

Power-Efficient Sampling

Satish Mulleti, *Member, IEEE*, Timur Zirtiloglu, *Student Member, IEEE*, Arman Tan, *Student Member, IEEE*, Rabia Tugce Yazicigil, *Senior Member, IEEE*, Yonina C. Eldar, *Fellow, IEEE*

Abstract—Analog-to-digital converters (ADCs) facilitate the conversion of analog signals into a digital format. While the specific designs and settings of ADCs can vary depending on their applications, it is crucial in many modern applications to minimize their power consumption. The significance of low-power ADCs is particularly evident in fields like mobile and handheld devices reliant on battery operation. Key parameters of the ADCs that dictate the ADC's power are its sampling rate, dynamic range, and number of quantization bits. Typically, these parameters are required to be higher than a threshold value but can be reduced by using the structure of the signal and by leveraging preprocessing and the system application needs. In this review, we discuss four approaches relevant to a variety of applications.

I. INTRODUCTION

Real-world signals, including speech, audio, biomedical data, radar readings, communication transmissions, and more, inherently exist in an analog format. Transforming these signals into a digital representation is often a more efficient way to process, compress, store, and transmit them when considering factors like power efficiency, noise resistance, and algorithms' adaptability. Therefore, the digital conversion of analog signals is central to any application involving real-world signals.

To achieve this digital representation, analog-to-digital converters (ADCs) are employed. Given that ADCs are an essential component in many applications, their design and selection of parameters, such as the sampling rate, the number of bits for quantization, and the dynamic range (DR), are critical factors in effectively addressing the requirements of a particular application. For instance, when it comes to accurately reconstructing bandlimited signals from uniformly spaced samples, the ADC's sampling rate should meet or exceed the Nyquist rate, which is twice the maximum frequency component of the analog signal. In a similar vein, the dynamic range of the ADC should exceed that of the signal to prevent signal clipping during the sampling process, and a greater number of bits results in reduced quantization errors.

In emerging fields like wearable biosensors, wireless sensor networks, voice-activated interfaces (such as Amazon's Alexa, Apple's Siri, Microsoft's Cortana, and Google's Assistant), Internet of Things sensors, and more, the power consumption of ADCs holds great importance. Take, for example, wearable biosensors, which can monitor parameters like blood pressure, electrocardiogram readings, and epilepsy seizures. These

devices often operate within stringent resource limitations, including power and space constraints, especially since many of them rely on battery power with finite energy reserves. In such applications, efficiently reducing power requirements for the sensor's ADCs is essential to ensure uninterrupted monitoring of the signals. Likewise, in any voice-activated interface, the voice detection sensors remain active continuously, underscoring the significance of minimizing the power demands of their ADCs. Furthermore, in applications that deal with wideband signals, the power consumption of the ADC is notably high due to the need for a high sampling rate, and this element constitutes a significant portion of the overall power requirements of the system.

In this review, we explore techniques and frameworks to lower the ADC power and the entire system power. A common theme is to enable analog preprocessing prior to sampling together with postprocessing on the signal in order to reduce the ADC power while still enabling signal recovery. The recovery may involve either reconstructing the original analog signals from the collected samples or estimating certain parameters. The power consumption of a conventional ADC, which captures uniformly spaced instantaneous samples of the analog signal, increases in tandem with the sampling rate, the number of bits used, and the dynamic range of the ADC. To address this challenge, we delve into theoretical lower limits on the power consumption of ADCs [1] and also examine a few figures-of-merit (FoMs) employed for assessing commercially available ADCs [1], [2]. Once these relationships are established, various frameworks can be applied to mitigate the impact of these factors on power requirements.

We consider the following four approaches for power reduction: (a) *Sub-Nyquist sampling*, where the signal's structure is used to reduce the sampling rate [3]; (b) *Modulo-ADC* where a low-dynamic range ADC preceded with a folding circuit is used to sample high-DR signals without clipping [4]–[7]; (c) *Asynchronous sampling* where time instants at which a signal crosses a set of thresholds are used as the discrete representation [8]–[12]; and (d) *Low-bit quantization* methods in which the number of bits is reduced when the end task is to recover a few parameters of the signal [13], [14].

The majority of these frameworks comprise three primary components: (i) an analog pre-processing hardware or an analog front-end, (ii) a traditional low-power ADC, and (iii) a digital processing module. The pre-processing component alters the analog signal in a manner that allows the ADC to function at a reduced sampling rate, a lower bit rate, or a narrower dynamic range compared to an ADC lacking pre-processing. The digital processing segment encompasses algorithms that can potentially reverse the impact of the pre-processing stage and includes reconstruction or recovery algorithms. We detail

S. Mulleti is with the Department of Electrical Engineering, Indian Institute of Technology (IIT) Bombay, India. T. Zirtiloglu, A. Tan and R. T. Yazicigil are with the Electrical and Computer Engineering Department, Boston University, Boston, MA, USA, and Y. C. Eldar is with the Faculty of Math and Computer Science, Weizmann Institute of Science, Israel. Email: mulleti.satish@gmail.com, timurz@bu.edu, armantan@bu.edu, rty@bu.edu, yonina.eldar@weizmann.ac.il

Bounds on Power Consumption of a Conventional ADC

To compute lower bounds on the power, we consider two major components of a conventional ADC: (a) a sample and hold (S/H) circuit that samples and retains the value of the analog signal for quantization; (b) a quantizer that converts the held value from S/H to a bit stream (See Fig. 1) using comparators. The total power consumption is the sum of that of the S/H circuit and the comparators. To derive the total power, we assume that the capacitors used in the S/H and quantizers are chosen to be large enough such that the thermal noise is less than or equal to the quantization error [1]. For deriving the quantization error, we assume that the input signal $f(t)$ lies in the amplitude range $[-A, A]$ and the quantizer has n bits. With these assumptions, the power consumption in the S/H is given as

$$P_S = 24kTf_s2^{2n}, \quad (1)$$

where k is Boltzmann's constant and T is the absolute temperature in degrees Kelvin. In the quantizer part, an expression for power consumption in each comparator is derived as

$$P_C = 4n \ln 2 V_{\text{eff}} C_C f_s A, \quad (2)$$

where C_C is the load capacitance of the latch comparator. The quantity V_{eff} is a parameter of the comparator and is defined as $V_{\text{eff}} = g_m I_D$ where I_D is the supply current of the comparator, and g_m is the minimum transconductance of the comparator to make a decision within the sampling period. Hence, the total power of a conventional flash ADC is

$$P = P_S + (2^n - 1)P_C = 24kTf_s2^{2n} + 4(2^n - 1)n \ln 2 V_{\text{eff}} C_C f_s A. \quad (3)$$

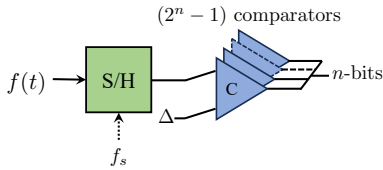


Fig. 1. A conventional flash-type ADC with a sample and hold circuit operating at a sampling rate f_s , and a quantizer consisting of 2^n comparators where n is the number of bits per sample. We assume that the analog signal $f(t)$ is bounded: $|f(t)| \leq A$. The quantization step size is $\Delta = A/(2^{n-1})$.

these components' integrated design, considering theoretical and practical perspectives. In the theoretical realm, we present assurances for achieving perfect signal recovery and examine algorithms concerning critical parameters like the sampling rate, number of bits, and dynamic range. Subsequently, we explore the influence of these parameters on power consumption and hardware design. In addition, we also discuss existing hardware prototypes for these various approaches [5], [15]–[22].

In the following section, we discuss bounds and FoMs, which will be used to assess the power-saving aspects of the considered frameworks.

II. POWER CONSUMPTION OF AN ADC AND FoMs

We first discuss bounds on the power consumption of a conventional ADC, which are used throughout the paper to compare different ADCs. The details are given in the box *Bounds on Power Consumption of a Conventional ADC*. We refer the readers to [1] and references therein for details on these derivations.

From (3), we infer that the power consumption increases linearly with the sampling rate and exponentially with the

number of bits. These observations are based on theoretical analysis, and in practice, the power dependency on the sampling rate and number of bits could be different. To understand the practical aspect, we consider the surveys considered in [2] and a recent one by Murmann [23]. These reports compare various ADCs in terms of different FoMs. As an example, Walden's FoM for a conventional ADC is given as [2]

$$\text{FOM}_W = \frac{P}{2^{\text{ENOB}} f_s}, \quad (4)$$

where P is the power consumption of the ADC, f_s is the sampling rate, and ENOB is the effective number of bits. An empirical study of the available ADCs in [23] shows that the FOM_W does not change as f_s changes for $f_s \leq 100$ MHz. This implies that P is proportional to f_s . However, for $f_s > 100$ MHz, the behavior of FOM_W as a function of f_s reveals that the increase in power consumption due to the sampling frequency exceeds linear increment, which means that the actual power saving of sub-Nyquist is more than the theoretical power saving. Further, the reviews reveal that, in commercially available ADCs, the sampling rate and number of bits are typically not independent. Specifically, the number of bits decreases with an increased sampling rate owing to design constraints.

