$c(S_{t_i}, T \,|\, \hat{\theta}_{S_{t_i}} = 2) = (T-1)\sigma_{\max}^2 + \frac{\rho}{T}. \quad (4)$$

We analyze the performance of the proposed greedy sampling scheme in Theorem 2. Before presenting this theorem, we introduce some notations. Suppose that, for all S_{t_i}, $P_e^{\text{low}} \leq P_e(S_{t_i}) \leq P_e^{\text{up}}$. By considering an upper bound on $P_e(S_{t_i})$, we define

$$\begin{aligned} T_0^{\text{low}} = \arg\min_{T} \quad (1 - P_e^{\text{up}}) \sum_{l=1}^{T-1} (1 - \alpha_0^{2l}) & \\ + P_e^{\text{up}}(T-1)\sigma_{\max}^2 + \frac{\rho}{T}. & \quad (5) \end{aligned}$$

T_0^{up} is defined similarly by considering a lower bound on $P_e(S_{t_i})$. Analogously, T_1^{up} and T_1^{low} can be defined.

Also, d_0^{up}, an upper bound on the expected reconstruction distortion per sample given $\hat{\theta}_{S_{t_i}} = 0$ is defined as follows:

$$d_0^{\text{up}} = \frac{1}{T_0^{\text{up}}} \big\{ (1 - P_e^{\text{up}}) \sum_{\ell=1}^{T_0^{up}-1} (1 - \alpha_0^{2\ell}) + P_e^{up}(T_0^{up} - 1)\sigma_{\max}^2 \big\}.$$

Quantities d_0^{low}, d_1^{up} and d_1^{low} are defined similarly.

The following theorem provides analytical upper and lower bounds on the average sampling rate and the expected reconstruction distortion of the greedy sampling scheme introduced in Algorithm 1.

Theorem 2: Consider a Markovian signal defined in (1) over a large enough time interval $[0, T_{\text{tot}}]$. An achievable rate-distortion pair (R, D) of the greedy sampling scheme of Algorithm 1 can be bounded as follows:

$$\begin{aligned} \frac{1}{2T_0^{\text{up}}} + \frac{1}{2T_1^{\text{up}}} \leq \quad R \quad & \leq \frac{1}{2T_0^{\text{low}}} + \frac{1}{2T_1^{\text{low}}} \\ \frac{d_0^{\text{low}}}{2} + \frac{d_1^{\text{low}}}{2} \leq \quad D \quad & \leq \frac{d_0^{\text{up}}}{2} + \frac{d_1^{\text{up}}}{2}. \end{aligned}$$

Remarks:

1) The proof of Theorem 2 with analytical solutions of optimization setups can be found in the full version of this paper on arXiv. In this paper, we compare analytical results with simulation ones in Section V.
2) In the first step of this sampling function, we need to estimate $\theta_{S_{t_i}}$ by using m most recently taken samples of the signal and compute its probability of error $P_e(\hat{\theta}_{S_{t_i}})$. If m is not too large, a maximum likelihood estimation can be used. However, when the order of the sampling function is large and/or autocorrelation coefficients change continuously, a maximum likelihood estimator may not be practically interesting. In these cases, we can use previously taken samples within a window of size W from the last sample (i.e., all taken samples from time $t_i - W + 1$ to the time t_i) to update or estimate autocorrelation coefficients to use in the sampling function. A gradient-based update can be used as follows: Suppose at the sampling state $S_{t_{i-1}}$, the set of estimated autocorrelation coefficients is $\{\hat{r}(j) : j \geq 1\}$. By taking a sample at time n_i, these coefficients are updated as follows:

$$\hat{r}(j) := \hat{r}(j) + \gamma(X(t_i)X(t_i - j) - \hat{r}(j)) \quad (6)$$

for all possible j's, where := represents an update sign, and $\gamma > 0$ is a gradient step size. This gradient-based update method can be useful when W is not large.

IV. DYNAMIC PROGRAMMING-BASED TANS

In greedy TANS, sampling functions are derived based on minimizing the sampling cost at each sampling state. Hence, it does not take into account the *quality* of next sampling states. Intuitively, the larger the sampling increment at the sampling state S_{t_i}, the lower the quality of the next sampling state. Therefore, in general, greedy methods may not provide optimal sampling functions with respect to the overall sampling cost.

We consider *quality* of next sampling states in dynamic programming-based TANS methods.

A solution of this Bellman-Ford equation (BFE) ([7]) can provide an optimal solution for the TANS sampling problem when the reconstruction function is causal. However, finding this optimal solution in most of cases can be computationally difficult. Here, we propose sampling functions based on approximate dynamic programming (ADP) algorithms. We define a *quality* function $q(S_{t_i})$ for each sampling state S_{t_i}. A greedy solution is used to define this quality function.

