A Modulo-Based Architecture for Analog-to-Digital Conversion

Or Ordentlich, Gizem Tabak, Pavan Kumar Hanumolu, Andrew C. Singer, and Gregory W. Wornell

***Abstract*—Systems that capture and process analog signals must first acquire them through an analog-to-digital converter. While subsequent digital processing can remove statistical correlations present in the acquired data, the dynamic range of the converter is typically scaled to match that of the input analog signal. The present paper develops an approach for analog-to-digital conversion that aims at minimizing the number of bits per sample at the output of the converter. This is attained by reducing the dynamic range of the analog signal by performing a modulo operation on its amplitude, and then quantizing the result. While the converter itself is universal and agnostic of the statistics of the signal, the decoder operation on the output of the quantizer can exploit the statistical structure in order to unwrap the modulo folding. The performance of this method is shown to approach information theoretical limits, as captured by the rate-distortion function, in various settings. An architecture for modulo analog-to-digital conversion via ring oscillators is suggested, and its merits are numerically demonstrated.**

I. Introduction

Analog-to-digital converters (ADCs) are an essential component in any device that manipulates analog signals in a digital manner. While digital systems have benefited tremendously from scaling, their analog counterparts have become increasingly challenging. Consequently, it is often the case that the ADC constitutes the main bottleneck in a system, both in terms of power consumption and real estate, and in terms of the quality of the system's output. Developing more efficient ADCs is therefore of great interest [1], [2].

The quality of an ADC is measured via the tradeoff between various parameters such as power consumption, size, cost of manufacturing, and the distortion between the input signal and its digitally-based representation. For the sake of a unified, technology-independent, discussion, it is convenient to restrict the characterization of an ADC quality to three basic parameters: 1) The number of analog samples per second F_S; 2) The number of "raw" output bits R the ADC produces per sample (before subsequent possible lossless compression); 3) The mean squared error (MSE) distortion D between the input signal and a reconstruction that is based on the output of the ADC.

While different applications may require different tradeoffs between F_S, R and D, it is always desirable to design the ADC such that all three parameters are as small as possible. The focus of this work is on the quantization rate R. For a given sampling frequency F_S, and a given target distortion D, our goal is to design ADCs that use the smallest possible number of raw output bits per sample.

The problem of analog-to-digital conversion can be seen as an instance of the lossy source coding/lossy compression problem [3]–[5], as the output of an ADC is a binary sequence, which represents the analog source. A unique key feature of the analog-to-digital conversion problem is that the encoding of the source is carried out in the analog domain, while the decoding procedure is purely digital. Given the limitations of analog processing, it is therefore generally only practical to exploit the source structure at the decoder. Hence, the type of source coding schemes that are suitable for data conversion, are those that approach fundamental limits without requiring knowledge of the source structure at the encoder. In addition, latency and complexity constraints in data conversion, typically preclude the use of schemes other than those based on scalar quantization.

The input signal to an ADC is often known to have structure that could be exploited to reduce the overall bit rate of its representation, R. In our analysis, it will be convenient to express this structure using a stochastic model for the input. Consequently, throughout the paper, we will model the input to the ADC as a stationary stochastic Gaussian process $X(t)$, whose power spectral density (PSD) encapsulates the assumed structure. More generally, we will sometimes also consider the problem of analog-to-digital conversion of a vector $\mathbf{X}(t) = \{X_1(t), \ldots, X_K(t)\}$ of jointly stationary stochastic Gaussian processes, via K parallel ADCs, the input to each one of them is one of the K processes.

Under such stochastic modeling, rate-distortion theory [3] provides the fundamental lower bound $F_s \cdot R > R_X(D)$ for any ADC (and corresponding decoder) that achieves distortion D, where $R_X(D)$ is the rate-distortion function of the process $X(t)$ in bits per second. In general, achieving the rate-distortion function of a source requires using sophisticated high-dimensional quantizers, whereas analog-to-digital conversion is invariably done via scalar uniform quantizers. Thus, achieving this lower bound with ADCs seems overly optimistic. Nevertheless, as we shall see, approaching the

Manuscript received December 8, 2017; revised April 16, 2018 and July 16, 2018; accepted July 17, 2018. Date of publication August 6, 2018; date of current version September 27, 2018. This work was supported in part by the Israel Science Foundation under Grant 1791/17, in part by the National Institutes of Health under Grant EB025327A, in part by SONIC, and in part by the U.S. National Science Foundation under Grant CCF-1717610. The guest editor coordinating the review of this paper and approving it for publication was Dr. Yonina Eldar. *(Corresponding author: Or Ordentlich.)*

O. Ordentlich is with the Rachel and Selim Benin School of Computer Science and Engineering, Hebrew University of Jerusalem, Jerusalem 91904, Israel (e-mail: or.ordentlich@mail.huji.ac.il).

G. Tabak, P. K. Hanumolu, and A. C. Singer are with the Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, IL 61801-2918 USA (e-mail: tabak2@illinois.edu; hanumolu@illinois.edu; acsinger@illinois.edu).

G. W. Wornell is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: gww@mit.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTSP.2018.2863189

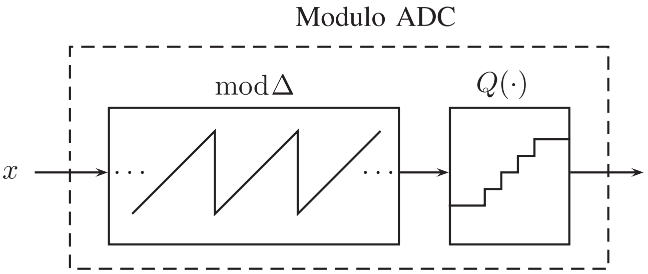


Fig. 1. A schematic illustration of the modulo ADC.

rate-distortion bound, up to some inevitable loss due to the one-dimensional nature of the quantization, is sometimes possible by a simple modification of the scalar uniform quantizer, namely, a *modulo ADC*, followed by a digital decoder that efficiently exploits the source structure.

Instead of sampling and quantizing the process $X(t)$, a modulo ADC samples and quantizes the process $[X(t)] \bmod \Delta$, where the modulo size Δ is a design parameter. See Figure 1. Equivalently, a modulo ADC can be thought of as a standard uniform scalar ADC with step-size δ and an arbitrarily large dynamic range/support, but that outputs only the R least significant bits in the description of each sample, where $2^R = \frac{\Delta}{\delta}$, such that the encoding rate is R. The benefit of applying the modulo operation on $X(t)$ is in reducing its dynamic range/support, which in turn enables a reduction of the number of bits per sample produced by the ADC, without increasing the quantizer's step-size. This operation, which corresponds to disregarding coarse information about $X(t)$, will otherwise substantially degrade the source reconstruction. However, by properly accounting for the modulo operation and appropriately choosing its parameter Δ, we can unwrap the modulo operation with high probability using previous samples of $X(t)$ and exploiting the (redundant) structure in the signal.

Following standard system design methodology, in the performance analysis of a modulo ADC, we distinguish between two events: 1) The no-overload event $\bar{\mathcal{E}}_{\mathrm{OL}}$ where the decoder was able to correctly unwrap the modulo operation. We require the MSE distortion, conditioned on this event, to be at most D; 2) The overload event $\mathcal{E}_{\mathrm{OL}}$ where the decoder fails in unwrapping the modulo operation. We require the probability of this event $\Pr(\mathcal{E}_{\mathrm{OL}})$ to be small, but do not concern ourselves with the MSE distortion conditioned on the occurrence of this event.

A. Our Contributions

This work further develops the modulo ADC framework in three complementary directions, as specified below.

1) Oversampled Modulo ADC: We show that a modulo ADC can be used as an alternative to $\Sigma\Delta$ converters. A $\Sigma\Delta$ converter is based on oversampling the input process $X(t)$, i.e., sampling above the Nyquist rate, in conjunction with noise-shaping, which pushes much of the energy of the quantization noise to high frequencies, where there is no signal content. See Figure 2. The noise shaping operation requires incorporating an elaborate mixed signal feedback circuit. In particular, the circuit first generates the quantization noise, which necessitates using not only an ADC, but also an accurately-matched digital-to-analog converter (DAC), and then applies an analog filter. The analog nature of the signal processing makes it challenging to use filters of high-orders, which in turn limits performance.

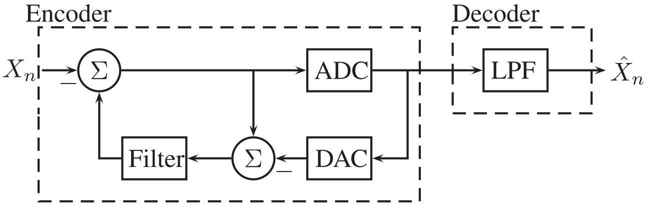


Fig. 2. Schematic architecture for oversampled $\Sigma\Delta$ converter. $\{X_n\}$ is obtained by sampling the process $X(t)$.

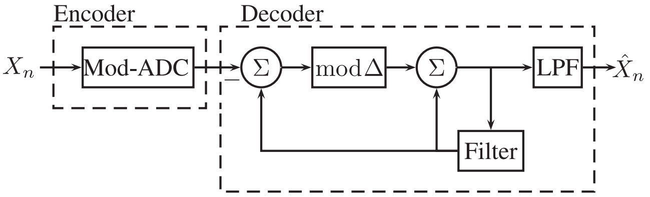


Fig. 3. Schematic architecture for oversampled modulo ADC. The same architecture, without the low-pass filter (LPF) is also suitable for modulo ADC for a general stationary process. $\{X_n\}$ is obtained by sampling the process $X(t)$.

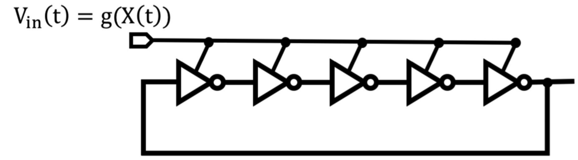


Fig. 4. A schematic illustration of a ring oscillator with $N = 5$ inverters. The states of all N inverter are measured every T_S seconds.

We develop an alternative architecture (Section III) that shifts much of the complexity to the decoder, whereas the "encoder" is simply a modulo ADC. See Figure 3. The parameter Δ in the modulo ADC, as well as the coefficients of the prediction filter in Figure 3, depend only on the bandwidth B of the input process $X(t)$ and on its variance σ^2, and not on the other details of its PSD. Similarly, the MSE distortion between the input process and its reconstruction, depends only on B and σ^2. Thus, the developed architecture is as agnostic as $\Sigma\Delta$ converters to the statistics of the input process. Furthermore, for a flat-spectrum process, the distortion is within a small gap, due to one-dimensionality of the encoder, from the information theoretic limit.

2) A Phase-Domain Implementation of Modulo ADC via Ring Oscillators: We develop a modulo ADC implementation that performs the modulo reduction inherently as part of the analog signal acquisition process. As the phase of a periodic waveform is always measured modulo 2π, a natural class of candidates are ADCs that first convert the input voltage into phase, and then quantize that phase. A notable representative within this class, which has been extensively studied in the literature [6], [7], is the *ring oscillator ADC*.

Consider a closed-loop cascade of N inverters, where N is an odd number, all controlled with the same voltage $V_{dd} = V_{\text{in}}$, see Figure 4. This circuit, which will be described in detail in Section IV, oscillates between $2N$ states, corresponding to the values ('low' or 'high', represented by '0' or '1') of each of the N inverters. See Figure 5. The oscillation frequency is controlled by V_{dd}. Due to the oscillating nature of the circuit, if we sample its state every T_S seconds, we cannot tell how many "state

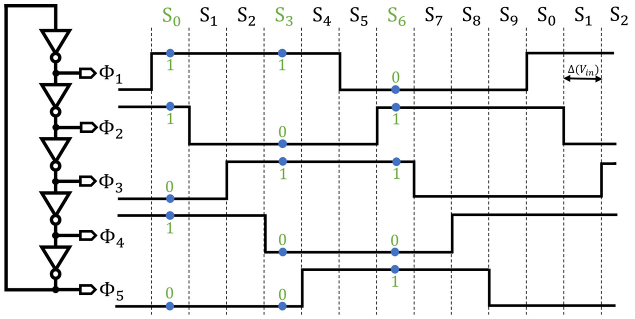


Fig. 5. An example of the evolution of the states of the inverters in a ring oscillator.

changes" occurred between two consecutive samples, but we are able to determine this number modulo $2N$. Thus, by setting V_{dd} to $V_{dd}(t) = g(X(t))$, where $X(t)$ is the analog signal to be converted to a digital one and $g(\cdot)$ is a function to be specified, we obtain a modulo ADC. The input-output relation of this modulo ADC is characterized in Section IV, and depends on the response time of the inverters to change in their input, as a function of V_{dd}.

In practice, the modulo operation realized in this way deviates from the ideal characteristic of Figure 1 in a variety of ways. Accordingly, we perform several numerical experiments to evaluate and optimize the performance of an oversampled ring oscillator modulo ADC, and compare it to the performance of an ideal modulo ADC as well as to a $\Sigma\Delta$ converter. The results demonstrate that despite the non-idealities in the ring oscillator implementation, in some regimes, this architecture holds substantial potential for improvement over existing ADCs.

3) Modulo ADCs for Jointly Stationary Processes: There is great interest in designing efficient ADCs for applications where the number of sensors/antennas observing a particular process is greater than the number of degrees-of-freedom (per time unit) governing its behavior. Thus, there is a redundancy at the receiver that can be exploited. However, as this redundancy can be spread across time and space, traditional ADC architectures, as well as the modulo ADC architectures described in Section II-A and II-B, are insufficient. In this part of the paper, we show how to address this problem via a natural extension of the modulo ADC framework.

As an example we will consider the problem of wireless communication. It is by now well established that using receivers, as well as transmitters, with multiple antennas, dramatically increases the achievable communication rates over wireless channels [8], [9]. However, adding antennas comes with the price of requiring multiple expensive and power hungry RF chains. For traditional ADC architectures, power and cost scale linearly with the number of receive antennas, which motivates an alternative solution.

It is often the case, that the signals observed by the different receive antennas are highly correlated, in time and in space. As an illustrative example, consider the case where the transmitter has one antenna, whereas the receiver has $K > 1$ antennas. We can model the signal observed at each of the antennas, after sampling, as

$$Y_n^k = h_n^k * X_n + Z_n^k,\ k = 1, \dots, K,\ n = 1, \dots, N, \quad (1)$$

where $\{X_n\}$ is the process emitted by the transmitter, $\{h_n^k\}$ is the kth channel impulse response, and $\{Z_n^k\}$ are independent additive white Gaussian noise (AWGN) processes.