In the aforementioned analysis, we do not observe an explicit dependency of power consumption on the DR. Rather, power consumption is an implicit function of DR, as shown next. Consider two ADCs with dynamic ranges $[-\lambda_1, \lambda_1]$ and $[-\lambda_2, \lambda_2]$ respectively, where $\lambda_1 > \lambda_2$. Let the resolutions of the two ADCs be N_1 and N_2 bits, respectively. Then, the quantization noise powers are given as $\frac{1}{12} \left(\frac{\lambda_1}{2^{N_1-1}} \right)^2$ and $\frac{1}{12} \left(\frac{\lambda_2}{2^{N_2-1}} \right)^2$. To keep the same quantization error for both the ADCs, the number of bits should satisfy the equality $N_1 =$

$N_2 + \frac{\lambda_1}{\lambda_2} \log 2$. Hence, for a high-DR ADC, one must choose a larger number of bits to keep the same quantization error level. Following (3), this results in higher power consumption.

In summary, the power consumption of an ADC increases with the sampling rate, number of bits, and DR. In the following, we discuss how to reduce these quantities and eventually the power consumption without compromising the quality of the reconstruction or task. In general, these quantities are reduced by analog preprocessing steps, which could consist of either single-channel or multiple-channel front-end circuits. These front ends could be realized using off-the-shelf components or custom-designed analog and radio-frequency circuits with low-power consumption by co-designing digital signal processing and analog preprocessing. Hence, the overall power consumption of the circuit can be made lower than the corresponding conventional circuit.

III. SUB-NYQUIST SAMPLING

In most practical applications, analog signals are either assumed bandlimited or maximally bandlimited. In the latter case, one assumes that most (e.g., 99%) of the signal's energy lies in a frequency range known as the signal's effective bandwidth. These signals are sampled according to the well-known Shannon-Nyquist sampling framework [3]. In this framework, the signals are first passed through an anti-aliasing filter to project the signal to a bandlimited space and then sampled uniformly at the Nyquist rate, which is twice the maximum frequency content of the signal. Since the power consumption of ADCs increases with the sampling rate, the Shannon-Nyquist framework leads to higher power consumption for wideband applications such as ultra-wideband radar, cognitive radio, and wideband communications, where the bandwidth of the signals is on the order of GHz.

A solution to this problem is to explore signal structures beyond bandlimitedness and reduce the sampling rate below the Nyquist rate (sub-Nyquist sampling) by using the signal structure or model. Many signals encountered in practice have sparse representations either in the time or frequency domains, which leads to sub-Nyquist sampling. Several review articles on sub-Nyquist sampling exist from both theoretical aspects [3], [17], [24], [25] and from practical considerations [22], [26]. Hence, in this review, we keep discussion on the sub-Nyquist theory short and focus on the power savings that can be obtained through this approach. In particular, we consider two signal models, finite-rate-of-innovation (FRI) signals [27]–[31] and multiband signals [16], [32]–[35], and discuss their power efficiency aspects.

A. Finite-Rate-of-Innovation Signals

In many applications, the signals can be specified by a finite number of parameters per unit time interval. As the rate of innovation (RoI) or the number of parameters specifying the signal per unit time interval is finite, these classes of signals are known as FRI signals. These signals can be sampled at their RoI; if the RoI is lower than the Nyquist rate, then it is sub-Nyquist sampling.

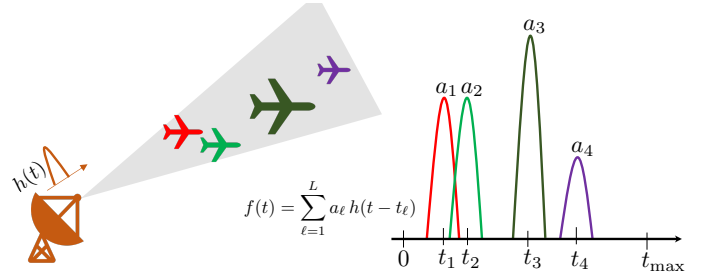


Fig. 2. Application of FRI signals in radar imaging: A known transmit pulse $h(t)$ is reflected from sparsely located targets; the received signal is modeled as in (5) where time delays and amplitudes parameterize the signal.

Commonly used FRI signals are streams of known pulses representing signals in time-of-flight applications such as radar, ultrasound, optical coherence tomography, and more. An example is depicted in Fig. 2 where a known transmit pulse $h(t)$ reflects from sparsely located L point targets. The received signal is given by

$$f(t) = \sum_{\ell=1}^L a_{\ell} h(t - t_{\ell}), \quad (5)$$

where the amplitude a_{ℓ} and delay $t_{\ell} \in (0, t_{\max}]$ denote size and location of the ℓ -th target, and t_{\max} is maximum time delay. The signal $f(t)$ is specified by $2L$ parameters: $\{a_{\ell}, t_{\ell}\}_{\ell=1}^L$ and hence is an FRI signal. The Nyquist rate of these signals depends on the essential bandwidth of the pulse $h(t)$, which is generally very high due to its short time duration. However, due to the FRI nature, the parameters can be determined from sub-Nyquist samples.

In Fig. 3, we depict a frequently employed FRI sampling and reconstruction framework consisting of a sampling and a reconstruction block. The sampling kernel $g(t)$ acts like an anti-aliasing filter in bandlimited sampling and plays a key role in the sampling block. The kernel removes additional information from the FRI signal and ensures that the filtered output $y(t) = (f * g)(t)$ is in a form such that the FRI parameters can be uniquely determined from low-rate samples of $y(nT_s)$. The reconstruction block, typically, consists of a digital signal processing (DSP) unit that linearly combines samples of $y(nT_s)$ and a parameter estimation method that determines the FRI parameters $\{a_{\ell}, t_{\ell}\}_{\ell=1}^L$.

The sampling and reconstruction blocks are interdependent and must be designed in unison. To elaborate on this, we consider a compactly supported sum-of-sincs (SoS) kernel with impulse response given as [30]

$$g(t) = \text{rect}\left(\frac{t}{T_g}\right) \sum_{k=-K}^K e^{jk\omega_0 t}, \quad (6)$$

where $\omega_0 > 0$ and $\left(\frac{t}{T_g}\right) = 1$ for $t \in [0, T_g]$ and zero elsewhere. If the support of the filter T_g is sufficiently larger than that of the signal $f(t)$, then a part of the filtered signal