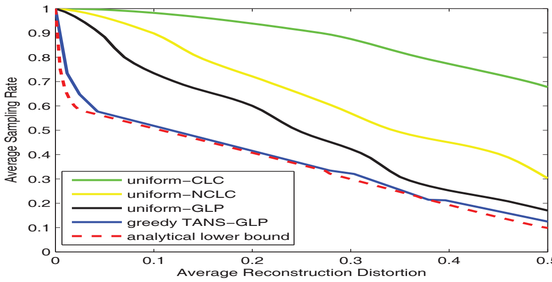

Fig. 3. Comparison of average sampling rate versus average reconstruction distortion for a Markovian signal for various methods.

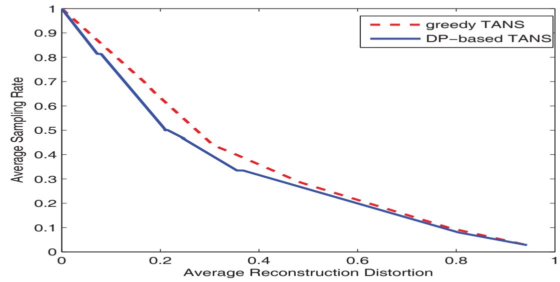

Fig. 4. Comparison of a dynamic programming-based TANS with greedy TANS for a Markovian signal model.

Then, a sampling function can be computed as follows:

$$T_i = \arg\inf_{T} \quad c(S_{t_i}, T) + \beta E[q(S_{t_{i+1}})]. \tag{7}$$

where $0 < \beta < 1$ and $q(S_{t_i}) = \gamma T_i^{\text{greedy}}$. The following describes a TANS algorithm based on approximate dynamic programming for Markovian signals:

Algorithm 3: An approximate dynamic programming-based sampling function for the Markovian signal of (1) can be summarized as follows:

- Step i,0: Compute $\hat{\theta}_{S_{t_i}}$, $P_e(S_{t_i})$ and $q(\widehat{S}_{t_{i+1}})$.
- Step i,1: Compute $T_i = \min c(S_{t_i}, T|\hat{\theta}_{S_{t_i}}) + \beta q(\widehat{S}_{t_{i+1}})$.
- Step i,2: Take a sample at time $t_i + T_i$.
- Step i,3: Compute S_{t_i+1}. Repeat.

V. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed sampling schemes by simulations and compare their performance against uniform sampling (in uniform sampling, for a given rate $R = 1/T_{\text{uni}}$ where T_{uni} is not an integer number, the ith sample is taken at time $t_i = \text{round}(T_{\text{uni}})$.).

Figure 3 shows rate-distortion curves achieved by simulations for greedy TANS and compares it with analytical lower bound and various uniform sampling schemes. We assume $\alpha_0 = 0.01$ and $\alpha_1 = 0.99$. We use a maximum likelihood estimation block with $m = 10$ to estimate the state of the underlying Markov chain. For greedy TANS, we use generalized linear prediction (GLP) filter as the reconstruction method. Note that, this reconstruction is causal. For uniform sampling, we use three reconstruction methods: causal line-connecting (CLC), non-causal line-connecting (NCLC) and GLP filtering. Note that, the proposed greedy TANS outperforms other uniform sampling schemes (even the uniform sampling with non-causal reconstruction method). Also, the proposed greedy TANS performs closely to the analytical lower bound curve.

Figure 4 illustrates performance of a TANS scheme based on approximate dynamic programming for a Markovian signal explained in Algorithm 3. Here, we assume that underlying Markov chain transition probabilities are 0.1 (i.e., $p_{01} = p_{10} = 0.1$). The signal parameters are assumed to be $\alpha_0 = 0.7$ and $\alpha_1 = 0.99$. As illustrated in this figure, a TANS scheme based on ADP outperforms the greedy one.

VI. CONCLUSIONS

In this paper, we introduced a new framework for an adaptive nonuniform sampling scheme called *time-stampless adaptive nonuniform sampling*, TANS. The key idea of this framework is that *time increments between samples are computed by using a function of previously taken samples*. Therefore, keeping sampling times (time stamps), except initialization times, is not necessary. We introduced two methods to design sampling functions: a greedy method, and a method based on dynamic programming. We analyzed the performances of these schemes by computing (or bounding) their trade-offs between sampling rate and expected reconstruction distortion for Markovian signals.

VII. REFERENCES

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

[2] F. A. Marvasti, *Nonuniform Sampling: Theory and Practice*. Plenum Publishers Co., 2001.

[3] D. Wei, “Sampling based on local bandwidth,” in *Master Thesis, MIT*, 2007.

[4] J. D. Gibson, “Sequentially adaptive backward prediction in ADPCM speech coders,” *IEEE Trans. Commun.*, vol. 26, no. 1, pp. 145–150, Jan. 1978.

[5] V. K. Goyal, J. Zhuang, and M. Vetterli, “Transform coding with backward adaptive updates,” *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1623–1633, Jul. 2000.

[6] S. Feizi, V. Goyal, and M. Médard, “Locally adaptive sampling,” in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. IEEE, 2010, pp. 152–159.

[7] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995.