Since all K output processes $\{Y_n^1\}, \dots, \{Y_n^K\}$ in (1) are noisy and filtered versions of of the same input process, they will typically be highly correlated. However, this correlation may be spread in time (the n-axis) and in space (the k-axis). As an extreme example, assume $\{X_n\}$ is an iid process, and the filters simply incur different delays, i.e., $h_n^k = \delta_{n-k}$ for $k = 1, \dots, K$. While each individual process $\{Y_k^n\}$ is white, and each vector $(Y_n^1, \dots, Y_n^K)$, $n = 1, \dots, N$ has a scaled identity covariance matrix, the vector process $\{\{Y_n^1\}, \dots, \{Y_n^K\}\}$ is highly correlated. One must therefore jointly process the time and the spatial dimensions in order to exploit this correlation.

This phenomenon, where the signals observed by the different ADCs are highly correlated, is not unique to the wireless communication setup, and appears in many other applications, e.g., multi-array radar. It is, however, taken to the extreme in *massive MIMO* [10], where the number of antennas at the base station is of the order of tens or even hundreds, while the number of users it supports may be substantially fewer.

In Section VI we develop an architecture that uses modulo ADCs, one for each receive antenna, in order to exploit the space-time correlation of the processes. We develop a low-complexity decoding algorithm for unwrapping the modulo operations. This algorithm combines the idea of performing prediction in time, of the quantized vector process from its past, with that of integer-forcing source decoding [11], which is used for exploiting spatial correlations in the prediction error vector. See Figure 6. In the limit of small D, the excess-rate of the developed analog-to-digital conversion scheme with respect to the information theoretic lower bound, is shown to reduce to that of the integer-forcing source decoder.

B. Related Work

The idea of using modulo ADCs/quantizers for exploiting temporal correlations within the input process $X(t)$ towards reducing the quantization rate R, dates back, at least, to [12], where a quantization scheme, called modulo-PCM, was introduced. A decoding scheme for unwrapping the modulo operation, based on maximum-likelihood sequence detection [13], was further proposed in [12], and a heuristic analysis was performed, based on prediction of $X(t)$ from its past, which shows that modulo-PCM can approach the Shannon lower bound under the high-resolution assumptions. In Section II-A, we develop a more complete analysis of modulo quantization, the details of which are required for the application we discuss in Section III.

The architecture from Figure 3 is based on using a prediction filter at the decoder, as a part of the modulo unwrapping process, as was hinted at in [12] (see also [14]). In agreement with the literature on differential pulse-code modulation (DPCM) at the late 1970s (see e.g. [15]), the authors in [12] proposed to design the prediction filter as the optimal one-step predictor of the unquantized process $\{X_n\}$ from its past. As shown in [16], this design criterion is sub-optimal, and the "correct" design criterion is to take this filter as the one-step predictor of the *quantized* process from its past. The difference between the two design criteria is significant for oversampled processes, which are the focus of Section III, whose PSD is zero at high frequencies, as in those frequencies the signal-to-distortion ratio

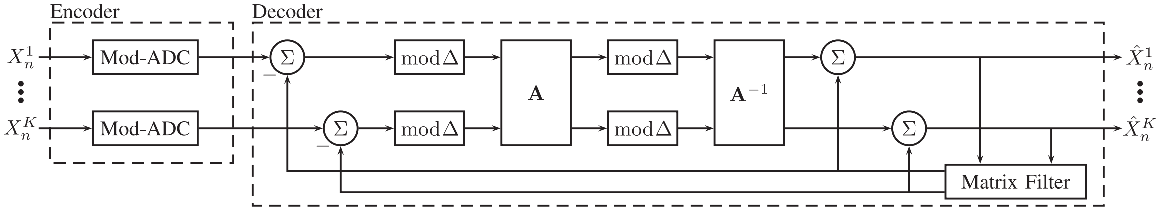


Fig. 6. Schematic architecture for Modulo ADCs for jointly stationary processes.

is zero, no matter how small the quantization noise is. Our analysis in Section III reveals that designing the modulo size Δ and the prediction filter with respect to a quantized flat-spectrum input process, results in a *universal* system. This means, that this system attains the same distortion D for all input processes that share the same support for the PSD and the same variance.

The use of modulo ADCs/quantizers was also studied by Boufounos in the context of quantization of oversampled signals [17] (see also [18]). In particular, it is shown in [17] that by randomly embedding a measurement vector in $\mathbb{R}^K$ onto an $M \gg K$ dimensional subspace, and using a modulo ADC for quantizing each of the coordinates of the result, one can attain a distortion that decreases exponentially with the oversampling ratio, with high probability. In Section III we consider a similar setup, where an oversampled analog signal, with oversampling ratio $L > 1$, i.e. F_s is L times greater than the Nyquist frequency, is digitized by a modulo ADC. In the language of [17], this corresponds to embedding $\mathbf{X} \in \mathbb{R}^K$ to an $M = LK$ dimensional space by zero-padding followed by interpolation, which is indeed a linear operation. We show that for this particular "embedding" not only is the decay of MSE distortion exponential in the oversampling ratio, but the attained distortion is information-theoretically optimal, up to a constant loss, which is explicitly characterized, due to the scalar nature of the quantizer. Moreover, under this "embedding", a simple low-complexity decoding algorithm exists, whereas for the random projection case studied in [17], no computationally efficient decoding algorithm was given. One advantage, on the other hand, of the approach from [17], is that it is applicable to 1-bit modulo ADCs, whereas the performance of the scheme from Section III typically becomes attractive starting from $R \gtrsim 2$ bits per sample, due to reasons that will become clearer in the sequel (see discussion around eq. (12)).

Very recently, Bhandari *et al.* have addressed the question of what is the minimal sampling rate that allows for exact recovery of a bandlimited finite-energy signal, from its modulo-reduced sampled version [19] (see also [20]). They have found that a sufficient condition for correct reconstruction is sampling above the Nyquist rate by a factor of $2\pi e$, regardless of the size of the modulo interval. The analysis in [19] did not take quantization noise into account, which corresponds to $R = \infty$ and $D = 0$ in our setup.

The merits of a modulo ADC for distributed analog-to-digital conversion of signals correlated in space, but not in time, were demonstrated in [11]. A low-complexity decoding algorithm, for unwrapping the modulo operation, was proposed and its performance was analyzed. It was demonstrated via numerical experiments that the performance is usually quite close to the information theoretic lower bounds (See also [21]). In Section II-B, we summarize the decoding scheme from [11] and the corresponding performance analysis, as those will be needed in Section VI, where we develop a modulo ADC architecture for analog-to-digital conversion of jointly stationary processes. The decoding algorithm for this setup, as well as its performance analysis, is inspired by the ideas and techniques from Sections II-A and II-B.

As modulo reduction can be viewed as a one dimensional deterministic instance of *binning*, in a broader sense, modulo quantization is closely related to Wyner-Ziv's source coding with side information setup and to its channel coding dual, which is the Gel'fand-Pinsker setup [22]. In the latter context, we further note that modulo quantization is widely used for communication over intersymbol interference channels [23], [24]. Recently, Hong and Caire [25] considered modulo ADCs as potential candidates for the front end of receivers in a cloud radio access network (CRAN), employing compute-and-forward [26] based protocols.

Note that the although the concept of modulo ADC is reminiscent of *folding ADCs* [27], an important difference is that unlike the latter, the former does not keep track of the number of folds that occurred and, moreover, its functionality does not depend on this number, i.e., it does not saturate for large inputs. In unwrapping the modulo operation at the decoder, the missing information about number of folds is recovered, and we are able to attain the same D with smaller rate.

Finally, another related line of work, is that of compressed sampling, see, e.g., [28]–[30], where the goal is to design universal and efficient ADCs with a small sampling frequency F_S, under the assumption that the input signal occupies only a small portion of its total bandwidth, but the exact support is unknown.

C. Organization

The rest of the paper is organized as follows. In Section II we formally define the modulo ADC and study its performance for stationary scalar input processes, and for random vectors (spatial correlation). Section III develops the use of oversampled modulo ADCs as a substitute for $\Sigma\Delta$ converters, and analyzes the tradeoffs this architecture achieves. In Section IV we introduce an implementation of modulo ADCs via ring oscillators and establish the corresponding input-output mathematical model. Numerical experiments for evaluating the performance of ring oscillators based oversampled modulo ADCs are performed in Section V. Section VI proposes to use parallel modulo ADCs for digitizing jointly stationary processes. The paper concludes in Section VII.

II. Preliminaries on Ideal Modulo ADC

Let $\Delta \in \mathbb{R}^+$ be a positive number, and define the $\mathrm{mod}\Delta$ operation as

$$[x] \bmod \Delta \triangleq x - \Delta \left\lfloor \frac{x}{\Delta} \right\rfloor \in [0, \Delta),$$

where the floor operation $\lfloor x \rfloor$ returns the largest integer smaller than or equal to x. By definition, we have that for any $x, y \in \mathbb{R}$ and $\Delta > 0$

$$[[x] \bmod \Delta + y] \bmod \Delta = [x + y] \bmod \Delta. \tag{2}$$

An R-bit modulo ADC with resolution parameter α, or (R, α) *mod-ADC*, maps a real input $x \in \mathbb{R}$ to R bits, by computing

$$[x]_{R,\alpha} \triangleq [\lfloor \alpha x \rfloor] \bmod 2^R \in \{0, 1, \ldots, 2^R - 1\},$$

and producing a binary representation of it. Note that $\lfloor \alpha x \rfloor$ is the output of an infinite support scalar quantizer with step size $1/\alpha$, and $[x]_{R,\alpha}$ is a wrapped version of it. In the sequel we will demonstrate that in various scenarios an appropriately designed decoder can recover $\lfloor \alpha x \rfloor$ from its wrapped version $[x]_{R,\alpha}$, with high probability, based on temporal/spatial correlations of the ADCs input signal.

We can write $[x]_{R,\alpha}$ as

$$[x]_{R,\alpha} = [\alpha x + (\lfloor \alpha x \rfloor - \alpha x)] \bmod 2^R = [\alpha x + z] \bmod 2^R. \tag{3}$$

The term $z = \lfloor \alpha x \rfloor - \alpha x \in (-1, 0]$ in (3), is the quantization error of a uniform scalar quantizer $\lfloor \alpha x \rfloor$, and is clearly a deterministic function of x. Nevertheless, throughout this paper we will model z as additive uniform noise $Z \sim \mathrm{Unif}((-1, 0])$ statistically independent of x, such that the (R, α) mod-ADC will be modeled as a *stochastic channel* with input x and output Y, related as

$$Y = [\alpha x + Z] \bmod 2^R. \tag{4}$$

The modulo additive noise channel model (4) for an (R, α) mod-ADC can be rigorously justified via the use of *subtractive dithers*. Specifically, we can use a random variable $U \sim \mathrm{Unif}([0, 1))$, statistically independent of x, which we refer to as a *dither*, and feed $\tilde{x} = x + U/\alpha$ to the (R, α) mod-ADC instead of feeding x. The output of the modulo ADC in this case will be

$$\begin{aligned} [\tilde{x}]_{R,\alpha} &= [\alpha \tilde{x} + (\lfloor \alpha \tilde{x} \rfloor - \alpha \tilde{x})] \bmod 2^R \\ &= [\alpha x + U + (\lfloor \alpha x + U \rfloor - (\alpha x + U))] \bmod 2^R. \end{aligned}$$

Subtracting U from $[\tilde{x}]_{R,\alpha}$ and reducing the result modulo 2^R, we obtain

$$\begin{aligned} &[[\tilde{x}]_{R,\alpha} - U] \bmod 2^R \\ &= \left[[\alpha x + U + (\lfloor \alpha x + U \rfloor - (\alpha x + U))] \bmod 2^R - U\right] \bmod 2^R \\ &= [\alpha x + (\lfloor \alpha x + U \rfloor - (\alpha x + U))] \bmod 2^R, \end{aligned}$$

where the last equality follows from the distributive law of modulo (2). Note that for every $x \in \mathbb{R}$, the random variable $Z = \lfloor \alpha x + U \rfloor - (\alpha x + U)$ is uniformly distributed over $(-1, 0]$, and is therefore independent of x [31, Lemma 1]. Thus, with subtractive dithers, the additive noise model (4) is exact. We note that even when dithering is not used, under suitable conditions this model predicts performance quite accurately [32].

Although the modulo operation entails loss of information in general, in many situations it is possible to unwrap it, i.e., reconstruct $\alpha x + Z$ from $Y = [\alpha x + Z] \bmod 2^R$ with high probability.[1] In particular, let

$$\tilde{Y} = \left[Y + \frac{1}{2}2^R\right] \bmod 2^R - \frac{1}{2}2^R, \tag{5}$$

and note that conditioned on the *no-overload* event

$$\mathcal{E}_{\overline{\mathrm{OL}}} \triangleq \left\{\alpha x + Z \in \left[-\frac{1}{2}2^R, \frac{1}{2}2^R\right)\right\},$$

we have that $\tilde{Y} = \alpha x + Z$. Thus, if $\Pr(\mathcal{E}_{\overline{\mathrm{OL}}})$ is close to 1, the modulo operation has no effect with high probability. Note that $\Pr(\mathcal{E}_{\mathrm{OL}}) = \Pr\left(|\alpha x + Z| > \frac{1}{2}2^R\right)$ is identical to the probability that a standard uniform quantizer with dynamic range (support) $2^R/\alpha$ is in *overload*. Thus, when thinking of x as a single observation, it is unclear what the advantages of a modulo ADC are with respect to a traditional uniform ADC. However, as we illustrate below, the modulo ADC allows exploitation of the statistical structure of the acquired *signal* in a much more efficient manner than the standard ADC.

The following lemma is proved using Chernoff's bound, and will be useful in the sequel for bounding $\Pr(\mathcal{E}_{\overline{\mathrm{OL}}})$ in various scenarios.