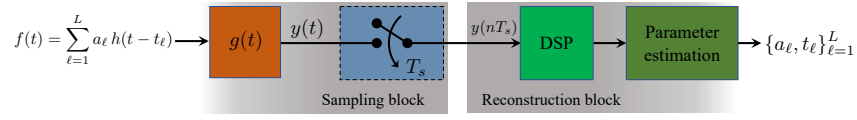


Fig. 3. A schematic of FRI sampling and reconstruction framework

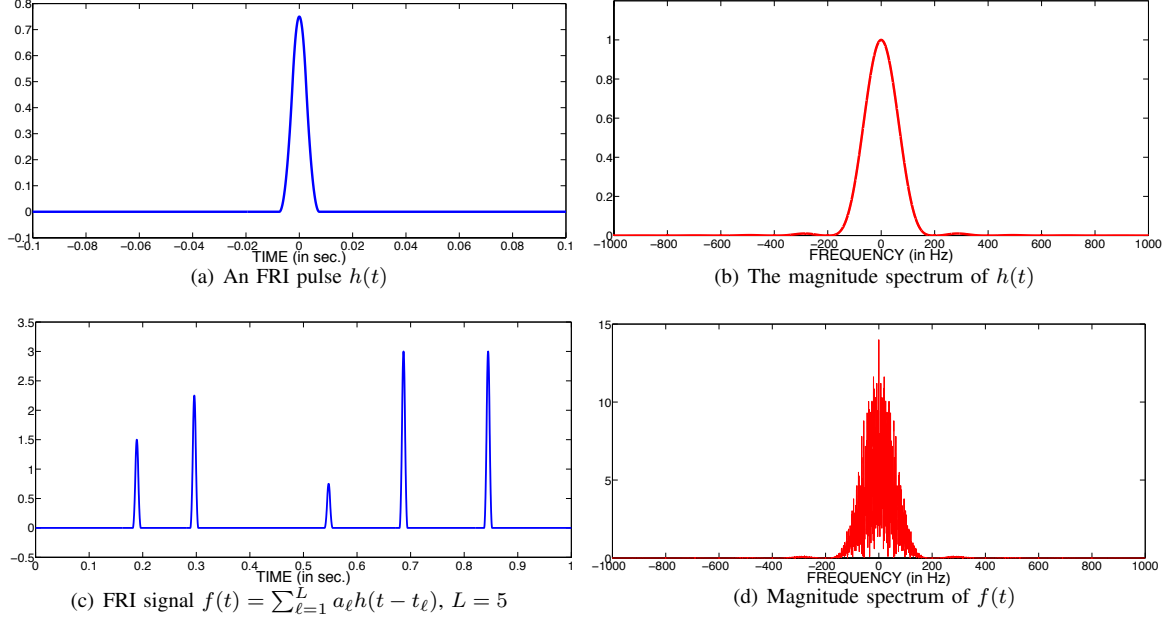


Fig. 4. An illustration of sub-Nyquist sampling of FRI signals: (a) $h(t)$ is a time-limited pulse with its magnitude spectrum $|H(\omega)|$ shown in (b). In (c), an FRI signal is shown consisting of $L = 5$ delayed copies of $h(t)$ with $t_{\max} = 1$ sec. The magnitude spectrum of $f(t)$ shown in (d) shows that the Nyquist sampling rate is 400 Hz; however, the FRI-based framework requires sampling at 11 Hz.

can be written as

$$y(t) = \sum_{k=-K}^K F(k\omega_0) e^{-jk\omega_0 t}, \quad (7)$$

where $F(\omega)$ is the continuous-time Fourier transform (CTFT) of $f(t)$ (cf. [30], [31] for details). By choosing $K \geq L$ and measuring $2K + 1$ or more uniform samples $y(nT_s)$, one can determine the Fourier samples $F(k\omega_0)$ uniquely provided that $\omega_s = \frac{2\pi}{T_s} \geq (2K + 1)\omega_0$.

From the Fourier samples, one can construct the following sequence

$$S(k\omega_0) = \frac{F(k\omega_0)}{H(k\omega_0)} = \sum_{\ell=1}^L a_\ell e^{jk\omega_0 t_\ell}, \quad k = -K, \dots, K \quad (8)$$

where $F(\omega)$ and $H(\omega)$ are CTFTs of $f(t)$ and $h(t)$, respectively, and ω_0 is sampling interval in the Fourier domain. The sequence $S(k\omega_0)$ consists of a linear combination of complex exponentials with discrete frequencies $\{\omega_0 t_\ell\}_{\ell=1}^L$. The parameters $\{a_\ell, t_\ell\}_{\ell=1}^L$ can be uniquely determined from the sequence $S(k\omega_0)$ by using high-resolution spectral estimation methods [36, Ch. 4] provided that $\omega_0 t_{\max} < 2\pi$ and $K \geq L$. In this approach, we assume that $H(k\omega_0) \neq 0, k = -K, \dots, K$, which can always be ensured in practice as $h(t)$ has a narrow duration in time and therefore large bandwidth.

Similar spectral-estimation-based reconstruction can also

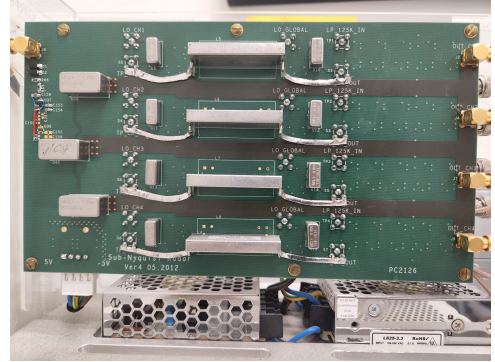


Fig. 5. A hardware prototype of sub-Nyquist radar [15].

be achieved by using a lowpass filter or a Gaussian kernel [27]. However, these kernels have infinite support and are, hence, impractical to realize. Other choices of compactly supported sampling kernels include polynomial and exponential-generating kernels [28], [37]. All the kernels have a common working principle: Spread the information of the FRI pulses such that each sample of the filtered signal contains sufficient information about the amplitude and time delays of all the pulses.

From the inequalities $\omega_s = \frac{2\pi}{T_s} \geq (2K + 1)\omega_0$ and $|\omega_0 t_{\max}| < 2\pi$, we infer that the lowest possible sampling rate is $(2L + 1) \frac{2\pi}{t_{\max}}$. The sampling rate is independent of the

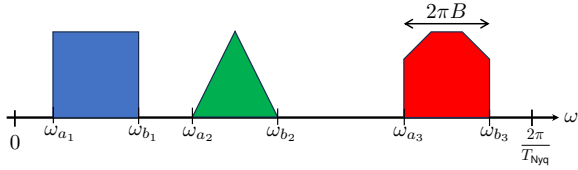


Fig. 6. Spectrum of a multiband signal consisting of three disjoint bands. The maximum frequency of the signal is $2\pi/T_{\text{Nyq}}$, and the Nyquist rate is $1/T_{\text{Nyq}}$ Hz. In this example, for simplicity, we consider a signal with a positive spectrum, whereas signals in practice can have both positive and negative parts.

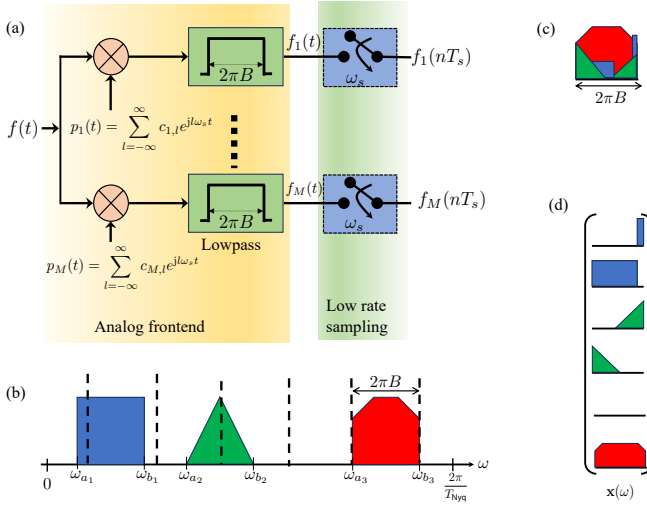


Fig. 7. An M -channel MWC-based mixing and sub-Nyquist framework for multiband signals [16], [34], [35].

bandwidth of the FRI signal and only a function of the number of pulses and maximum time delay. Indeed, this results in sub-Nyquist sampling, as illustrated in the subsequent example. In Fig. 4, we show a particular example of an FRI signal where $h(t)$ is a time-limited pulse constructed from a time-scaled, third-order B-spline function. The FRI signal consists of $L = 5$ pulses with a maximum time delay of one second. The magnitude spectra of the pulse $h(t)$ and $f(t)$ show that the essential bandwidth is 200 Hz, which shows that the Nyquist rate is 400 Hz. However, by using the FRI nature of the signal, we know that it can be sampled at a minimum rate of $2L + 1$ or 11 Hz. Hence, there is a reduction in the sampling rate by 36 times in this example, and consequently, power saving will be of the same order. However, one may not achieve such a large sampling rate reduction in practice due to noise and other practical limitations such as non-ideal sampling kernels. For example, hardware prototypes for sub-Nyquist radar systems, where the sampling and reconstruction are based on the FRI principle, are discussed in [15]. A hardware board is shown in Fig. 5. The hardware prototypes demonstrate that using the received signals' FRI nature, the radar targets are estimated from the samples measured at $1/30$ of the Nyquist rate. As the sampling rate is reduced by a factor of 30 in this case, following (3), power consumption is reduced by at least a factor of 30. If the operating frequencies are in the range of a few hundred MHz, then the power consumption goes down further, as discussed earlier, with the help of FOM_W.

B. Sub-Nyquist Sampling of Multiband Signals

To simultaneously communicate several bandlimited signals over a common channel, signals are modulated to different carrier frequencies, as shown in Fig. 6. Following the Shannon-Nyquist sampling framework, such multiband signals can be sampled at twice the maximum frequency bandwidth of the signal. However, as shown in the example in Fig. 6, the signal's spectrum is zero over several intervals, and the information-bearing part resides in a few bands. Mathematically, consider a set of N disjoint frequency intervals $\mathcal{M} = \cup_{n=1}^N [\omega_{b_n} - \omega_{a_n}] \subset [0, 2\pi/T_{\text{Nyq}}]$ where ω_{a_n} and ω_{b_n} denotes lower and upper band edges of the n -th band. Then $\mathcal{B}_{\mathcal{M}}$ denotes a set of signals whose spectrum is restricted to the interval \mathcal{M} . Specifically, if $f(t) \in \mathcal{B}_{\mathcal{M}}$ then $F(\omega) = 0, \omega \notin \mathcal{M}$. The Nyquist rate of the signals in this class is $1/T_{\text{Nyq}}$, which could be much larger than the spectral support $|\mathcal{M}|$.

By observing the sparsity of the signal's spectrum, Landau [38] showed that a multiband signal is uniquely identifiable from its stable samples measured at a rate equal to its spectral support (referred to as Landau rate). Hence for signals in $\mathcal{B}_{\mathcal{M}}$, sampling at a rate $|\mathcal{M}|$ is sufficient. Periodic non-uniform sampling and multi-coset sampling patterns can be used for sampling and reconstructing multi-band signals [32], [33]. Most of these approaches show that perfect reconstruction can be achieved by sampling the signals at the Landau rate, provided that the band edges $\{\omega_{a_n}, \omega_{b_n}\}_{n=1}^N$ are known.

A blind sampling and reconstruction framework was presented in a series of papers [16], [34], [35], where the authors proved that the minimum sampling rate is twice Landau's rate and not more than the Nyquist rate in the blind setup. In these blind-multiband works, the authors assumed that $\omega_{b_n} - \omega_{a_n} = 2\pi B$ and B and N are known. Then, the set of multiband signals can be uniquely identifiable from stable samples measured at a rate $\min\{2NB, 1/T_{\text{Nyq}}\}$. An oversampling by a factor of two is the price paid for blindness. The signal reconstruction algorithms in these works first estimate the unknown spectral supports and then use a conventional algorithm to reconstruct the signals from the samples. For support recovery, the algorithms rely on a similar principle as in FRI sampling: to spread or mix the information so that the desired information is determined from fewer samples relying on notions from compressed sensing.

For mixing the information, the authors proposed two multichannel frameworks that are based on multi-coset sampling [33], quadrature analog-to-information converter (QAIC) [39], [40], and the modulated-wideband converter (MWC) [16], [17], [35]. The principle idea of mixing is based on aliasing of the spectrum. Reconstruction is achieved by applying sparse-recovery methods from compressive sensing [41] as discussed next. For an M -channel sampling framework, an MWC-based mixing and sampling framework is shown in Fig. 7. The mixing is performed by multiplying $f(t)$ with a set of M periodic signals $p_m(t) = \sum_{l=-\infty}^{\infty} c_{m,l} e^{j\omega_s l t}$, $m = 1, \dots, M$, where $\omega_s \geq 2\pi B$ and $c_{m,l}$ s are Fourier coefficients of the periodic signals. This results in modulation in the frequency domain and, as a result, the lowpass spectrum of the product $f(t)p_m(t)$ is a linear combination of all the bands as shown in

Fig. 7(c). The mixed spectrum is captured from the sampled signal by passing the product $f(t)p_m(t)$ through a lowpass filter and then sample uniformly at its Nyquist rate. The discrete-time Fourier transform of the samples is related to the spectrum of the continuous-time multiband signal as [34], [35]

$$F_m(e^{j\omega T_s}) = \sum_{\ell=-L}^L c_{m,\ell} F(\omega - \ell\omega_s), \quad \omega \in [-\omega_s/2, \omega_s/2], \quad (9)$$

where $\omega_s = \frac{2\pi}{T_s} < \omega_{\text{Nyq}}$ is the sampling rate in rad/sec. The summation in (9) is due to aliasing because of sub-Nyquist sampling, and its finiteness is a consequence of the bandlimitedness of $f(t)$. Re-writing (9) in matrix-vector notation, we have

$$\mathbf{y}(\omega) = \mathbf{A}\mathbf{x}(\omega) \in \mathbb{C}^M, \quad \omega \in [-\omega_s/2, \omega_s/2], \quad (10)$$

where $\mathbf{A}_{m,\ell} = c_{m,\ell}$, $\mathbf{y}(\omega) = [F_1(e^{j\omega T_s}), \dots, F_M(e^{j\omega T_s})]^T$, and $\mathbf{x}(\omega) = [F(\omega - L\omega_s), \dots, F(\omega + L\omega_s)]^T \in \mathbb{C}^{2L+1}$. The sparsity of the spectrum $F(\omega)$ ensures that the vector $\mathbf{x}(\omega)$ is sparse for each $\omega \in [-\omega_s/2, \omega_s/2]$ provided that $\omega_s \geq 2\pi B$ (See Fig. 7(d)). Algorithms and conditions for recovering $\mathbf{x}(\omega)$ from $\mathbf{y}(\omega)$ are derived in [34], [35] where the authors showed that the minimum number of the hardware channels should be greater than the number of bands, that is, $M \geq 2N$. Hence, for reconstruction, the minimum sampling rate is $2NB$, which is twice the rate for unique identifiability. In practice, the number of channels can be reduced below $2N$, as low as one channel, by designing a specific set of periodic signals $p_m(t)$ [34].

The periodic signals in an MWC are typically realized using a sign-alternating sequence or pseudorandom noise (PN) sequence generator. The PN sequence has to alternate signs a sufficiently large number of times within a time period to ensure that the sparse spectrum is estimated from the compressed measurements. Thus, these PN mixers operate at the Nyquist rate of the maximum input signal frequency. For example, for the multiband signal shown in Fig. 6, the mixers operate at $2\pi/T_{\text{Nyq}}$. By co-designing the analog preprocessing with digital signal processing, QAIC applies a band-pass filter and downconverts the high-frequency input signal. Specifically, by assuming that the spectra of the multiband signal lie in a frequency interval $[\omega_{\min}, \omega_{\max}]$, where $\omega_{\max} \leq 2\pi/T_{\text{Nyq}}$ and $\omega_{\min} > 0$, the downconverted signal spectrum lie in the range $[0, \omega_{\max} - \omega_{\min}]$. With this change in the spectrum, the Nyquist rate of the signal is reduced by ω_{\min} , and the same amount reduces the rate of the PN sequence. A comparison of the lowpass filter-based MWC and a bandpass filter-based QAIC is shown in Fig. 9. If $\omega_{\min} \gg 0$ for the application of interest, the power savings introduced by the band-pass QAIC architecture due to the reduction of the clock rate and sequence length of the PN sequence generator compared to the low-pass MWC architecture would be significant up to an order of magnitude for the analog preprocessing block interfacing with the sub-Nyquist sampling ADC as demonstrated in [39].

In the FRI and multiband frameworks, power reduction is achieved by reducing the sampling rates. However, in both these approaches, analog front ends are used to facilitate sub-

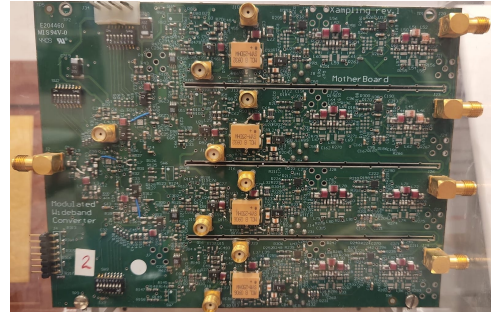


Fig. 8. A hardware realization of the MWC consisting of $M = 4$ channels [16], [17].

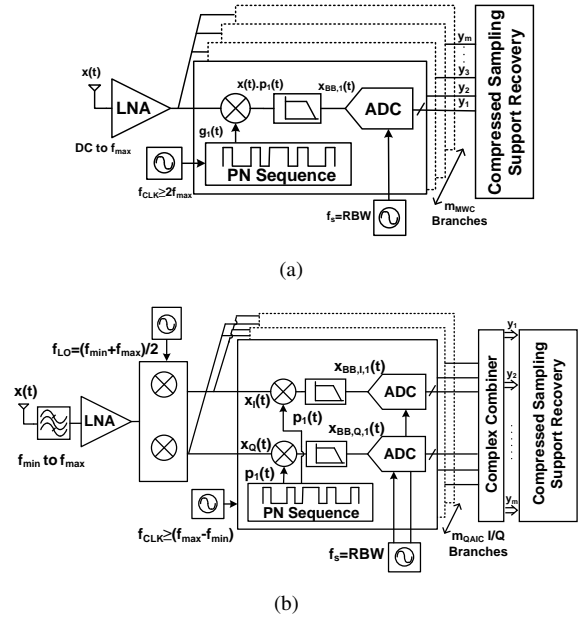


Fig. 9. Hardware prototypes for Sub-Nyquist sampling known as compressed sampling analog-to-information converter architectures; (a) Low-pass Modulated Wideband Converter (MWC) [16], [17], (b) Band-pass Quadrature Analog-to-Information Converter (QAIC) [39], [42].