Lemma 1 ([33, Lemma 4], [34, Theorem 7]): Consider the random variable $Z_{\mathrm{eff}} = \sum_{\ell=1}^{L} \alpha_\ell Z_\ell + \sum_{k=1}^{K} \beta_k U_k$ where $\{Z_\ell\}_{\ell=1}^{L}$ are iid Gaussian random variables with zero mean and some variance σ_z^2, $\{U_k\}_{k=1}^{K}$ are iid random variables, statistically independent of $\{Z_\ell\}_{\ell=1}^{L}$, uniformly distributed over the interval $[-\rho/2, \rho/2)$ for some $\rho > 0$, and $\{\alpha_\ell\}_{\ell=1}^{L}$ and $\{\beta_k\}_{k=1}^{K}$ are arbitrary real (deterministic) numbers. Let $\sigma_{\mathrm{eff}}^2 \triangleq \mathbb{E}(Z_{\mathrm{eff}}^2) = \sigma_z^2 \sum_{\ell=1}^{L} \alpha_\ell^2 + \frac{\rho^2}{12} \sum_{k=1}^{K} \beta_k^2$. Then for any $\tau > 0$

$$\Pr(Z_{\mathrm{eff}} > \tau) = \Pr(Z_{\mathrm{eff}} < -\tau) \leq \exp\left\{-\frac{\tau^2}{2\sigma_{\mathrm{eff}}^2}\right\}.$$

A. Modulo ADCs for Scalar Stationary Processes

In this subsection we consider the case where an (R, α)-mod ADC, as described above, is applied on a scalar stationary process. We develop a corresponding decoder and analyze its performance, including the effects of the choices of α and R.

Let $\{X_n\}$ be a zero-mean discrete-time stationary Gaussian stochastic process, obtained by sampling a stationary Gaussian process $X(t)$ every T_S seconds. Let

$$Y_n = [\alpha X_n + Z_n] \bmod 2^R, \; n = 1, 2, \ldots$$

be the process obtained by applying an (R, α) mod-ADC on the process $\{X_n\}$, where $\{Z_n\}$ is a $\mathrm{Unif}((-1, 0])$ iid noise, and let

$$V_n = \alpha X_n + Z_n, \; n = 1, 2, \ldots$$

be its non-folded version. Our goal is to design a decoder that recovers V_n from the outputs of the modulo ADC, $\{Y_n\}$, with high probability. To that end, we assume the decoder has access to $\{V_{n-1}, \ldots, V_{n-p}\}$, an assumption that will be justified in the sequel, and that it knows the auto-covariance function $C_X[r] = \mathbb{E}\left[X_n X_{n-r}\right]$ of $\{X_n\}$. We apply the following algorithm (See also Figure 3 for a schematic illustration):

Inputs: Y_n, $\{V_{n-1}, \ldots, V_{n-p}\}$, $\{C_X[r]\}$, R, α.

[1] Here, the term "high probability" is used to state that this probability can be made as high as desired by increasing R. We explicitly quantify the relation between R and the desired "no-overload" probability.

Output: Estimates $\hat{V}_n$, $\hat{X}_n$, for V_n and X_n, respectively.
Algorithm:

1) Compute the optimal linear MMSE predictor for V_n from its last p samples

$$\hat{V}_n^p = \sum_{i=1}^{p} h_i \cdot \left(V_{n-i} + \frac{1}{2}\right) - \frac{1}{2}, \tag{6}$$

where $\{h_n\}$ is a p-tap prediction filter, computed based on $\{C_X[r]\}$ and α, and the shift by $1/2$ compensates for $\mathbb{E}(Z_n)$.

2) Compute

$$W_n = [Y_n - \hat{V}_n^p] \bmod 2^R$$
$$\tilde{W}_n = \left[W_n + \frac{1}{2}2^R\right] \bmod 2^R - \frac{1}{2}2^R.$$

3) Output $\hat{V}_n = \hat{V}_n^p + \tilde{W}_n$, and $\hat{X}_n = \frac{\hat{V}_n + \frac{1}{2}}{\alpha}$.

Remark 1: Note that $\{h_n\}$ is the p-tap prediction filter for the *quantized* process $\{V_n\}$ from its past, rather than for $\{X_n\}$ from its past. While the loss for using the latter, instead of the former, becomes insignificant when high-resolution assumptions apply, it can be arbitrarily large for oversampled processes, for which high-resolution assumptions never hold [16], [35]. The filter coefficients $\{h_n\}$ need only be computed once, and can then be used for all times.

The following proposition characterizes the performance of the algorithm above. All logarithms in this paper are taken to base 2, unless stated otherwise.

Proposition 1: Let $\hat{V}_n^p$, $\hat{V}_n$ and $\hat{X}_n$ be as defined in the algorithm above, and let $\sigma_p^2 = \mathbb{E}(V_n - \hat{V}_n^p)^2$. We have that

$$\Pr(\mathcal{E}_{\mathrm{OL}_n}) \triangleq \Pr(\hat{V}_n \neq V_n) \leq 2\exp\left\{-\frac{3}{2}2^{2\left(R - \frac{1}{2}\log(12\sigma_p^2)\right)}\right\}, \tag{7}$$

and

$$D = \mathbb{E}[(X_n - \hat{X}_n)^2 | \mathcal{E}_{\overline{\mathrm{OL}_n}}] \leq \frac{1}{12\alpha^2(1 - \Pr(\mathcal{E}_{\mathrm{OL}_n}))}, \tag{8}$$

where the event $\mathcal{E}_{\overline{\mathrm{OL}_n}} = \{\hat{V}_n = V_n\}$ is the complement of the event $\mathcal{E}_{\mathrm{OL}_n} = \{\hat{V}_n \neq V_n\}$.

Proof: Let $E_n^p \triangleq V_n - \hat{V}_n^p$ be the pth order prediction error of the process $\{V_n\}$, and note that its variance $\sigma_p^2 = \mathbb{E}(E_n^p)^2$ is invariant to n due to stationarity. We have that

$$\begin{aligned} W_n &= [Y_n - \hat{V}_n^p] \bmod 2^R \\ &= \left[[V_n] \bmod 2^R - \hat{V}_n^p\right] \bmod 2^R \\ &= \left[V_n - \hat{V}_n^p\right] \bmod 2^R \\ &= [E_n^p] \bmod 2^R, \end{aligned} \tag{9}$$

where equation (9) follows from the modulo distributive law (2), and constitutes the key advantage of the modulo operation for exploiting temporal correlations. Note that $\tilde{W}_n \in [-\frac{1}{2}2^R, \frac{1}{2}2^R)$ is a cyclically shifted version of $W_n \in [0, 2^R)$, as in (5). Therefore, conditioned on the event

$$\mathcal{E}_{\overline{\mathrm{OL}_n}} = \left\{|E_n^p| < \frac{1}{2}2^R\right\}$$

we have that $\tilde{W}_n = E_n^p$.

Note that E_n^p is a zero-mean linear combination of statistically independent Gaussian and uniform random variables, such that Lemma 1 applies, and we have that

$$\begin{aligned} \Pr\,(\mathcal{E}_{\mathrm{OL}_n}) &= \Pr(\tilde{W}_n \neq E_n^p) \\ &= \Pr\left(|E_n^P| > \frac{1}{2}2^R\right) \\ &\leq 2\exp\left\{-\frac{2^{2R}}{8\sigma_p^2}\right\} \\ &= 2\exp\left\{-\frac{3}{2}2^{2\left(R - \frac{1}{2}\log(12\sigma_p^2)\right)}\right\}, \end{aligned} \tag{10}$$

Whenever $\mathcal{E}_{\overline{\mathrm{OL}_n}}$ occurs, we have that $\hat{V}_n = V_n$, and consequently

$$\hat{X}_n = X_n + \frac{Z_n + \frac{1}{2}}{\alpha}$$

and

$$\begin{aligned} \mathbb{E}[(X_n - \hat{X}_n)^2 | \mathcal{E}_{\overline{\mathrm{OL}_n}}] &= \mathbb{E}\left[\left(\frac{Z_n + \frac{1}{2}}{\alpha}\right)^2 \middle| \mathcal{E}_{\overline{\mathrm{OL}_n}}\right] \\ &= \frac{1}{\alpha^2}\frac{\mathbb{E}(Z_n + 1/2)^2 - \Pr(\mathcal{E}_{\mathrm{OL}_n})\mathbb{E}[(Z_n + 1/2)^2 | \mathcal{E}_{\mathrm{OL}_n}]}{\Pr(\mathcal{E}_{\overline{\mathrm{OL}_n}})} \\ &\leq \frac{1}{12\alpha^2(1 - \Pr(\mathcal{E}_{\mathrm{OL}_n}))}. \end{aligned} \tag{11}$$

■

Proposition 1 shows that we can make $\Pr(\mathcal{E}_{\mathrm{OL}_n})$ as small as $2e^{-\frac{3}{2}2^{2\delta}}$ by choosing

$$R = \frac{1}{2}\log(12\sigma_p^2) + \delta. \tag{12}$$

For example, taking $\delta = 2$ bits, results in an overload probability smaller than 10^{-10}. In particular, unless we take a very small δ, we have that $1 - \Pr(\mathcal{E}_{\mathrm{OL}_n}) \approx 1$, and consequently, by Proposition 1, we will have $D \approx 1/12\alpha^2$. Thus, to simplify expressions in the analysis that follows, we assume $D = 1/12\alpha^2$. We note the tradeoff in choosing α: on the one hand, increasing α decreases the MSE distortion D, but on the other hand the prediction error variance σ_p^2 of the process $V_n = \alpha X_n + Z_n$ increases with α such that the required rate R for avoiding overload errors increases. Thus, the tradeoff between D and the required quantization rate is controlled through the parameter α. We now turn to characterize the tradeoff the developed scheme achieves.

Let $h(A)$ denote the differential entropy of the random variable A, and $h(A|B)$ the conditional differential entropy of A given the random variable B [5]. Recall that for a stationary Gaussian process $\{X_n\}$ with PSD $S_X(e^{j\omega})$ we have that [36]

$$h(X_n | X_{n-1}, \ldots) = \frac{1}{2\pi}\int_{\pi}^{\pi} \frac{1}{2}\log\left(2\pi e S_X(e^{j\omega})\right) d\omega, \tag{13}$$

and in particular $h(X_n|X_{n-1},\ldots)=-\infty$ if and only if $S_X(e^{j\omega})=0$ over a measurable subset of $[-\pi,\pi)$. Shannon's lower bound [3], states that the number of bits per sample R produced by any quantizer that attains an MSE distortion D must satisfy

$$R(D)\geq R_{\text{SLB}}(D)\triangleq h(X_n|X_{n-1},\ldots)-\frac{1}{2}\log(2\pi eD).$$

It is well-known [3] that for Gaussian processes with finite $h(X_n|X_{n-1},\ldots)$, Shannon's lower bound is asymptotically tight, i.e., $\lim_{D\to 0}R(D)-R_{\text{SLB}}(D)=0$.

Proposition 2: If $h(X_n|X_{n-1},\ldots)>-\infty$, then

$$\lim_{D\to 0}\lim_{p\to\infty}\frac{1}{2}\log(12\sigma_p^2)=R_{\text{SLB}}(D).$$

Proof: We can write

$$\frac{1}{2}\log(12\sigma_p^2)=\frac{1}{2}\log\left(\frac{\frac{\sigma_p^2}{\alpha^2}}{\frac{1}{12\alpha^2}}\right)=\frac{1}{2}\log\left(\frac{\mathbb{E}(E_n'^p)^2}{D}\right). \tag{14}$$

where $E_n'^p$ is the pth order prediction error of the process $X_n+Z_n/\alpha=X_n+\sqrt{D}\tilde{Z}_n$, where $\tilde{Z}_n\sim\text{Unif}([-\sqrt{12},0))$ iid.

For a Gaussian process $\{X_n\}$, the condition $h(X_n|X_{n-1},\ldots)>-\infty$ is equivalent to

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\log\left(S_X(e^{j\omega})\right)d\omega>-\infty. \tag{15}$$

As a consequence of (15), we have that

$$\begin{aligned}\lim_{D\to 0}\frac{1}{2\pi}\int_{-\pi}^{\pi}&\frac{1}{2}\log\left(2\pi e\left(S_X(e^{j\omega})+D\right)\right)d\omega\\&=h(X_n|X_{n-1},\ldots).\end{aligned} \tag{16}$$

By Paley-Wiener's theorem [37], we have that

$$\lim_{p\to\infty}\mathbb{E}(E_n'^p)^2=2^{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left(S_X(e^{j\omega})+D\right)d\omega}. \tag{17}$$

Combining (16) and (17), we obtain that

$$\lim_{D\to 0}\lim_{p\to\infty}\mathbb{E}(E_n'^p)^2=2^{2h(X_n|X_{n-1},\ldots)-2\pi e},$$

for processes with finite entropy rate $h(X_n|X_{n-1},\ldots)$. The result now follows by rearranging terms. ■

For the practically important case where $\{X_n\}$ is obtained by oversampling the process $\{X(t)\}$, which is studied in Section III, the assumption $h(X_n|X_{n-1},\ldots)>-\infty$ of Proposition 2 does not hold. Nevertheless, we will show that the modulo ADC achieves performance that is close to the information theoretic limits.

Above, we have assumed that the decoder has access to the non-folded samples $\{V_{n-1},\ldots,V_{n-p}\}$. To justify this assumption, an *initialization* step is needed, where the decoder acquires the first p consecutive samples $\{V_1,\ldots,V_p\}$, or estimates of these samples. Once those are obtained, we can apply the algorithm described above, sample-by-sample, and assume the estimate $\hat{V}_n$ produced by the algorithm at time n is correct, and can be used as an input for the algorithm in the next p steps. All samples $V_{p+1},\ldots,V_N$ will be recovered correctly, as long as no overload error occurred within the $N-p$ decoding steps. Thus, by the union bound, we see that the first $N-p$ samples are recovered correctly with probability at least $1-2Ne^{-\frac{3}{2}2^{2\delta}}$.[2]

One conceptually simple way of performing the initialization, i.e., obtaining $\{V_1,\ldots,V_p\}$ is by using a standard scalar quantizer with high-rate for the first p samples. Although the high power consumption of such a quantizer will have a negligible effect on the total power consumption, due to the fact it is used only for a small fraction of the time, this approach has the disadvantage of having to include two ADCs, a high-rate standard ADC and a modulo ADC withing the system. Alternatively, one can perform the initialization using only an R bit modulo ADC in one of the two following ways:

1) Increase α gradually until it reaches its final value. For the first sample, α_1 will be chosen such that $V_1=\alpha_1X_1+Z_1$ is w.h.p. within the modulo interval, such that no prediction is needed. Next, we can use V_1 in order to predict $V_2=\alpha_2X_2+Z_2$, which allows to use $\alpha_2>\alpha_1$ such that the prediction error is still within the modulo interval. Continuing this way, we can keep increasing α until convergence.
2) We can collect a long vector of outputs from the modulo ADC, say $\{Y_1,\ldots,Y_K\}$, $K>p$, and unwrap the modulo operation via the integer-forcing source coding scheme described in the next subsection. The amount of computations per sample required in this method is greater than that of the "steady state", i.e., after initialization is complete, but since initialization is rarely performed, the effect on the total complexity is negligible.

B. Modulo ADCs for Random Vectors

In this subsection we consider the case where K *identical* (R,α)-mod ADCs, as described above, are applied on a random vector, one on each component of the vector. We develop a corresponding decoder and analyze its performance, including the effects of the choices of α and R.

Let $\mathbf{X}\sim\mathcal{N}(\mathbf{0},\mathbf{\Sigma})$ be a K-dimensional Gaussian random vector with zero mean and covariance matrix $\mathbf{\Sigma}$. Let

$$Y_k=[\alpha X_k+Z_k]\bmod 2^R,\ k=1,\ldots K,$$

be obtained by applying K identical (R,α) mod-ADCs, each applied to a different coordinate of the vector $\mathbf{X}$, where the quantization noises $Z_k\sim\text{Unif}((-1,0])$, $k=1,\ldots,K$, are iid, and let

$$V_k=\alpha X_k+Z_k,\ k=1,\ldots K,$$

be its non-folded version. Our goal is to recover $\mathbf{V}\triangleq[V_1,\ldots,V_K]^T$ from the outputs $\mathbf{Y}\triangleq[Y_1,\ldots,Y_K]^T$ of the modulo ADCs with high probability.

To that end, we now review a sub-optimal low-complexity decoder, proposed in [11], dubbed the *integer-forcing* (IF) source decoder, see Figure 7. Let $\frac{\mathbf{1}}{\mathbf{2}}$ be a K-dimensional vector with all entries equal to $\frac{1}{2}$, and $\mathbf{I}$ be the identity matrix. The decoding algorithm works as follows.

Inputs: $\mathbf{Y}$, $\mathbf{\Sigma}$, R, α.

Output: Estimates $\hat{\mathbf{V}}_{\text{IF}}$, and $\hat{\mathbf{X}}_{\text{IF}}$, for $\mathbf{V}$ and $\mathbf{X}$, respectively.

Algorithm:

[2] Note that conditioning on the event that no overload error occurred until time n, changes the statistics of E_n^p. Thus, applying the union bound correctly here requires some more care. See [35] for more details.

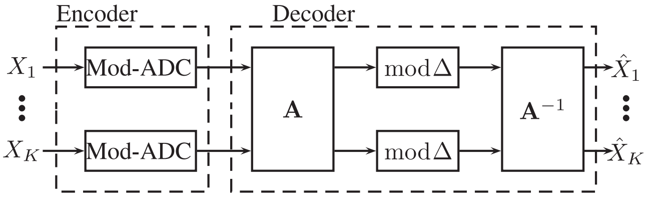


Fig. 7. Schematic architecture for modulo ADC for random vectors.

1) Solve

$$\begin{aligned}\mathbf{A} &= [\mathbf{a}_1|\cdots|\mathbf{a}_K]^T \\ &= \operatorname*{argmin}_{\substack{\bar{\mathbf{A}}\in\mathbb{Z}^{K\times K}\\ |\bar{\mathbf{A}}|\neq 0}} \max_{k=1,\ldots,K} \frac{1}{2}\log\left(\bar{\mathbf{a}}_k^T\left(\mathbf{I}+12\alpha^2\boldsymbol{\Sigma}\right)\bar{\mathbf{a}}_k\right),\end{aligned} \tag{18}$$

where $|\mathbf{A}|$ denotes the absolute value of $\det(\mathbf{A})$.

2) For $k = 1,\ldots,K$, compute

$$\begin{aligned}\bar{g}_k &\triangleq \left[\mathbf{a}_k^T\left(\mathbf{Y}+\frac{\mathbf{1}}{\mathbf{2}}\right)\right] \bmod 2^R, \\ \tilde{g}_k &\triangleq \left[\bar{g}_k + \frac{1}{2}2^R\right] \bmod 2^R - \frac{1}{2}2^R,\end{aligned} \tag{19}$$

and set $\tilde{\mathbf{g}} = [\tilde{g}_1,\ldots,\tilde{g}_K]^T$.

3) Output $\hat{\mathbf{V}}_{\text{IF}} = \mathbf{A}^{-1}\tilde{\mathbf{g}}$, and $\hat{\mathbf{X}}_{\text{IF}} = \frac{\hat{\mathbf{V}}_{\text{IF}}}{\alpha}$.

Remark 2: The optimization problem (18) requires a computational complexity exponential in K, in general (unless P=NP). However, the problem of finding the optimal integer matrix $\mathbf{A}$, need only be solved once for each covariance matrix $\boldsymbol{\Sigma}$ and α. Thus, even if the solution to this problem is computationally expensive, its cost is normalized by the number of times this solution is used. In practice, one can apply the LLL algorithm [38] in order to obtain a sub-optimal $\mathbf{A}$ with polynomial complexity in K.

The next proposition, adapted from [11, Theorem 2] characterizes the performance of modulo ADCs with the decoder above.

Proposition 3: Let $\mathbf{A} = [\mathbf{a}_1|\cdots|\mathbf{a}_K]^T$ be the matrix found in step 1 of the algorithm above, and define

$$R_{\text{IFSC}}(\mathbf{A}) = \max_{k=1,\ldots,K} \frac{1}{2}\log\left(\mathbf{a}_k^T\left(\mathbf{I}+12\alpha^2\boldsymbol{\Sigma}\right)\mathbf{a}_K^T\right). \tag{20}$$

We have that

$$\Pr(\mathcal{E}_{\text{OL}}) = \Pr(\hat{\mathbf{V}}_{\text{IF}} \neq \mathbf{V}) \leq 2K\exp\left\{-\frac{3}{2}\cdot 2^{2(R-R_{\text{IFSC}}(\mathbf{A}))}\right\},$$

and

$$D_k = \mathbb{E}\left[\left(X_k - \hat{X}_{k,\text{IF}}\right)^2 \middle| \mathcal{E}_{\overline{\text{OL}}}\right] \leq \frac{1}{12\alpha^2(1-\Pr(\mathcal{E}_{\text{OL}}))},$$

for all $k = 1,\ldots,K$, where the event $\mathcal{E}_{\overline{\text{OL}}} = \{\hat{\mathbf{V}}_{\text{IF}} = \mathbf{V}\}$ is the complement of the event $\mathcal{E}_{\text{OL}} = \{\hat{\mathbf{V}}_{\text{IF}} \neq \mathbf{V}\}$.

The main idea behind the decoder above is the simple observation that for any vector $\mathbf{a} = [a_1,\ldots,a_k]^T \in \mathbb{Z}^K$ and any vector $\mathbf{h} = [h_1,\ldots,h_K]^T \in \mathbb{R}^K$ we have that

$$\left[\sum_{k=1}^K a_k[h_k] \bmod 2^R\right] \bmod 2^R = \left[\sum_{k=1}^K a_k h_k\right] \bmod 2^R. \tag{21}$$

Proof: By the identity (21), we have that the quantities $\bar{g}_k$, computed in step 2 of the algorithm, satisfy

$$\bar{g}_k = \left[\mathbf{a}_k^T\left(\mathbf{Y}+\frac{\mathbf{1}}{\mathbf{2}}\right)\right] \bmod 2^R = [g_k] \bmod 2^R,$$

where

$$g_k \triangleq \mathbf{a}_k^T\left(\mathbf{V}+\frac{\mathbf{1}}{\mathbf{2}}\right).$$

Furthermore, $\tilde{g}_k \in [-\frac{1}{2}2^R, \frac{1}{2}2^R)$ is merely a cyclically shifted version of $\bar{g}_k \in [0, 2^R)$. Thus, $\tilde{g}_k = g_k$ if and only if $g_k \in [-\frac{1}{2}2^R, \frac{1}{2}2^R)$. Consequently, $\hat{\mathbf{V}}_{\text{IF}} \neq \mathbf{V}$ if and only if the event

$$\mathcal{E}_{\text{OL}} = \bigcup_{k=1}^K \left\{|g_k| \geq \frac{1}{2}2^R\right\},$$

occurs. Thus, by the union bound,

$$\Pr(\mathcal{E}_{\text{OL}}) = \Pr(\hat{\mathbf{V}}_{\text{IF}} \neq \mathbf{V}) \leq \sum_{k=1}^K \Pr\left(|g_k| \geq \frac{1}{2}2^R\right). \tag{22}$$

The random variable g_k has zero mean, variance $\sigma_k^2 = \mathbf{a}_k^T\left(\alpha^2\boldsymbol{\Sigma}+\frac{1}{12}\mathbf{I}\right)\mathbf{a}_k$, and satisfies the conditions of Lemma 1. We therefore have that

$$\begin{aligned}\Pr\left(|g_k| \geq \frac{1}{2}2^R\right) &\leq 2\exp\left\{-\frac{2^{2R}}{8\sigma_k^2}\right\} \\ &= 2\exp\left\{-\frac{3}{2}\cdot 2^{2\left(R-\frac{1}{2}\log(12\sigma_k^2)\right)}\right\} \\ &= 2\exp\left\{-\frac{3}{2}\cdot 2^{2\left(R-\frac{1}{2}\log\left(\mathbf{a}_k^T\left(\mathbf{I}+12\alpha^2\boldsymbol{\Sigma}\right)\mathbf{a}_k\right)\right)}\right\}.\end{aligned}$$

Substituting this into (22) and recalling the definition of $R_{\text{IFSC}}(\mathbf{A})$, gives

$$P_e \leq 2K\exp\left\{-\frac{3}{2}\cdot 2^{2(R-R_{\text{IFSC}}(\mathbf{A}))}\right\}. \tag{23}$$

Conditioned on the event $\mathcal{E}_{\overline{\text{OL}}}$, i.e., the event that $\mathcal{E}_{\text{OL}}$ did not occur, we have that for all $k = 1,\ldots,K$

$$\begin{aligned}D_k &= \mathbb{E}\left[\left(X_k - \hat{X}_{k,\text{IF}}\right)^2 \middle| \mathcal{E}_{\overline{\text{OL}}}\right] \\ &= \mathbb{E}\left[\left(\frac{Z_k+\frac{1}{2}}{\alpha}\right)^2 \middle| \mathcal{E}_{\overline{\text{OL}}}\right] \leq \frac{1}{12\alpha^2(1-\Pr(\mathcal{E}_{\text{OL}}))},\end{aligned}$$

where the last inequality follows similarly to (11). ■

As in the previous subsection, we set

$$R = R_{\text{IFSC}}(A) + \delta, \tag{24}$$

such that

$$\Pr(\mathcal{E}_{\text{OL}}) \leq 2K\exp\left\{-\frac{3}{2}\cdot 2^{2\delta}\right\}, \tag{25}$$

and set $D = 1/12\alpha^2$, which is a good approximation for the upper bound we derived on D_k, provided that δ is not too small. Consequently, we can write

$$R_{\text{IFSC}}(\mathbf{A}, D) \triangleq \max_{k=1,\ldots,K} \frac{1}{2}\log\left(\mathbf{a}_k^T\left(\mathbf{I} + \frac{1}{D}\mathbf{\Sigma}\right)\mathbf{a}_k\right). \quad (26)$$

The tradeoff between rate, distortion and error probability achieved by the (R, α) mod-ADC with an integer-forcing decoder is therefore characterized by equations (24), (25), and (26). To put this result in context, we recall the information theoretic benchmark [11]

$$R_{\text{bench}}^{\text{BT}}(D) \triangleq \frac{1}{2K}\log\left|\mathbf{I} + \frac{1}{D}\mathbf{\Sigma}\right|,$$

that approximates the minimal quantization rate, per quantizer, required by any computationally and delay unlimited system in order to achieve MSE of at most D in the reconstructions of each X_k, $k = 1, \ldots, K$. Thus,

$$R_{\text{IFSC}}(\mathbf{A}, D) - R_{\text{bench}}^{\text{BT}}(D) = \frac{1}{2}\log\left(\frac{\max_{k=1,\ldots,K} \mathbf{a}_k^T\left(\mathbf{I} + \frac{1}{D}\mathbf{\Sigma}\right)\mathbf{a}_k}{\left|\mathbf{I} + \frac{1}{D}\mathbf{\Sigma}\right|^{\frac{1}{K}}}\right). \quad (27)$$

It is easy to show that the right hand side of (27) is non-negative [11]. However, typically it is possible to find an integer matrix $\mathbf{A}$ for which the gap is quite small, and under certain distributions of practical interest on $\mathbf{\Sigma}$, the cumulative distribution function (CDF) of this gap can be characterized [21]. A comprehensive comparison between $R_{\text{IFSC}}(D)$ and $R_{\text{bench}}^{\text{BT}}(D)$ was performed in [11], and it was demonstrated that they are usually quite close.

III. Oversampled Modulo-ADC

In Section II-A we have demonstrated the effectiveness of the modulo ADC architecture for acquiring stochastic processes that are correlated in time. In particular, we have shown that the performance of a modulo ADC depends on the variance of the prediction error of the process $\{V_n = \alpha X_n + Z_n\}$, rather than the variance of V_n itself. However, when designing an ADC, it is desirable to impose as few constraints as possible on the signals that will be fed to the ADC. Therefore, assuming that $\{X_n\}$ is such that $\{V_n\}$ is highly predictable may be too restrictive.

Nevertheless, recalling that the process $\{X_n\}$ is obtained by sampling a continuous-time process $X(t)$, we observe that if the sampling rate is higher than Nyquist's rate, $\{X_n\}$ will be bandlimited,[3] and consequently, $\{V_n\}$ will be highly predictable no matter what the precise PSD of $\{X_n\}$ happens to be. In fact, this observation can be viewed as the rationale underlying $\Sigma\Delta$-conversion. In particular, a $\Sigma\Delta$-converter is information theoretically equivalent to a differential pulse-code modulator (DPCM) whose input is a bandlimited signal with flat spectrum [35].