Nyquist sampling. For example, in Fig. 7, M channels are used where each channel consists of a multiplier, a lowpass filter, and a PN sequence generator. These circuits can be built by using existing low-power circuits and hence do not add significant power gain to the overall circuits.

As in the case of FRI sampling, power saving is also applicable for multiband signals when the minimum sampling rate, $2NB$ Hz, is less than the Nyquist rate. Again, the savings amount depends on the actual sampling rate ratio to the Nyquist rate. In [16], [17], Mishali et al. proposed hardware prototypes for MWC-based sub-Nyquist samplers for $B = 19$ MHz, $N = 6$, and $1/T_{\text{Nyq}} = 2$ GHz, the signals can be sampled and reconstructed at a rate of 280 MHz by using a $M = 4$ channel MWC (See Fig. 8). By using the multiband structure, there is a seven-fold reduction in the sampling rate, which reduces the power requirements of the ADC by the same amount.

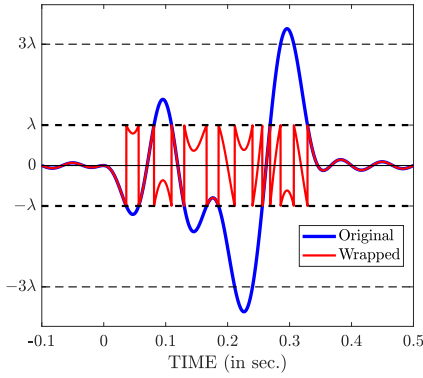


Fig. 10. An illustration of modulo folding of a bandlimited signal; the ADC's DR is $[-\lambda, \lambda]$, and signal is folded to stay within the range.

IV. MODULO-ADC

The higher the DR of an ADC, the higher the number of bits required to keep a low quantization error, which results in high power consumption. Hence, it is desirable to keep the DR low. However, to avoid clipping due to saturation, the signal's DR must be within that of ADC's. In many applications, it is not always possible to premeditate the signal's DR and then choose an ADC with a suitable DR. Moreover, the available ADCs have a fixed DR in many scenarios. An attenuator can be applied to address these mismatches in DRs that scale down the signal to fit into ADC's DR. This solution may not be suitable when the amplitude of the signal varies significantly over time. For such a signal, small components of the signals are scaled down below the noise floor and hence cannot be detected.

A. Theory and Algorithms

A modulo-folding operation was suggested by Bhandari et al. [4] that folds the signal before sampling such that its variations are contained within the ADC's DR. To be precise, let $f(t)$ be an analog signal to be sampled, which is bounded as $|f(t)| \leq A$. The ADC's DR is $[-\lambda, \lambda]$ where $\lambda < A$. To process $f(t)$ through the ADC, it is first folded as $f_\lambda(t) = \mathcal{M}_\lambda(f(t))$ where $\mathcal{M}(\cdot)$ is the folding operation defined as $\mathcal{M}_\lambda(a) = (a + \lambda) \bmod 2\lambda - \lambda$ for a real-valued a . The folding operation ensures that $|f_\lambda(t)| \leq \lambda$. The folded signal is then sampled using a low-DR ADC, which results in measurements $f_\lambda(nT_s)$ where $1/T_s$ is samples/sec. The folding operation does not discard any information of the analog signals; hence, the samples $f_\lambda(nT_s)$ retain necessary information of $f(t)$. In fact, mathematically, it was shown that bandlimited signals are uniquely identifiable from modulo samples provided that they are sampled above their Nyquist rate [4], [6]. Based on this discussion, the key message is that the dynamic range of a modulo-ADC, an ADC preceded by a modulo operation, is much higher (theoretically unlimited) than the ADC's DR. Hence, the modulo-based approach is termed a high-DR ADC. An example of modulo folding for a bandlimited signal is shown in Fig. 10 where the signal with high-dynamic range is folded to stay within the ADC's DR.

In the preceding paragraph, we mentioned that modulo-DR's ADC is unlimited ($A/\lambda \rightarrow \infty$), provided that the sampling rate is higher than the Nyquist rate. However, in practice, recovery of $f(t)$ from the folded samples $f_\lambda(nT_s)$ depends on A/λ and T_s . Typically, the reconstruction process works in two stages: first, determine the true samples $f(nT_s)$ from the folded samples by using an unfolding algorithm, and second, use standard reconstruction to get $f(t)$ from $f(nT_s)$. For bandlimited signals, the latter part requires that the sampling rate should be above the Nyquist rate. However, for unfolding, the signals are required to be oversampled where the oversampling factor ($\text{OF} = f_s/f_{\text{Nyq}}$) varies from algorithm to algorithm. Here f_{Nyq} is the Nyquist rate.

Unfolding algorithms [4]–[7] typically use the following decomposition of the modulo signal or samples.

$$f_\lambda(nT_s) = f(nT_s) + z(nT_s), \quad (13)$$

where $z(t)$ is a piecewise constant function whose values are multiple of 2λ and hence $z(nT_s) \in 2\lambda\mathbb{Z}$. Bhandari et al. [4] proposed a higher-order differences approach for unfolding. Let Δ^d be the d -th order sample difference operator. Then we have that $\mathcal{M}_\lambda(\Delta^d f_\lambda(nT_s)) = \mathcal{M}_\lambda(\mathcal{M}_\lambda(\Delta^d f(nT_s)) + \mathcal{M}_\lambda(\Delta^d z(nT_s)))$. The second term is zero as $\Delta^d z(nT_s) \in 2\lambda\mathbb{Z}$. Hence, if

$$|\Delta^d f_\lambda(nT_s)| < \lambda, \quad (14)$$

then we get $\mathcal{M}_\lambda(\Delta^d f_\lambda(nT_s)) = \Delta^d f(nT_s)$. In other words, we can determine the d -th order difference of the true samples from the folded samples provided that (14) is satisfied. The authors in [4] showed that the condition is satisfied for $\text{OF} \geq 17$, and they proposed an approach to recover $f(nT_s)$ from higher-order differences. The sampling rate is high for this algorithm, and the use of sample differences makes it unstable in the presence of noise.

A robust and sample-efficient algorithm is proposed in [6] where the authors used bandlimitedness and decay properties of the samples $f(nT_s)$ to extract them from $f_\lambda(nT_s)$. The main idea of the method and the corresponding optimization problem is discussed in the box *Beyond Bandwidth Residue Recovery (B^2R^2) Algorithm*. In this algorithm, for an acceptable accuracy of the reconstruction, the OF varies between 2 – 8 depending on the ratio $\theta = A/\lambda$ and the noise level. The modulo-ADC framework can also be extended to FRI and other classes of signals [7], [18]. For example, for the FRI sampling framework, a modulo folding operation is applied after the filtering stage (cf. Fig. 3), and an unfolding algorithm is used before the DSP block. As in the bandlimited case, it is required to sample above the rate of innovation for FRI signals when using a modulo-ADC. As a consequence, FRI sampling with modulo-ADC falls under the sub-Nyquist framework.

B. Power-Analysis of Modulo-ADC

We derive theoretical lower bounds on the power consumption of a modulo-ADC and compare it with a conventional ADC. To this end, we consider a conventional ADC with sampling rate f_s (which is equal to the Nyquist rate for a bandlimited signal or equal to RoI for an FRI signal) and

Beyond Bandwidth Residue Recovery (B^2R^2) Algorithm

We give a brief overview of the B^2R^2 algorithm [6] that can be used to unfold bandlimited signal samples. In the discussion, we assume that signals have finite energy and bandlimited to frequency interval $[-\omega_{\text{Nyq}}, \omega_{\text{Nyq}}]$. In addition, we assume that the sampling rate is above the Nyquist rate, that is, $\omega_s > 2\omega_{\text{Nyq}}$. The following two properties of a finite-energy bandlimited signal are used in this algorithm.

- 1) *Time-domain decay property*: From the Riemann-Lebesgue Lemma we have that $\lim_{|t| \rightarrow \infty} f(t) = 0$. The time-domain decay implies that for every $\lambda > 0$ there exists an integer N_λ such that $|f(nT_s)| < \lambda$, for all $|n| > N_\lambda$.
- 2) *Frequency-domain decay property*: Since the signal is sampled above the Nyquist rate, the discrete-time Fourier transform of the samples $f(nT_s)$ are zero beyond the bandwidth, Mathematically, $F(e^{j\omega T_s}) = 0$ for $|\omega| < \omega_s/2$.

By using the decomposition in (13) and the time-domain decay property, we have that $z(nT_s) = 0, |n| > N_\lambda$ as $f_\lambda(nT_s) = f(nT_s)$ for those samples.

Further, by using the bandlimitedness of the samples, we get $F_\lambda(e^{j\omega T_s}) = Z(e^{j\omega T_s})$, $\omega \in \rho = (-\omega_s/2, -\omega_{\text{Nyq}}) \cup (\omega_{\text{Nyq}}, \omega_s/2)$. Hence, the samples of the residue are compactly supported over $n \in [-N_\lambda, N_\lambda]$, and its DTFT is known over a finite-length interval ρ . Let vectors $\mathbf{f}_\lambda, \mathbf{z} \in \mathbb{R}^{2N_\lambda+1}$ denote samples $f_\lambda(nT_s)$ and $z(nT_s)$, respectively, for $n \leq |N_\lambda|$. In addition, let \mathcal{F}_ρ be a partial DTFT operator that evaluates Fourier measurements at frequencies in the set ρ . Then, from the above discussion, we can write the following optimization problem

$$\min_{\mathbf{z}} C(\mathbf{z}) = \frac{1}{2} \|\mathcal{F}_\rho \mathbf{f}_\lambda - \mathcal{F}_\rho \mathbf{z}\|^2 \quad \text{s.t.} \quad \mathbf{z} \in \mathcal{S}_{N_\lambda}, \quad (11)$$

where \mathcal{S}_{N_λ} is a set of sequences that have support on the interval $[-N_\lambda, N_\lambda]$.

Problem (11) can be solved using a projected gradient descent (PGD) method starting from an initial point $\mathbf{z}^0 \in \mathcal{S}_{N_\lambda}$. The steps at the k -th iteration are

$$\begin{aligned} \mathbf{y}^k &= \mathbf{z}^{k-1} - \gamma_k \nabla C(\mathbf{z}^{k-1}) \\ \mathbf{z}^k &= P_{\mathcal{S}_{N_\lambda}}(\mathbf{y}^k), \end{aligned} \quad (12)$$

where $\gamma_k > 0$ is a suitable step-size, $\nabla C(\mathbf{z}) = \mathcal{F}_\rho^* \mathcal{F}_\rho (\mathbf{z} - \mathbf{f}_\lambda)$ is the gradient of $C(\mathbf{z})$ and $P_{\mathcal{S}_{N_\lambda}}(\cdot)$ is the orthogonal projection onto \mathcal{S}_{N_λ} .

After estimating $z(nT_s)$ from the modulo samples, the unfolded samples are determined by (13).

input signals dynamic range A , as shown in Fig. 11. For a n -bit flash-type quantizer, the total power consumption is given by (3), which is also indicated in the table in Fig. 11. To determine the power of the modulo-ADC, we determine the power consumption of its constituents: a modulo-folding circuit and a conventional low-DR ADC (cf. Fig. 11). The low-DR ADC has a DR range $[-\lambda, \lambda]$ with n bit resolution and operates at an oversampling rate $\text{OF} f_s$. Following the derivations presented in Section II and assuming that the capacitance values of the low-DR ADC are the same as that of the conventional ADC, the power of the sampler and the comparator of the low-DR ADC are given as

$$\frac{24kT \text{OF} f_s 2^{2n}}{\theta^2}, \quad \text{and} \quad \frac{4n \ln 2 \text{OF} f_s C_C 2\lambda V_{eff}}{\theta},$$

respectively. By comparing these expressions with those of the conventional ADC, we note that the power consumption of the low-DR ADC P' is given as

$$P' = \frac{\text{OF}}{\theta^2} P, \quad (15)$$

where $P = P_S + P_C$ is the power of a conventional ADC.

From the analysis up to this point, we infer that for $\text{OF} < \theta^2$, the S/H with quantization part of a modulo-ADC consumes less power than a conventional ADC. For example, in the

modulo-ADC hardware prototype presented in [19], $\theta \leq 8$ and $\text{OF} \leq 5$, which is a factor of 12.8 reduction in power consumption. However, the feedback circuitry of a modulo-ADC will add additional power requirements, which we will consider next.

The goal of the modulo-circuit, or the feedback circuit, is to fold an input signal $f(t) \in [-A, A]$ to a dynamic range $[-\lambda, \lambda]$. The circuit consists of two comparators and a digital voltage generator (DVG). The input to the comparators is the modulo signal $f_\lambda(t)$. The comparators trigger whenever their input signal crosses $\pm\lambda$. The DVG reacts to the trigger signals and generates a residual signal $z(t) \in 2\lambda\mathbb{Z}$, which is added to the input signal to keep it within $[-\lambda, \lambda]$. Hence, we focus on the power utilization of the capacitors and the DVG. To this end, it is required to find an expression of the operating frequency f_{FB} of the feedback loop. The frequency f_{FB} depends on how fast the capacitors have to switch, and it could be higher than the sampling rate of the ADC ($= \text{OF} f_s$). Specifically, for any signal $f(t)$, f_{FB} denotes an upper bound on the number of times it crosses amplitude levels $\pm(2m+1)\lambda, m \in \mathbb{N}$. To derive an upper bound in f_{FB} for a bandlimited signal, we use the fact that the number of zeros of a bandlimited function is equal to the Nyquist rate. Then, by using a derivation similar to that in [7, Appendix], it can be shown that $f_{FB} = 2\theta f_{\text{Nyq}}$ where f_{Nyq} is the Nyquist rate.

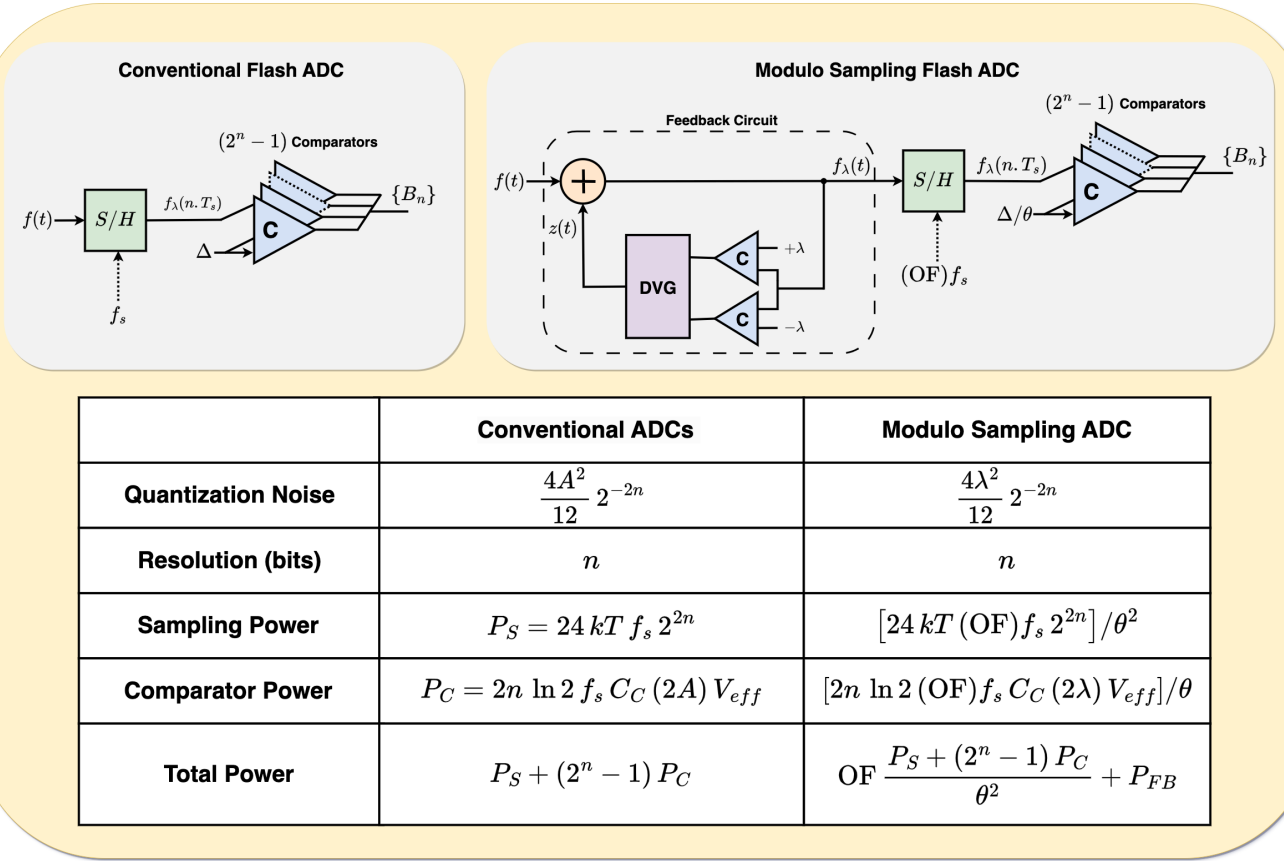


Fig. 11. Power consumption in a modulo sampling ADC and conventional ADC.