While having many advantages, the implementation of $\Sigma\Delta$ converters is more involved than that of traditional scalar uniform quantizers. The main challenge in the design of $\Sigma\Delta$ converters is the need to produce the quantization error, and then apply a filter to this analog signal. A major obstacle is that the generation of the quantization error requires to first quantize the current sample, then apply a digital-to-analog converter (DAC) to produce the analog representation of the quantizer's output, and finally to subtract this representation from the original sample. See Figure 2. The quantizer and the DAC need to be matched as otherwise the produced quantization error is inaccurate. This, however, turns out to be quite difficult to achieve, unless the quantizer is a simple sign detector (1-bit quantizer).

To circumvent the challenges listed above, we develop an oversampled modulo ADC architecture, as an alternative to $\Sigma\Delta$-conversion. The *only* assumptions made on the input process $\{X(t)\}$ is that it is bandlimited with maximal frequency at most B, and that its variance is at most σ^2. The developed universal architecture is as follows. See Figure 3.

Analog-to-digital conversion: The process $X(t)$ is uniformly sampled every $T_S = 1/2LB$ seconds, $L > 1$, such that the sampling rate is L times above Nyquist's rate. Each sample of the obtained discrete-time process $\{X_n\}$ is then discretized using an (R, α) mod-ADC, resulting in the quantized process $\{Y_n = [\alpha X_n + Z_n] \bmod 2^R\}$.

As above, we define the unfolded process $\{V_n = \alpha X_n + Z_n\}$. The decoding procedure assumes $\{V_{n-1}, \ldots, V_{n-p}\}$ are given, and computes an estimate for V_n, based on Y_n.

Inputs: Y_n, $\{V_{n-1}, \ldots, V_{n-p}\}$, σ^2, L, R, α.

Outputs: Estimates $\hat{V}_n$ and $\hat{X}_n$ for V_n and X_n, respectively.

Algorithm: The algorithm is exactly the same as that in Section II-A, with only one difference. Here $\{C_X[r]\}$ is unknown. Thus, for the computation of the p-tap prediction filter $\{h_n\}$, we assume the PSD of $\{X_n\}$ is

$$S_X(e^{j\omega}) = \begin{cases} L\sigma^2 & \omega \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right) \\ 0 & \omega \notin \left[-\frac{\pi}{L}, \frac{\pi}{L}\right) \end{cases}, \quad (28)$$

even though this assumption may, and is most likely to, be wrong.

Final post-processing: After collecting a long sequence of estimates $\{\hat{X}_1, \ldots, \hat{X}_N\}$ we apply a non-causal low pass filter

$$G(e^{j\omega}) = \begin{cases} \dfrac{12\alpha^2 L\sigma^2}{1 + 12\alpha^2 L\sigma^2} & \text{if } \omega \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \\ 0 & \text{if } \omega \notin \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \end{cases}$$

on them, to obtain the sequence $\{\hat{X}_1^{\text{LPF}}, \ldots, \hat{X}_N^{\text{LPF}}\}$.

The advantages over $\Sigma\Delta$ conversion are clear: the only processing done in the analog domain is sampling and applying a modulo ADC, whereas all filtering operations are done digitally at the decoder.

Proposition 1 provides an upper bound on the error probability $\Pr(\mathcal{E}_{\text{OL}_n}) = \Pr(\hat{V}_n \neq V_n)$ in terms of $R - \frac{1}{2}\log(12\sigma_p^2)$. However, Proposition 2, which characterizes the scaling of $\frac{1}{2}\log(12\sigma_p^2)$ with D, does not apply here for two reasons. The first is that we use a mismatched prediction filter here, due to the unknown PSD of $\{X_n\}$, and the second is that whatever the exact PSD truns out to be, it is assumed to be supported on the frequency interval $[-\frac{\pi}{L}, \frac{\pi}{L}]$, such that $h(X_n|X_{n-1}, \ldots) = -\infty$, and the high-resolution assumption never holds. Instead, we prove the following.

[3] We say that a discrete-time process $\{X_n\}$ is bandlimited, if there exists some $\gamma < \pi$ such that $S_X(e^{j\omega}) = 0$ for all $\omega \in (-\pi, -\gamma) \cup (\gamma, \pi)$. Since our analysis takes quantization noise into account, it is quite robust to slight deviations from the assumption that $S_X(e^{j\omega})$ is strictly band limited. In particular, as long as $S_X(e^{j\omega}) \ll D$, for all $\omega \in (-\pi, -\gamma) \cup (\gamma, \pi)$, where D is the target MSE distortion, our analysis remains valid.

Proposition 4: Let $\{X_n\}$ be a zero-mean stationary process with variance $\mathbb{E}(X_n^2) \leq \sigma^2$ and PSD supported in frequency interval $[-\frac{\pi}{L}, \frac{\pi}{L}]$. Let $V_n = \alpha X_n + Z_n$ where $Z_n \sim \text{Unif}([-1, 0))$, and $\hat{V}_n^p$ be as in (6), where $\{h_n\}$ is the optimal linear MMSE p-tap prediction filter for V_n, from its past samples $\{V_{n-1}, \ldots, V_{n-p}\}$, designed under the assumption that $S_X(e^{j\omega})$ is as in (28). Then

$$\lim_{p\to\infty} 12\sigma_p^2 \leq \left(1 + 12\alpha^2 L\sigma^2\right)^{\frac{1}{L}}.$$

Proof: Let

$$S_{\tilde{V}}(e^{j\omega}) = \begin{cases} \alpha^2 L\sigma^2 + 1/12 & \omega \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \\ 1/12 & \omega \notin \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \end{cases}, \tag{29}$$

and let $H_p(e^{j\omega})$ be the frequency response of the prediction filter $\{h_n\}$, which is designed with respect to (29). Further, let $H(e^{j\omega}) = \lim_{p\to\infty} H_p(e^{j\omega})$. By the basic principles of optimal linear MMSE prediction, we have that

$$S_{\tilde{V}}(e^{j\omega})|1 - H(e^{j\omega})|^2 = 2^{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(S_{\tilde{V}}(e^{j\omega}))d\omega}. \tag{30}$$

Therefore, combining (29) and (30), we see that

$$|1 - H(e^{j\omega})|^2 = \begin{cases} \left(1 + 12\alpha^2 L\sigma^2\right)^{\frac{1}{L}-1} & \omega \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \\ \left(1 + 12\alpha^2 L\sigma^2\right)^{\frac{1}{L}} & \omega \notin \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \end{cases}. \tag{31}$$

Applying this filter on the "actual" process $V_n = \alpha X_n + Z_n$, whose PSD is

$$S_V(e^{j\omega}) = \begin{cases} \alpha^2 S_X(e^{j\omega}) + 1/12 & \omega \in \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \\ 1/12 & \omega \notin \left[-\frac{\pi}{L}, \frac{\pi}{L}\right] \end{cases},$$

we get

$$\begin{aligned} \lim_{p\to\infty} 12\sigma_p^2 &= \lim_{p\to\infty} 12\mathbb{E}(V_n - \hat{V}_n^p)^2 \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} S_V(e^{j\omega})|1 - H(e^{j\omega})|^2 d\omega \\ &= \frac{\left(1 + 12\alpha^2 L\sigma^2\right)^{\frac{1}{L}}}{2\pi}\left[\int_{\omega\notin[-\pi/L,\pi/L]} 1 d\omega \right. \\ &\quad \left. + \int_{-\pi/L}^{\pi/L} \left(1{+}12\alpha^2 L\sigma^2\right)^{-1}\left(1{+}12\alpha^2 S_X(e^{j\omega})\right) d\omega\right] \\ &\leq \left(1 + 12\alpha^2 L\sigma^2\right)^{\frac{1}{L}}, \end{aligned} \tag{32}$$

where the last inequality follows from our assumption that $\frac{1}{2\pi}\int_{-\pi/L}^{\pi/L} S_X(e^{j\omega})d\omega = \mathbb{E}(X_n^2) \leq \sigma^2$. ■

It follows from Proposition 1 combined with Proposition 4, that for large p and a quantization rate of roughly

$$R = \delta + \frac{1}{L}\frac{1}{2}\log\left(1 + 12\alpha^2 L\sigma^2\right), \tag{33}$$

the proposed system achieves $\Pr(\mathcal{E}_{\text{OL}_n}) \leq 2\exp\{-\frac{3}{2}2^{2\delta}\}$, for all input processes with bandwidth $\leq B$ and variance $\leq \sigma^2$.

After low-pass filtering with $G(e^{j\omega})$, we get by a similar analysis to that done in Section II-A and in [35], that for long enough N such that the discrete Fourier transform (DFT) of N consecutive samples of $\{X_n\}$ have negligible energy in frequencies above π/L, we have that

$$\begin{aligned} D &= \mathbb{E}\left[(X_n - \hat{X}_n^{\text{LPF}})^2 \middle| \bigcap_{n=1}^{N}\{\hat{V}_n = V_n\}\right] \\ &\leq \frac{\sigma^2}{1 + 12\alpha^2 L\sigma^2}\frac{1}{1 - \Pr\left(\bigcap_{n=1}^{N}\{\hat{V}_n = V_n\}\right)} \\ &\leq \frac{\sigma^2}{1 + 12\alpha^2 L\sigma^2}\frac{1}{1 - N\Pr(\mathcal{E}_{\text{OL}_n})} \\ &\leq \frac{\sigma^2}{1 + 12\alpha^2 L\sigma^2}\frac{1}{1 - 2N\exp\{-\frac{3}{2}2^{2\delta}\}}. \end{aligned} \tag{34}$$

Thus, for large enough δ such that the total overload probability is small, i.e.,

$$2N\exp\left\{-\frac{3}{2}2^{2\delta}\right\} \ll 1, \tag{35}$$

we have that our system achieves distortion $\approx D$ with

$$R = \frac{1}{L}\frac{1}{2}\log\left(\frac{\sigma^2}{D}\right) + \delta. \tag{36}$$

The term $\frac{1}{L}\frac{1}{2}\log(\frac{\sigma^2}{D})$ is the rate-distortion function of a source with PSD as in (28). Thus, up to the loss of δ bits per sample, due to the one dimensional quantizer we are using, whose size is dictated by (35), our system is optimal in the following minimax sense: no system can attain a better tradeoff between R and D simultaneously for all processes with bandwidth at most B and variance at most σ^2.

The multiplicative increase in quantization rate of the developed system, with respect to the fundamental rate-distortion limit, is $(\frac{1}{2}\log(\frac{\sigma^2}{D}) + L\delta)/(\frac{1}{2}\log(\frac{\sigma^2}{D}))$. If $X(t)$ were sampled at its Nyquist rate, rather than L times above it, standard uniform scalar quantization would have achieved similar overload probability and distortion with only a $(\frac{1}{2}\log(\frac{\sigma^2}{D}) + \delta)/(\frac{1}{2}\log(\frac{\sigma^2}{D}))$ multiplicative increase in rate with respect to the fundamental limit. Thus, oversampling combined with the architecture developed here produces a total number of bits-per-second which is greater than that required by an ADC operating at the Nyquist rate. The disadvantage of the latter approach is that it requires to use a high-resolution quantizer for each sample, whereas the scheme developed here, allows to reduce the number of quantization bits per sample, at the expanse of an increased sampling rate. Thus, just like $\Sigma\Delta$ conversion, the scheme developed here allows to replace slow but high-resolution ADCs, with fast low-resolution ones.

IV. Implementation via Ring Oscillators

In this Section we develop an architecture for a circuit implementing a modulo ADC, and provide a mathematical model for its input-output characteristic. Our implementation is essentially based on converting the input voltage into phase, which can naturally only be observed modulo 2π, and then quantizing the phase. To that end, we use *ring oscillator ADCs*, as described next.

Consider a closed-loop cascade of N inverters, where N is an odd number, all controlled with the same voltage V_{dd}, see Figure 4. This circuit, which is referred to as a ring oscillator can act as an ADC with sampling period T_s, when V_{dd} is set to $V_{in}(t) = g(X(t))$, where $X(t)$ is the analog signal to be converted to a digital one and $g(\cdot)$ is a function to be specified, and the state ('0' or '1', corresponding to 'low' or 'high') of each inverter is measured every T_s seconds.

It is well known that the time it takes for a non-ideal inverter's output to respond to a change in its input is a function of V_{dd} [39], which we denote by $\Delta(V_{dd}) > 0$. Taking this delay into account, a moment of reflection reveals that at each time instance, exactly one pair of adjacent inverters are at the same state whereas all other pairs of adjacent inverters are at distinct states. Denote by $I \in \{1, \ldots, N\}$ the index of the first inverter within the pair that shares the same state, and denote its state by $B \in \{0, 1\}$, i.e., the adjacent pair of inverters with the same state are inverter I and inverter $[I + 1] \bmod N$, and their state is B. With this notation, we can uniquely identify the states of all N inverters at time t with the number $Q_t = (I_t - 1) + N \cdot [I_t + B_t] \bmod 2 \in \{0, \ldots, 2N - 1\}$. See Figure 5. By sampling the states of all N inverters every T_s seconds, we gain access to the discrete-time process $\{Q_{nT_s}\}$.