Based on f_{FB} , the power consumption of the comparators can be modeled similarly to (2). However, the power consumption of the comparator is proportional to $\ln \theta$ instead of $2n \ln 2$ since the comparators need to resolve the voltage of 2λ and generate a voltage level of A at the output. Therefore, we rewrite the lower bound as (16) where $C_{C,FB}$ is the latch comparator load capacitance. The power consumption of the DVG is modeled as the power consumption of a Capacitive Digital-to-Analog Converter (C-DAC) since it scales down or up the output voltage in multiples of 2λ with reference voltage A . We estimate the average power consumption of the DVG as (17) where C_{DAC} and N are the total C-DAC capacitance and bit resolution of the C-DAC in the feedback circuit respectively [43],

$$P_{C,FB} = \ln \theta f_{FB} C_{C,FB} V_{eff} A, \quad (16)$$

$$P_{DVG} = P_{DAC} = 0.66 \frac{f_{FB} C_{DAC} A^2}{N + 1}, \quad (17)$$

$$P_{FB} = P_{DVG} + 2 P_{C,FB}. \quad (18)$$

The total power consumption of the feedback circuit shown in (18) is the sum of DVG and two comparators. The total

power consumption of the modulo-ADC is given as

$$P_{\text{mod-ADC}} = P' + P_{FB} = \frac{OF}{\theta^2} P + P_{FB}. \quad (19)$$

For a fixed f_s , A , and n , we note that power reduction in the sampling and quantization part of the modulo-ADC is inversely proportional to θ^2 whereas that in the modulo-circuit is proportional to $\theta \ln \theta$. The capacitance of the C-DAC can be minimized to reduce the power consumption of the DVG of a modulo-circuit under the noise and linearity limitations. For example, with a unit capacitance value of 10 fF, a 4-bit C-DAC can be used that allows up to $\theta = 16$. In this case, we have $C_{DAC} = 2^4(10 \text{ fF}) = 160 \text{ fF}$. Now, if A and f_{FB} are assumed to be in 3.3 V and 10 MHz ranges, respectively, this results in $P_{DVG} \simeq 2.3 \mu\text{W}$ which is much less compared to the power consumption of the state-of-the-art ADC in [23] that operates close to this frequency range. Hence, by carefully optimizing the feedback circuit, a modulo-ADC can have lower power requirements compared to a conventional ADC for any $\theta > 1$.

In the context of power saving for modulo-ADC, we would also like to emphasize that the oversampling factor, which depends on the unfolding algorithm, may not be independent of θ . In fact, as shown via simulations in [6], the OF for most algorithms increases with θ . In such cases, the power saving in a modulo ADC still holds if OF does not increase faster than θ^2 . Hence, designing sample-efficient algorithms plays a key role in the power-saving aspect of a modulo-ADC.

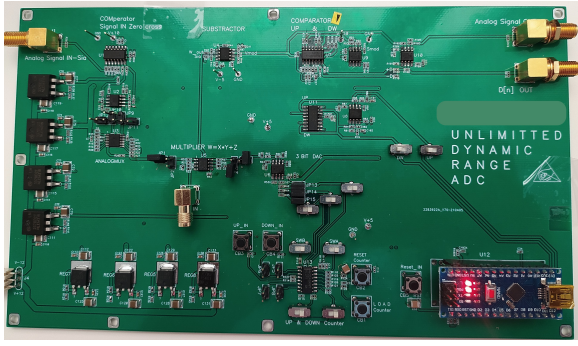


Fig. 12. A hardware prototype of modulo-ADC [19].

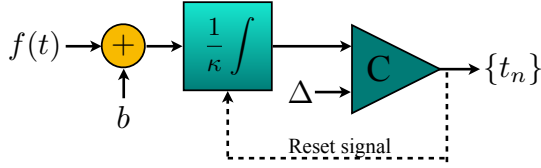


Fig. 13. A schematic of IF-TEM

A few works presented hardware prototypes of modulo-ADCs [5], [18], [19]. The emphasis in these works is to show that one can fold the analog signal to a low-dynamic range and the applicability of the unfolding algorithms. In [5], [18], the authors consider different signal models, hardware limitations, and algorithms; however, they did not provide details of the hardware implementation. Moreover, the results are presented for signals up to 300 Hz. A hardware prototype, shown in Fig. 12, for signals up to 10 kHz, with detailed circuitry, is presented in [19]. The authors showed that, for FRI signals, the system operates six times below the Nyquist rate, and the modulo ADC can sample signals that are eight times larger than the ADC's DR.

V. ASYNCHRONOUS SAMPLING

Sampling approaches discussed in the previous sections are based on uniform sampling where a continuous-time signal $f(t)$ is represented by $f(nT_s)$, $T_s > 0$, $n \in \mathbb{Z}$. A key advantage of uniform sampling is that it allows the application of standard Fourier analysis. This, in turn, relates the spectra of $f(t)$ and $f(nT_s)$ and often leads to closed-form expressions for reconstruction, such as Shannon-Nyquist recovery. Despite being widely used, a shortcoming of uniform sampling is that it measures samples at the same rate, usually dictated by the maximum signal component, and ignores whether the signal is changing rapidly or not locally. Specifically, the instantaneous bandwidth of most real-world signals varies across time, even if their overall bandwidth is fixed. In such scenarios, it is power-efficient to measure samples at a low rate in the regions with small changes and vice-versa.

To this end, event-based samplers such as level-crossing sampling [8] and integrated and fire time-encoding machines (IF-TEM) [9]–[12] are applied where the sampling locations are a function of the analog signal. In these frameworks, analog signals are represented by a set of time instants $\{t_n\}$ instead of their values at a specific time instants as in uniform

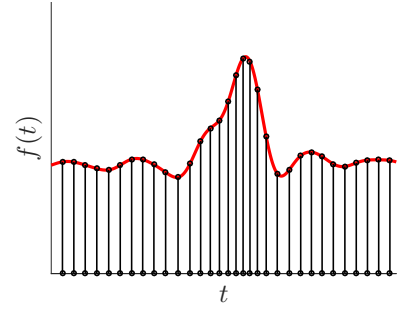


Fig. 14. IF-TEM samples of a bandlimited signal: the samples are dense when the signal changes rapidly compared to the regions of slow variations.

and random sampling. For example, in a popular level-crossing sampling architecture called zero-crossing detector, the signal is represented by its zero-crossing instants [8]. In general level-crossing, a signal is represented by time instants at which the signal crosses a set of predefined amplitude levels.

VI. THEORY AND ALGORITHMS

The level-crossing-based approach requires multiple comparators, one for each level, which could be inefficient for hardware implementation, although more power-efficient implementations were demonstrated [44]. On the other hand, IF-TEMs use a single comparator to time encode analog signals. In an IF-TEM, an analog signal $f(t)$ is added to a bias b such that $b + f(t) \geq 0$. This is ensured by choosing the bias as $b > A$ where $|f(t)| \leq A$. Thus, bias b is determined by the input signal dynamic range. The positive signal is scaled as $\frac{1}{\kappa}(b + f(t))$ where $\kappa > 0$ and then integrated. The resulting signal is compared with a threshold Δ as shown in Fig. 13. When the value of the output of the integrator reaches threshold Δ , defined by input signal level, bias b , and signal bandwidth, the time instant is recorded, and the integral is reset. The process is repeated to get the time encodings $\{t_n\}$, which are digital representations of $f(t)$ and is a function of IF-TEM parameters $\{b, \kappa, \Delta\}$. As discussed next, the reconstruction of $f(t)$ is based on non-uniform sampling-based recovery methods.