A crucial observation is that the process Q_t cyclically oscillates in increments of $+1$ modulo $2N$. More formally stated, if $t' > t$ is the earliest time where $Q_{t'} \neq Q_t$, then $Q_{t'} = [Q_t + 1] \bmod 2N$. We designate by V_n the number of increments that occurred in the process $\{Q_t\}$ within the time interval $[nT_S, (n + 1)T_s)$, and define the output of the induced modulo ADC as

$$Y_n \triangleq [V_n] \bmod 2N = [Q_{(n+1)T_s} - Q_{nT_s}] \bmod 2N.$$

Next, we relate V_n to the process $V_{in}(t)$. To this end, we make the simplifying assumption that $X(t)$ is constant within each time interval $[nT_s, (n + 1)T_s)$, and consequently, so is $V_{in}(t)$. This assumption can be made exact by adding a sample-and-hold circuit to the system. Assuming the function $\Delta(V_{dd})$ is identical for all N inverters, we have that

$$Q_{nT_s} = \left[\left\lfloor \sum_{k=-\infty}^{n-1} \frac{T_s}{\Delta(V_{in}(kT_s))} \right\rfloor\right] \bmod 2N,$$

and consequently,

$$\begin{aligned} Y_n = &\left[\left[\left\lfloor \sum_{k=-\infty}^{n} \frac{T_s}{\Delta(V_{in}(kT_s))} \right\rfloor\right] \bmod 2N \right. \\ &\left. - \left[\left\lfloor \sum_{k=-\infty}^{n-1} \frac{T_s}{\Delta(V_{in}(kT_s))} \right\rfloor\right] \bmod 2N \right] \bmod 2N \\ = &\left[\left\lfloor \sum_{k=-\infty}^{n} \frac{T_s}{\Delta(V_{in}(kT_s))} \right\rfloor - \left\lfloor \sum_{k=-\infty}^{n-1} \frac{T_s}{\Delta(V_{in}(kT_s))} \right\rfloor\right] \\ &\bmod 2N, \end{aligned}$$

where the last equality follows from the modulo distributive law (2). Defining the "quantization error"

$$Z_n = \left\lfloor \sum_{k=-\infty}^{n} \frac{T_s}{\Delta(V_{in}(kT_s))} \right\rfloor - \sum_{k=-\infty}^{n} \frac{T_s}{\Delta(V_{in}(kT_s))} \in (-1, 0],$$

we can write

$$\begin{aligned} Y_n &= \left[\sum_{k=-\infty}^{n} \frac{T_s}{\Delta(V_{in}(kT_s))} + Z_n \right. \\ &\qquad \left. - \sum_{k=-\infty}^{n-1} \frac{T_s}{\Delta(V_{in}(kT_s))} - Z_{n-1} \right] \bmod 2N \\ &= \left[\frac{T_s}{\Delta(V_{in}(nT_s))} + Z_n - Z_{n-1} \right] \bmod 2N. \end{aligned}$$

Let us now define the function

$$f(x) = \frac{1}{\Delta(x)},$$

which corresponds to the oscillation frequency of our circuit, and is dictated by the characteristics of the inverters at hand, and let us also take the function $g(\cdot)$ to be affine, such that $V_{in}(t) = a + bX(t)$. We further define the discrete time process $X_n = X(nT_s)$, for all $n \in \mathbb{N}$. We have therefore obtained the model

$$Y_n = [T_s \cdot f(a + bX_n) + Z_n - Z_{n-1}] \bmod 2N. \qquad (37)$$

In general, the quantization noise process $\{Z_n\}$ is a deterministic function of the process $\{X_n\}$. Nevertheless, as in the analysis of the ideal modulo ADC, in the sequel we make the simplifying assumption that it is an iid process with $Z_n \sim \mathrm{Unif}((-1, 0])$.

If $f(\cdot)$ were an affine function itself, with an appropriate choice of the parameters a, b we could have induced the model

$$Y_n = [\alpha X_n + Z_n - Z_{n-1}] \bmod 2^R,$$

where $R = \log(2N)$, which is identical to the ideal (R, α) mod-ADC, up to the fact that the quantization noise $Z_n - Z_{n-1}$ is now a first order moving-average (MA) process rather than a white process. In practice, however, it is difficult to construct inverters for which $f(\cdot)$ is approximately affine within a large range. The effect of nonlinearities of $f(\cdot)$ on the performance of the modulo ADC is numerically studied in the next section.

V. Numerical Experiments

We have conducted numerical simulations for the performance of a ring oscillator based modulo ADC, where the input is an oversampled process, as in Section III. In our simulations, we have assumed that the inverters were produced using a CMOS technology. The corresponding function $f(V_{\text{in}})$ relating the input voltage to the output frequency of the oscillator, which was introduced in Section IV, is shown in Figure 8, as obtained using a PSpice simulation.

A. Design of System Parameters

In all our simulations, we have designed the modulo ADC and the corresponding decoder as described in Section III, i.e., under the assumption that the input signal $X(t)$ is a Gaussian

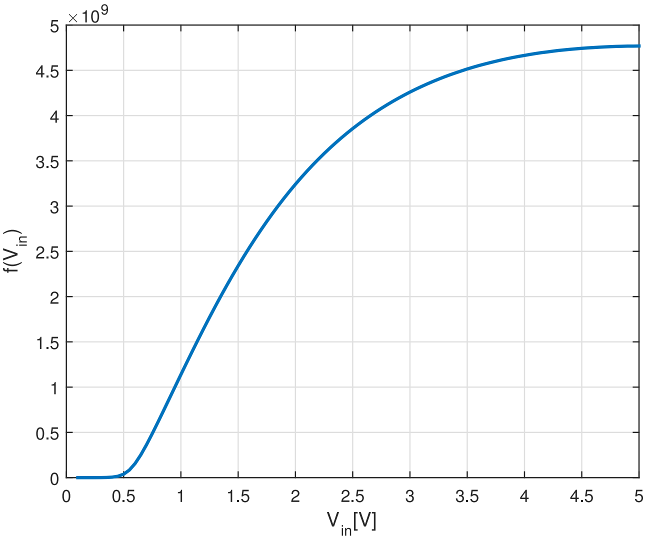


Fig. 8. The voltage to output frequency function $f(V_{\text{in}})$.

stationary process with zero mean and variance σ^2, whose PSD is flat within the frequency interval $[-B, B]$ and zero outside this interval. The sampling rate is a factor of $L > 1$ above the Nyquist rate, such that the sampling period is $T_s = \frac{1}{2LB}$ seconds.

Given the oversampling ratio L, the number of inverters N, and the above assumptions on the statistics of $X(t)$, the design of the modulo ADC and its corresponding decoder consists of:

1) Choosing the shift and scaling parameters a and b for the modulo ADC such that $V_{\text{in}}(t) = a + bX(t)$;
2) Designing the p-tap prediction filter $\{h_n\}$ for $V_n = T_s f(a + bX_n) + Z_n - Z_{n-1}$ given the past samples $\{V_{n-1}, \ldots, V_{n-p}\}$;
3) Designing a $2k+1$-tap noncausal smoothing filter $\{g_n\}$ for estimating X_n from $\{V_{n-k}, \ldots, V_{n+k}\}$.

The decoding procedure consists of recovering an estimate $\{\hat{V}_n\}$ for $\{V_n\}$ from the modulo ADC's outputs $\{Y_n = [T_s f(a + bX_n) + Z_n - Z_{n-1}] \bmod 2N\}$, by applying the decoding procedure described in Section III with the prediction filter $\{h_n\}$. Then, the estimate $\{\hat{X}_n\}$ is produced by applying the smoothing filter $\{g_n\}$ to the process $\{\hat{V}_n\}$, which is referred to as final post-processing in Section III . The filters $\{h_n\}$ and $\{g_n\}$ are chosen as the MMSE-optimal linear prediction and smoothing filters, respectively. Calculating the coefficients of $\{h_n\}$ requires knowledge of the second-order statistics of the process $\{V_n\}$. This in turn, can be (numerically) calculated from the pairwise distribution of $\{X_n, X_{n-m}\}$, $m = 0, \ldots, p$, which is fully characterized by our assumption that $\{X_n\}$ is a Gaussian process with PSD $S_X(e^{j\omega})$ as in (28). Calculating the coefficients of $\{g_n\}$ requires, in addition, the joint second-order statistics of the processes $\{X_n, V_n\}$, which can either be calculated numerically, or via Bussgang's Theorem [40].

We apply the developed modulo ADC architecture to processes of length T discrete samples. The parameters a and b are chosen as follows: Let $P_e = \Pr(\cup_{t=1}^{T} \hat{V}_t \neq V_t)$ be the block error probability of our decoder, and let ϵ be our target block error probability. For every a and b, we find the filters $\{h_n\}$ and $\{g_n\}$ as described above, and compute the corresponding $P_e = P_e(a, b)$ via Monte Carlo simulation for a Gaussian input process with PSD as in (28). Among all (a, b) for which $P_e(a, b) < \epsilon$, we choose the pair that results in the smallest MSE distortion $\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}(X_t - \hat{X}_t)^2$. The target block error probability for all of the setups we consider is $\epsilon = 10^{-3}$, and the block length we consider is $T = 2^{11}$. Roughly, these parameters correspond to allowing a per-sample overload error probability of $10^{-3} \cdot 2^{-11} \approx 4.89 \cdot 10^{-7}$.

B. Evaluation Method

The system was designed for a bandlimited Gaussian process with a flat PSD. Nevertheless, we would like it to achieve approximately the same MSE distortion and error probability for all bandlimited processes with the same variance, regardless of the PSD within that band. For an ideal modulo ADC and large p, this is indeed the case, as shown in Section III. To test to what extent this remains the case also for the ring oscillator based modulo ADC, we apply our system on two types of processes: 1) A Gaussian process with variance σ^2 and bandwidth B, whose PSD is flat within this band, for which the system was designed; 2) A sinusoidal waveform, whose frequency is chosen at random, uniformly on $[0, B)$, and whose amplitude is $\sqrt{2\sigma^2}$, such that its power is σ^2.

For each experiment, we also plot the theoretical performance of an ideal (R, α) mod-ADC, as well as those of a first-order $\Sigma\Delta$ (with the optimal 1-tap noise shaping filter) converter, both designed to achieve the same target block error probability for the bandlimited Gaussian stochastic process $X(t)$. Although overload errors have a different effect on $\Sigma\Delta$ converters and modulo ADCs, both systems fail to achieve their target distortions unless those are avoided.

In the ADC literature, it is quite common to measure the performance of a particular ADC for a sinusoidal input. One drawback of this approach is that the deterministic nature of the input signal allows to design the ADC such that overload errors *never* occur, without significantly increasing its dynamic range above the standard deviation of its input. For stochastic processes, even if Gaussianity is assumed, the dynamic range must be as large as multiple standard deviations of its input, in order to ensure a small overload probability. In our derivations, this is manifested through the rate backoff parameter δ, which dictates the ratio between the quantizer's dynamic range 2^R and the standard deviation of its input (which in our case is the prediction error processes).

In order to allow a unified presentation of the results for both Gaussian and sinusoidal processes, rather than plotting the rate $R_{\text{mod-ADC}}(D)$ required by the modulo ADC in order to achieve an MSE distortion D with target block error probability ϵ, we plot $R_{\text{mod-ADC}}(D) - \delta$, where

$$\delta = \frac{1}{2}\log\left(-\frac{2}{3}\ln\left(\frac{\epsilon}{2T}\right)\right). \tag{38}$$

This is consistent with traditional converter analyses that separate saturation effects from granularity ones [4], [37]. For our parameters $T = 2^{11}$, $\epsilon = 10^{-3}$, (38) evaluates to $\delta \approx 1.6717$ bits. Note that by (12), δ is the rate backoff required in order to attain block error probability below ϵ by an ideal modulo ADC, when the input process is Gaussian. A similar analysis reveals that the same rate backoff is also required for a $\Sigma\Delta$ converter to attain the same block error probability, under the same assumptions on the input process [35]. Thus, in all figures we also plot $R_{\Sigma\Delta}(D) - \delta$ rather than $R_{\Sigma\Delta}(D)$, where $R_{\Sigma\Delta}(D)$ is the rate needed by the $\Sigma\Delta$ converter to attain distortion D with block error probability below ϵ.

C. Results and Discussion

We have performed experiments for the parameters $L = 3$ and four different values of B: 100 Hz, 44.1 KHz, 100 KHz and 1 MHz. The value of σ^2 is immaterial, as it can be absorbed in the parameter b. The results are depicted in Figures 9(a), (b), (c) and (d), respectively. The results are based on Monte Carlo simulation, with 10^3 independent trials for each point in each figure. No overload errors were observed for the choices of a, b, $\{h_n\}$ and $\{g_n\}$ that correspond to each point in the figures, neither for the Gaussian processes and neither for the sinusoidal processes.

In general, the results indicate that the ring oscillator implementation of a modulo ADC is closer to the ideal modulo ADC for small bandwidths B and quantization rates R. In all figures we observe the same trend: for small enough R the curve of the SNR as a function of R for the ring oscillator modulo ADC is parallel to that of the ideal modulo ADC, and has a slope of $\approx 6L = 18$ dB/bit, in agreement with (36). Then, for large enough R the system's non-linearities "kick-in" and the slope significantly decreases. Eventually, for large enough R, the first-order $\Sigma\Delta$ converter outperforms the ring oscillator modulo ADC, as can be observed in Figure 9(d). Nevertheless, for moderate values of R, even for $B = 1$ MHz, the improvement over the $\Sigma\Delta$ converter can be as large as 17 dB.

The trends above are to be expected. Recall that the output of the corresponding modulo ADC is given by (37). If $b \cdot \sigma$ is small enough, the function $f(a + bX_n)$ resides in a small interval around $f(a)$ with high probability, and is well approximated by the linear function $f(a) + bf'(a)X_n$. Consequently, the output of the modulo ADC can be well approximated as

$$Y_n \approx [T_s b f'(a) X_n + Z_n - Z_{n-1} + T_s f(a)] \bmod 2N.$$

Since $T_s f(a)$ is known and can be removed, this is equivalent to a $(T_s b f'(a), \log(2N))$ mod-ADC, albeit with quantization noise $Z_n - Z_{n-1}$ rather than Z_n.

Typically, however, in order to get a large gain from using a modulo ADC rather than a standard uniform quantizer, we would like to use an (R, α) mod-ADC with $\alpha \cdot \sigma \gg \frac{1}{2}2^R$. Thus, in order to get a "useful" modulo ADC that is close to ideal, the two conditions (i) $b \cdot \sigma \ll 1$; (ii) $T_s f'(a) \cdot b \cdot \sigma \gg N$; should hold. These two conditions can only be satisfied simultaneously if $T_s f'(a) \gg 1$, i.e., when the sampling rate is low, relative to $f'(a)$.

For an ideal (R, α) mod-ADC with a given target overload error probability, as R increases α can also increase, resulting in a smaller distortion. Similarly, for the ring oscillator modulo ADC, the optimal choice of b should, in general, increase with R. For small rates, the optimal value of b is also small, such that the linear approximation for the function $f(\cdot)$ is not too bad. However, as R, and consequently b, increases, the nonlinearities start becoming significant and the slope of the SNR as a function of R becomes smaller.

VI. Modulo ADCs for Jointly Stationary Processes

In this section we develop a scheme that uses K parallel modulo ADCs for digitizing K jointly stationary processes, provide a corresponding low-complexity decoding algorithm, and characterize its performance.

Let $\{X_n^1\}, \ldots, \{X_n^K\}$ be K discrete-time jointly Gaussian stationary random processes, obtained by sampling the jointly Gaussian stationary processes $X_1(t), \ldots, X_K(t)$ every T_s seconds. Let

$$Y_n^k = [\alpha X_n^k + Z_n^k] \bmod 2^R, \; k = 1, \ldots, K, \; n = 1, 2, \ldots$$

be the processes obtained by applying K parallel (R, α) mod-ADCs, on $\{X_n^1\}, \ldots, \{X_n^K\}$, where the input to the kth modulo ADC is the process $\{X_n^k\}$, and $\{Z_n^k\}$ is a $\mathrm{Unif}((-1, 0])$ noise, iid in space and in time. Let

$$V_n^k = \alpha X_n^k + Z_n^k, \; k = 1, \ldots, K, \; n = 1, 2, \ldots$$

be the non-folded version of Y_k^n. Let $\mathbf{X}_n = [X_n^1, \ldots, X_n^K]^T$, and define $\mathbf{Y}_n$, $\mathbf{Z}_n$ and $\mathbf{V}_n$ similarly. Our goal is to recover the process $\{\mathbf{V}_n\}$ from the outputs of the modulo ADCs with high probability.

To achieve this goal, we employ a two-step procedure, combining the schemes from Section II-A and Section II-B: first we compute a predictor $\hat{\mathbf{V}}_n^p$ based on previous p samples $\{\mathbf{V}_{n-1}, \ldots, \mathbf{V}_{n-p}\}$ whose error is the vector $\mathbf{E}_n^p = \mathbf{V}_n - \hat{\mathbf{V}}_n^p$. By the same derivation as in Section II-A, we can produce $[\mathbf{E}_n^p] \bmod 2^R$ from $\mathbf{Y}_n$ and $\{\mathbf{V}_{n-1}, \ldots, \mathbf{V}_{n-p}\}$, where the modulo operation applied to a vector is to be understood as reducing each coordinate modulo 2^R. Now, our task is to decode a modulo-folded correlated random vector, which can be done via the integer-forcing decoder described in Section II-B. This relatively simple decoding procedure allows to efficiently exploit both temporal and spatial correlations. Below we describe it in more detail. See Figure 6. For all $\ell, m \in \{1, \ldots, K\}$, let $C_{\ell m}[r] = \mathbb{E}(X_n^\ell X_{n-r}^m)$.

Inputs: $\mathbf{Y}_n$, $\{\mathbf{V}_{n-1}, \ldots, \mathbf{V}_{n-p}\}$, $\{C_{\ell m}[r]\}$ for all $\ell, m \in \{1, \ldots, K\}$, R, α.

Outputs: Estimates $\hat{\mathbf{V}}_n$ and $\hat{\mathbf{X}}_n$ for $\mathbf{V}_n$ and $\mathbf{X}_n$, respectively.

Algorithm:

1) Compute the optimal linear MMSE predictor for $\mathbf{V}_n$ from its last p samples

$$\hat{\mathbf{V}}_n^p = \sum_{i=1}^{p} \mathbf{H}_i \cdot \left(\mathbf{V}_{n-i} + \frac{\mathbf{1}}{\mathbf{2}} \right) - \frac{\mathbf{1}}{\mathbf{2}},$$

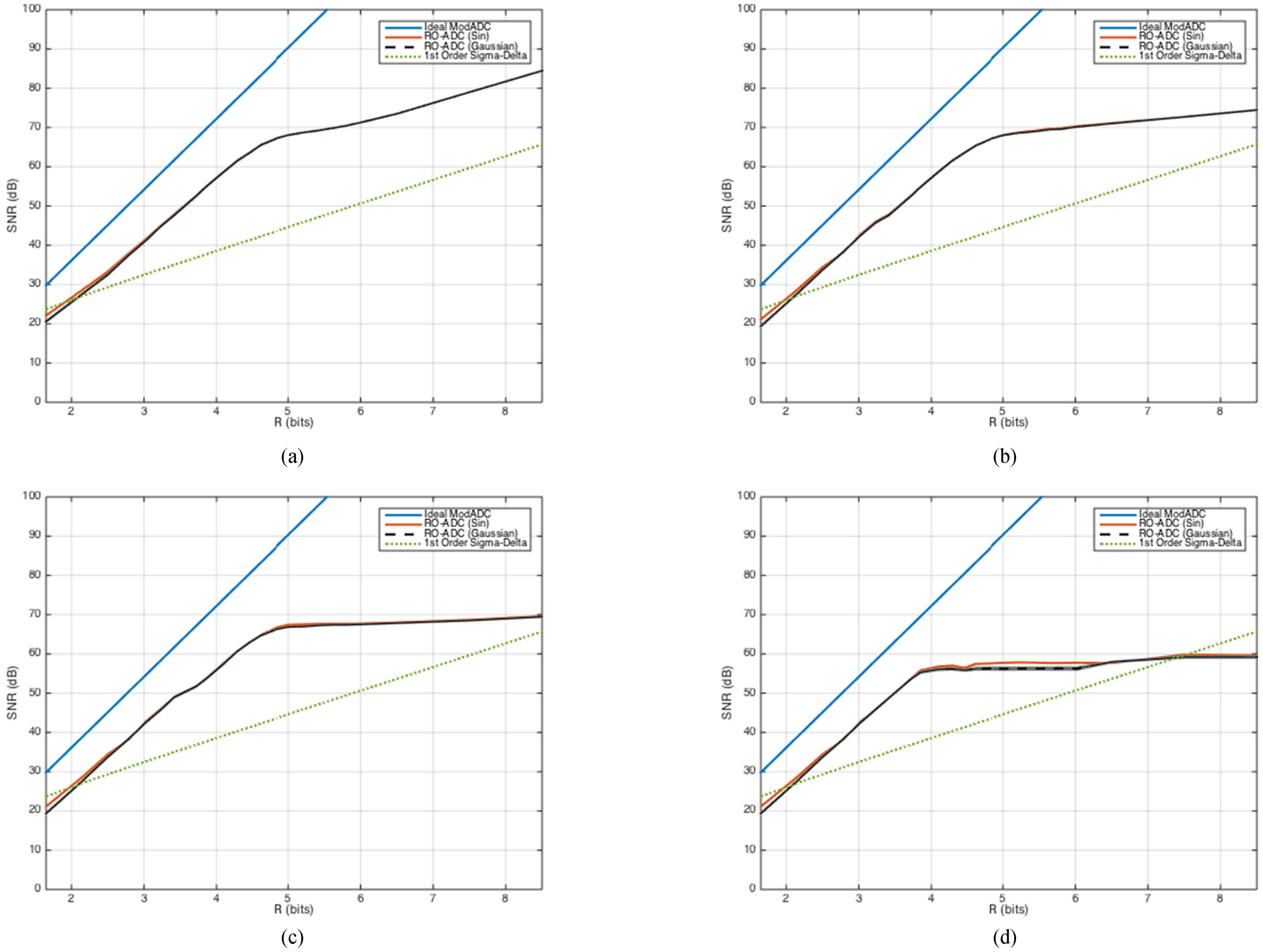
where $\{\mathbf{H}_n\}$ is a p-tap matrix prediction filter, $\mathbf{H}_i \in \mathbb{R}^{K \times K}$, for $i = 1, \ldots, p$, computed based on $\{C_{\ell m}[r]\}$ for all $\ell, m \in \{1, \ldots, K\}$ and α, and the shift by $\frac{\mathbf{1}}{\mathbf{2}}$ compensates for $\mathbb{E}(\mathbf{Z}_n)$.

2) Compute

$$\mathbf{W}_n = [\mathbf{Y}_n - \hat{\mathbf{V}}_n^p] \bmod 2^R,$$

where the modulo reduction is to be understood as taken component-wise.

3) Define the pth order prediction error $\mathbf{E}_n^p \triangleq \mathbf{V}_n - \hat{\mathbf{V}}_n^p$, and compute its covariance matrix $\boldsymbol{\Sigma}_p = \mathbb{E}\left[\mathbf{E}_n^p (\mathbf{E}_n^p)^T\right]$ based on $\{C_{\ell m}[r]\}$ for all $\ell, m \in \{1, \ldots, K\}$ and α. Note that $\boldsymbol{\Sigma}_p$ is indeed invariant with respect to n due to stationarity.



(a)

(b)

(c)

(d)

Fig. 9. Performance of ring oscillators based modulo ADC (RO-ADC). We plot SNR vs. quantization rate for a Gaussian process and for a sinusoidal waveform processes with a random frequency, uniformly distributed over $[0, B)$. For comparison we also plot the performance of an ideal (R, α) mod-ADC, as well as those of an ideal first-order $\Sigma\Delta$ converter. For all curves, SNR is defined as σ^2/D. The prediction filter has $p = 25$ taps, whereas the smoothing filter has $2k + 1$ taps for $k = 22$. (a) $B = 100$ Hz, $L = 3$. (b) $B = 44.1$ KHz, $L = 3$. (c) $B = 100$ KHz, $L = 3$. (d) $B = 1$ MHz, $L = 3$.

4) Solve

$$\begin{aligned}\mathbf{A} &= [\mathbf{a}_1|\cdots|\mathbf{a}_K]^T \\ &= \underset{\substack{\bar{\mathbf{A}}\in\mathbb{Z}^{K\times K}\\ |\bar{\mathbf{A}}|\neq 0}}{\operatorname{argmin}} \max_{k=1,\ldots,K} \frac{1}{2}\log\left(12\bar{\mathbf{a}}_k^T\boldsymbol{\Sigma}_p\bar{\mathbf{a}}_k\right). \qquad (39)\end{aligned}$$

5) For $k = 1, \ldots, K$, compute

$$\bar{g}_n^k \triangleq \left[\mathbf{a}_k^T\mathbf{W}_n\right] \bmod 2^R$$

$$\tilde{g}_n^k \triangleq \left[\bar{g}_n^k + \frac{1}{2}2^R\right] \bmod 2^R - \frac{1}{2}2^R,$$

and set $\tilde{\mathbf{g}}_n = [\tilde{g}_n^1, \ldots, \tilde{g}_n^k]^T$.

6) Compute

$$\hat{\mathbf{E}}_n^p = \mathbf{A}^{-1}\tilde{\mathbf{g}}_n,\;\; \hat{\mathbf{V}}_n = \hat{\mathbf{V}}_n^p + \hat{\mathbf{E}}_n^p,\;\; \hat{\mathbf{X}}_n = \frac{\hat{\mathbf{V}}_n + \frac{1}{2}}{\alpha}.$$

Proposition 5: Let $\mathbf{A} = [\mathbf{a}_1|\cdots|\mathbf{a}_K]^T$ be the matrix found in step 4 of the algorithm above, and define

$$R_{\mathrm{IFSC}}^{\mathrm{ST}}(\mathbf{A}) = \max_{k=1,\ldots,K} \frac{1}{2}\log\left(12\mathbf{a}_k^T\boldsymbol{\Sigma}_p\mathbf{a}_K^T\right). \qquad (40)$$

We have that

$$\Pr(\mathcal{E}_{\mathrm{OL}_n}) = \Pr(\hat{\mathbf{V}}_n \neq \mathbf{V}_n) \leq 2K\exp\left\{-\frac{3}{2}\cdot 2^{2\left(R - R_{\mathrm{IFSC}}^{\mathrm{ST}}(\mathbf{A})\right)}\right\},$$

and

$$D_n^k = \mathbb{E}\left[\left(X_n^k - \hat{X}_n^k\right)^2 \middle| \mathcal{E}_{\overline{\mathrm{OL}_n}}\right] \leq \frac{1}{12\alpha^2(1 - \Pr(\mathcal{E}_{\mathrm{OL}_n}))},$$

for all $k = 1, \ldots, K$, where the event $\mathcal{E}_{\overline{\mathrm{OL}_n}} = \{\hat{\mathbf{V}}_n = \mathbf{V}_n\}$ is the complement of the event $\mathcal{E}_{\mathrm{OL}_n} = \{\hat{\mathbf{V}}_n \neq \mathbf{V}_n\}$.

Proof: We first note that

$$\begin{aligned}\mathbf{W}_n &= [\mathbf{Y}_n - \hat{\mathbf{V}}_n^p] \bmod 2^R \\ &= \left[[\mathbf{V}_n] \bmod 2^R - \hat{\mathbf{V}}_n^p\right] \bmod 2^R \\ &= \left[\mathbf{V}_n - \hat{\mathbf{V}}_n^p\right] \bmod 2^R \\ &= [\mathbf{E}_n^p] \bmod 2^R,\end{aligned}$$

where the second equality follows from the modulo distributive law (2). By (21), we have that

$$\bar{g}_n^k \triangleq \left[\mathbf{a}_k^T \mathbf{W}_n\right] \bmod 2^R = \left[\mathbf{a}_k^T \mathbf{E}_n^p\right] \bmod 2^R = [g_n^k] \bmod 2^R,$$

where

$$g_n^k = \mathbf{a}_k^T \mathbf{E}_n^p. \qquad (41)$$

Furthermore, $\tilde{g}_n^k \in [-\frac{1}{2}2^R, \frac{1}{2}2^R)$ is merely a cyclically shifted version of $\bar{g}_n^k \in [0, 2^R)$. Thus, $\tilde{g}_n^k = g_n^k$ if and only if $g_n^k \in [-\frac{1}{2}2^R, \frac{1}{2}2^R)$. Consequently, $\hat{\mathbf{E}}_n^p \neq \mathbf{E}_n$, and therefore $\hat{\mathbf{V}}_n \neq \mathbf{V}_n$, if and only if the event

$$\mathcal{E}_{\mathrm{OL}_n} = \bigcup_{k=1}^{K} \left\{ |g_n^k| \geq \frac{1}{2}2^R \right\},$$

occurs. Now, repeating the same steps from the proof of Proposition 3, we arrive at the claimed bounds. ∎

Using Shannon's lower bound, and applying similar arguments as in [41], one can show that any quantization scheme for the source $\{\mathbf{X}_n\}$ that produces R bits/sample/coordinate and attains $\mathbb{E}(X_n^k - \hat{X}_n^k)^2 \leq D$, $k = 1, \ldots, K$, $n = 1, \ldots$, must have $R \geq \frac{1}{K}h(\mathbf{X}_n|\mathbf{X}_{n-1}, \ldots) - \frac{1}{2}\log(2\pi e D)$. Let $\mathbf{E}_n^{p*} = \mathbf{X}_n - \hat{\mathbf{X}}_n^p$, where $\mathbf{X}_n^p$ is the optimal pth order MMSE (linear) predictor of $\mathbf{X}_n$ from $\{\mathbf{X}_{n-1}, \ldots, \mathbf{X}_{n-p}\}$, and let $\mathbf{\Sigma}_p^* = \mathbb{E}\left[\mathbf{E}_n^{p*}(\mathbf{E}_n^{p*})^T\right]$. We have that

$$\begin{aligned} h(\mathbf{X}_n|\mathbf{X}_{n-1}, \ldots, \mathbf{X}_{n-p}) &= h(\mathbf{E}_n^{p*}|\mathbf{X}_{n-1}, \ldots, \mathbf{X}_{n-p}) \\ &\overset{(a)}{=} h(\mathbf{E}_n^{p*}) \overset{(b)}{=} \frac{1}{2}\log\left((2\pi e)^K |\mathbf{\Sigma}_p^*|\right), \end{aligned}$$

where (a) follows from the orthogonality principle of MMSE estimation [37], and (b) from the fact that $\mathbf{E}_n^{p*}$ is a Gaussian random vector [5]. Thus, for any quantization scheme we must have

$$R(D) \geq R_{\mathrm{SLB}}(D) \triangleq \frac{1}{2}\log\left(\frac{\lim_{p\to\infty}|\mathbf{\Sigma}_p^*|^{\frac{1}{K}}}{D}\right).$$

Similarly to previous subsections, we set $D = 1/12\alpha^2$, which is a good approximation for D_k^n, $k = 1, \ldots, K$, provided that $\delta = R - R_{\mathrm{IFSC}}^{\mathrm{ST}}(A)$ is not too small. The rate required by our scheme, as given in Proposition 5, depends on $12\mathbf{\Sigma}_p$, which corresponds to the prediction error covariance of the process $\tilde{\mathbf{X}}_n = \sqrt{12\alpha^2}\mathbf{X}_n + \tilde{\mathbf{Z}}_n = \frac{1}{\sqrt{D}}(\mathbf{X}_n + \sqrt{D}\tilde{\mathbf{Z}}_n)$, where $\tilde{\mathbf{Z}}_n = \sqrt{12}\mathbf{Z}_n$ is a random vector with unit variance iid entries. Let $\tilde{\mathbf{\Sigma}}_p$ be the pth order prediction error covariance of the process $\mathbf{X}_n + \sqrt{D}\tilde{\mathbf{Z}}_n$. We can rewrite the rate required by our scheme as

$$R_{\mathrm{IFSC}}^{\mathrm{ST}}(\mathbf{A}, D) \triangleq \frac{1}{2}\log\left(\frac{\max_{k=1,\ldots,K} \mathbf{a}_k^T \tilde{\mathbf{\Sigma}}_p \mathbf{a}_k}{D}\right).$$

Now, noting that if $h(\mathbf{X}_n|\mathbf{X}_{n-1}, \ldots) > -\infty$, we have that $\tilde{\mathbf{\Sigma}}_p \to \mathbf{\Sigma}_p^*$ as $D \to 0$, we obtain the following proposition.

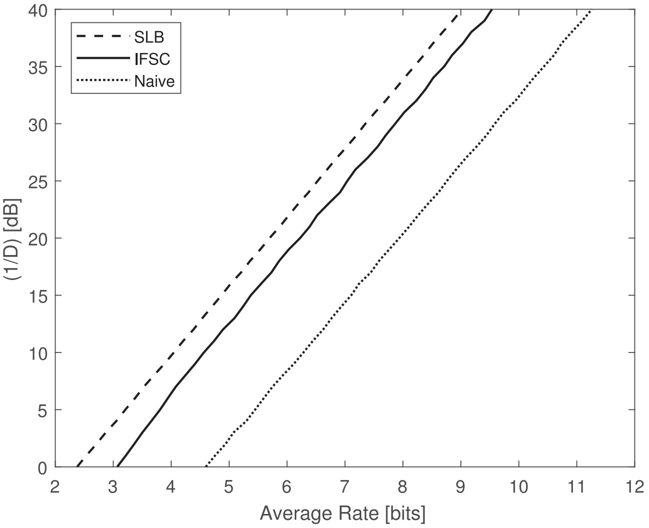


Fig. 10. Comparison between the average quantization rates $R_{\mathrm{IFSC}}^{\mathrm{ST}}(D)$, $R_{\mathrm{SLB}}(D)$, and $R_{\mathrm{naive}}(D)$. The setup is that of quantizing vector of stationary processes $\{X_n^1\}, \{X_n^2\}$ described in the end of Section VI, with $L = 5$ and $p = 24$.

Proposition 6: Assume $h(\mathbf{X}_n|\mathbf{X}_{n-1}, \ldots) \geq -\infty$, and let $\mathbf{\Sigma}^* \triangleq \lim_{p\to\infty} \mathbf{\Sigma}_p^*$. We have that

$$\begin{aligned} &\lim_{D\to 0}\lim_{p\to\infty} R_{\mathrm{IFSC}}^{\mathrm{ST}}(\mathbf{A}, D) - R_{\mathrm{SLB}}(D) \\ &= \frac{1}{2}\log\left(\frac{\max_{k=1,\ldots,K} \mathbf{a}_k^T \mathbf{\Sigma}^* \mathbf{a}_k}{|\mathbf{\Sigma}^*|^{\frac{1}{K}}}\right). \end{aligned} \qquad (42)$$

Thus, in the high-resolution regime, when taking large enough p, the gap between $R_{\mathrm{IFSC}}^{\mathrm{ST}}(\mathbf{A}, D)$ and the information theoretic lower bound is dictated by the loss of integer-forcing source decoder for a source whose covariance vector is $\mathbf{\Sigma}^*$. The right hand side of (42) is non-negative [11], but is typically quite small. To illustrate this, we generate two correlated processes $\{X_n^1\}$ and $\{X_n^2\}$ as follows: let $\{W_n^1\}, \{W_n^2\}, \{W_n^3\}$ be three iid $\mathcal{N}(0, 1)$ random processes. Let $X_n^1 = \sum_{i=0}^{L-1} h_i W_{n-i}^3 + W_n^1$, and $X_n^2 = \sum_{i=0}^{L-1} g_i W_{n-i}^3 + W_n^2$, where $\{h_n\}$ and $\{g_n\}$ are two filters, each with L taps. Clearly, when the filters have sufficiently strong taps the process $\{\mathbf{X}_n\} = [\{X_n^1\}, \{X_n^2\}]^T$ will be highly correlated in time and in space. In Figure 10 we plot the average rate required by the developed scheme, as well as $R_{\mathrm{SLB}}(D)$, and the rate required by a standard ADC that ignores spatial and temporal correlations entirely, denoted $R_{\mathrm{naive}}(D)$, with respect to to an iid $\mathcal{N}(0, 100)$ distribution on the $2L$ taps of $\{h_n\}$ and $\{g_n\}$. In the simulations performed, we took $L = 5$ and $p = 24$.

VII. Conclusions and Outlook

We have studied the modulo ADC architecture as an alternative approach for analog-to-digital conversion. The modulo ADC allows exploitation of the statistical structure of the input process digitally at the decoder without requiring the ADC to adapt itself to the input statistics. We have demonstrated the effectiveness of oversampled modulo ADCs as a simple substitute to $\Sigma\Delta$ converters, allowing an increase in the filter's order far beyond that which is possible in current $\Sigma\Delta$ converters, since for modulo ADC filtering is done digitally. Moreover, we have shown that, when used for digitizing jointly stationary processes, parallel modulo ADCs can efficiently exploit both temporal and spatial correlations.

An implementation of modulo ADCs via ring oscillators was developed, and the corresponding input-output function for the obtained modulo ADC was characterized in terms of the delay–V_{dd} profile of the inverters that construct the ring oscillator. We have then numerically studied the performance this implementation can attain for oversampled input processes, and compared it to those of $\Sigma\Delta$ converters.

There are several important challenges for future research. Perhaps most important is building a modulo ADC chip prototype. Although our simulations are based on the function $f(\cdot)$ measured from an actual (PSpice model of a) ring oscillator device, a hardware implementation is needed to fully assess the benefits of modulo ADCs. Furthermore, we would like to see whether it is possible to construct inverters with more favorable properties for ring oscillator-based modulo ADCs. In particular, we would like them to have a larger range where they are well approximated by an affine function. Another interesting avenue for future research is finding functions $g(\cdot)$ that can be implemented in the analog domain, such that the composition of function $f \circ g = f(g(\cdot))$ is more linear.

Acknowledgment

The authors are deeply grateful to Uri Erez, whose humbleness is the only reason for his absence from the authors list.

References

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

[2] B. Le, T. W. Rondeau, J. H. Reed, and C. W. Bostian, “Analog-to-digital converters,” *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 69–77, Nov. 2005.

[3] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1971.

[4] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1984.

[5] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley-Interscience, 2006.

[6] M. Hovin, A. Olsen, T. S. Lande, and C. Toumazou, “Delta-sigma modulators using frequency-modulated intermediate values,” *IEEE J. Solid-State Circuits*, vol. 32, no. 1, pp. 13–22, Jan. 1997.

[7] M. Z. Straayer and M. H. Perrott, “A 12-bit, 10-MHz bandwidth, continuous-time $\Sigma\Delta$ ADC with a 5-bit, 950-MS/s VCO-based quantizer,” *IEEE J. Solid-State Circuits*, vol. 43, no. 4, pp. 805–814, Apr. 2008.

[8] E. Telatar, “Capacity of multi-antenna Gaussian channels,” *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[10] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[11] O. Ordentlich and U. Erez, “Integer-forcing source coding,” *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1253–1269, Feb. 2017.

[12] T. Ericson and V. Ramamoorthy, “Modulo-PCM: A new source coding scheme,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 1979, vol. 4, pp. 419–422.

[13] G. Forney, “Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference,” *IEEE Trans. Inf. Theory*, vol. IT-18, no. 3, pp. 363–378, May 1972.

[14] V. Ramamoorthy, “A novel speech coder for medium and high bit rate applications using modulo-PCM principles,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 356–368, Apr. 1985.

[15] P. Noll, “On predictive quantizing schemes,” *Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1499–1532, May 1978.

[16] R. Zamir, Y. Kochman, and U. Erez, “Achieving the Gaussian rate-distortion function by prediction,” *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3354–3364, Jul. 2008.

[17] P. T. Boufounos, “Universal rate-efficient scalar quantization,” *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1861–1872, Mar. 2012.

[18] D. Valsesia and P. T. Boufounos, “Universal encoding of multispectral images,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2016, pp. 4453–4457.

[19] A. Bhandari, F. Krahmer, and R. Raskar, “On unlimited sampling,” in *Proc. Int. Conf. Sampling Theory Appl.*, Jul. 2017, pp. 31–35.

[20] A. Bhandari, F. Krahmer, and R. Raskar, “Unlimited sampling of sparse signals,” in *Proc. IEEE Intl. Conf. Acoust., Speech Signal Process.*, 2018, Paper 3510.

[21] E. Domanovitz and U. Erez, “Outage probability bounds for integer-forcing source coding,” in *Proc. IEEE Inf. Theory Workshop*, Kaohsiung, Taiwan, Nov. 2017, pp. 574–578.

[22] R. Zamir, S. Shamai, and U. Erez, “Nested linear/lattice codes for structured multiterminal binning,” *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.

[23] M. Tomlinson, “New automatic equalizer employing modulo arithmetic,” *Electron. Lett.*, vol. 7, pp. 138–139, Mar. 1971.

[24] H. Harashima and H. Miyakawa, “Matched-transmission technique for channels with intersymbol interference,” *IEEE Trans. Commun.*, vol. COM-20, no. 4, pp. 774–780, Aug. 1972.

[25] S.-N. Hong and G. Caire, “Compute-and-forward strategies for cooperative distributed antenna systems,” *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5227–5243, Sep. 2013.

[26] B. Nazer and M. Gastpar, “Compute-and-forward: Harnessing interference through structured codes,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[27] J. van Valburg and R. J. van de Plassche, “An 8-b 650-MHz folding ADC,” *IEEE J. Solid-State Circuits*, vol. 27, no. 12, pp. 1662–1666, Dec. 1992.

[28] R. Venkataramani and Y. Bresler, “Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals,” *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2173–2183, 2000.

[29] M. Vetterli, P. Marziliano, and T. Blu, “Sampling signals with finite rate of innovation,” *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[30] 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, Apr. 2010.

[31] U. Erez and R. Zamir, “Achieving $\frac{1}{2}\log(1+\mathrm{SNR})$ on the AWGN channel with lattice encoding and decoding,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[32] R. M. Gray, “Quantization noise spectra,” *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1220–1244, Nov. 1990.

[33] O. Ordentlich and U. Erez, “Precoded integer-forcing universally achieves the MIMO capacity to within a constant gap,” *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 323–340, Jan. 2015.

[34] C. Feng, D. Silva, and F. Kschischang, “An algebraic approach to physical-layer network coding,” *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7576–7596, Nov. 2013.

[35] O. Ordentlich and U. Erez, “Performance analysis and optimal filter design for sigma-delta modulation via duality with DPCM,” *Proc. IEEE Int. Symp. Inf. Theory*, 2015, pp. 321–325.

[36] R. M. Gray, “Toeplitz and circulant matrices: A review,” *Foundations Trends® Commun. Inf. Theory*, vol. 2, no. 3, pp. 155–239, 2006.

[37] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression* (Springer International Series in Engineering and Computer Science vol. 159). Berlin, Germany: Springer Science & Business Media, 2012.

[38] A. K. Lenstra, H. W. Lenstra, and L. Lovász, “Factoring polynomials with rational coefficients,” *Mathematische Annalen*, vol. 261, no. 4, pp. 515–534, 1982.

[39] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits: A Design Perspective*. London, U.K.: Pearson Education, 2003.

[40] J. J. Bussgang, “Crosscorrelation functions of amplitude-distorted Gaussian signals,” Res. Lab. Electron., MIT, Cambridge, MA, USA, Tech. Rep. 216, 1952.

[41] R. Zamir and T. Berger, “Multiterminal source coding with high resolution,” *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 106–117, Jan. 1999.

Authors’ photographs and biographies not available at the time of publication.