Following the sampling mechanism of IF-TEM discussed above, $f(t)$ and its time encodings are related as

$$\frac{1}{\kappa} \int_{t=t_{n-1}}^{t_n} (b + f(t)) dt = \Delta. \quad (21)$$

From this equation and by using the boundedness $|f(t)| < c$, the difference between consecutive firings are related as

$$\frac{\kappa\Delta}{(b+A)} \leq t_n - t_{n-1} \leq \frac{\kappa\Delta}{(b-A)}. \quad (22)$$

By using these bounds, the maximum and minimum number of time encodings per second or *firing rates* (FR) are given as

$$\text{FR}_{\max} = \frac{(b+A)}{\kappa\Delta}, \quad \text{and} \quad \text{FR}_{\min} = \frac{(b-A)}{\kappa\Delta}, \quad (23)$$

respectively. The notion of FR has a similar meaning to that of sampling rate in uniform sampling or sampling density (average number of samples per second) in random sampling. The

FRI Signal Reconstruction From IF-TEM Samples

To determine FRI parameters from IF-TEM samples, we assume that the FRI signal is filtered through an SoS kernel as in (6) and the output signal $y(t)$ (cf. (7)) is sampled using IF-TEM. Let the time encodings be $\{t_n\}_{n=1}^N$. Using (21), we compute the following amplitude measurements from time encodings: $y_n = \kappa\Delta - b(t_n - t_{n-1}) = \int_{t=t_{n-1}}^{t_n} y(t) dt$, $n = 1, \dots, N-1$. From trigonometric polynomial form of $y(t)$ in (7), the Fourier samples are related to the measurements as

$$y_n = \sum_{k \in \{-K, \dots, K\} \setminus 0} F(k\omega_0) \frac{(e^{jk\omega_0 t_{n+1}} - e^{jk\omega_0 t_n})}{jk\omega_0} + F(0)(t_{n+1} - t_n), \quad n = 1, \dots, N-1, \quad (20)$$

where $K \geq L$. It was shown that the Fourier samples $\{F(k\omega_0)\}_{k=-K}^K$ are uniquely determined from the measurements $\{y_n\}_{n=1}^{N-1}$ using (20) if $N \geq 2K + 2$ [12]. The condition $N \geq 2K + 2$ can be satisfied by choosing the IF-TEM parameters as $\text{FR}_{\min} > \frac{2K+1}{t_{\max}}$ where t_{\max} is maximum time delays of the FRI signal. The latter condition is similar to that for reconstructing bandlimited signals from IF-TEM samples where it is required that FR_{\min} is greater than the Nyquist rate.

After the Fourier samples are obtained from the time encodings, the FRI parameters can be computed using either Prony's approach or other robust methods, as discussed previously.

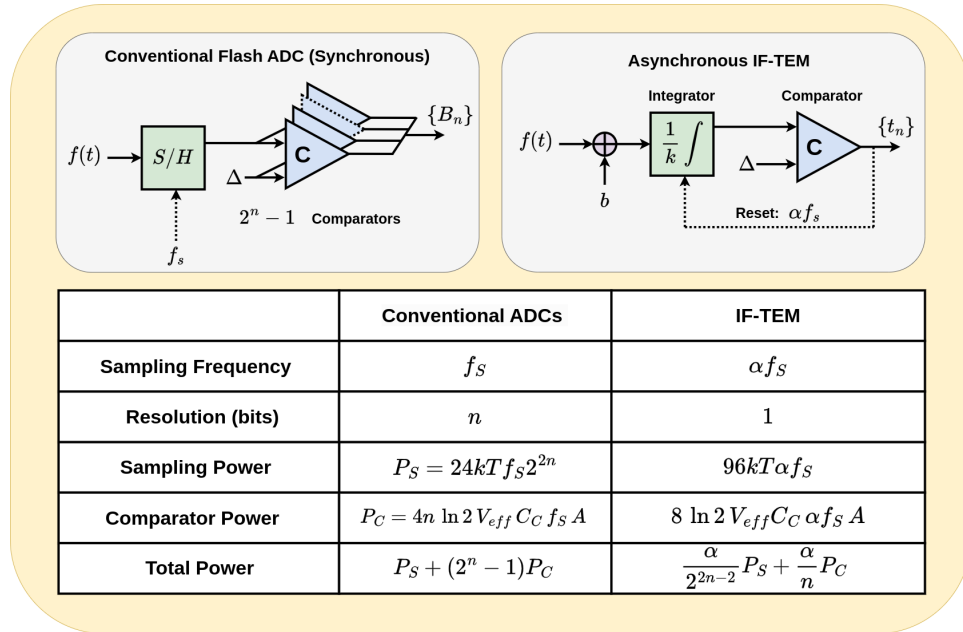


Fig. 15. Architecture and power consumption estimation comparison between conventional Flash ADCs and IF-TEM, showing corresponding power savings for the latter system.

difference is that FR is signal-dependent, whereas sampling density or rate is fixed. The FR is higher when the signal changes rapidly compared to when there are slow variations in the signal. An illustrative example is shown in Fig. 14.

For perfect reconstruction of $f(t)$, the IF-TEM parameters should be chosen such that FR_{\min} is above a certain rate. For example, Lazar and Tóth [9] used an iterative algorithm for the reconstruction of bandlimited signals and showed that the algorithm converges to the true signal provided that FR_{\min} is above the Nyquist rate. Similarly, [12] showed that FRI signals could be reconstructed from their time encodings if the minimum firing rate is set above the rate of innovation. The reconstruction approach is summarized in the box *FRI Signal Reconstruction From IF-TEM Samples*.

A. Power Efficiency of IF-TEMs

In the following, we determine the power an IF-TEM circuit consumes and compare it with that in a conventional ADC for the same input signal. The details of our analysis are depicted in Fig. 15. For the conventional ADC, we use (3) to represent its power consumption. For the IF-TEM, we note that there are three components: (i) an adder with a bias, (ii) an integrator with a reset rate αf_s , where α is a non-uniform sampling activity factor that reflects non-periodical sampling activity, and (iii) a comparator with a threshold Δ . Among these, the integrator and comparator are power-hungry blocks, and we shall focus on them in the following analysis.

Quantization of time-encodings utilized in IF-TEMs is usually not carried out in the same way as in conventional

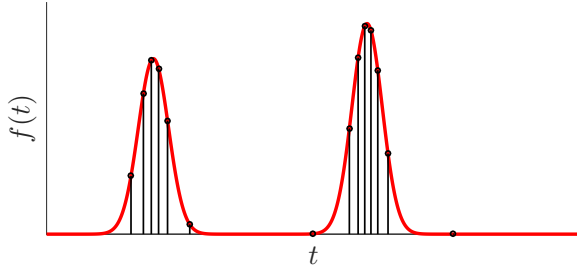


Fig. 16. IF-TEM sampling by using burst signals with positive amplitudes: We observe that there are no samples when the signal is inactive, whereas conventional ADC would sample during the entire time interval.

ADCs, but specialized time-to-digital circuits (TDCs) are used [45]. The underlying principle in TDCs is that time can be divided into measurable intervals that are quantized into digital representations or bits. There are several ways of realizing a TDC, such as analog-based, counter-based, or delay-line. (please refer to [45] for details). Among these, most digital implementation-based TDCs have an order of magnitude lower power consumption than typical analog circuits. This is the reason we did not account for the TDC in the power calculation of the IF-TEM framework.

For the calculation of the power consumption, we note that the integrator can be viewed as an active sample-and-hold (a capacitor and a switch) with $n = 1$ bit resolution [46], since the number of comparators defines the system resolution, similar to conventional flash ADCs. Hence, integrator power consumption is given as

$$P_{\text{INT}} = 96kT\alpha f_s, \quad (24)$$

where αf_s is the switching or reset rate of the integrator. It is purposely written in terms of the sampling rate f_s of the conventional ADC (See Fig. 15). The sampling activity factor α shows how the non-uniform switching rate of an integrator relates to the sampling frequency of conventional ADCs. For example, in the case when the average switching rate of an integrator approaches the conventional ADC sampling frequency f_s in case of a fast-changing signal, as illustrated in Fig. 14, α approaches or goes above the value of 1. If the average sampling rate (switching rate) of an IF-TEM is much lower than that of the conventional ADC in the case of burst signals, with activity only in some time intervals (as in Fig. 16), then $\alpha \ll 1$, an important case for event-driven low-power data conversion systems. Our power analysis reflects these changes in the activity factor.

Overall, The sampling activity factor $\alpha > 0$ is a quantity that depends on the signal to be sampled and the reconstruction methods, as discussed in the following. The reset rate has a similar notion as the f_{FB} in the case of modulo-ADC discussed in Section IV. To compute the comparator's power consumption, we use (2) with an assumption that the comparator has the same characteristics as the comparators in the flash ADC. Then, for $n = 1$, the comparator power is given as

$$P_{C,\text{IF-TEM}} = 8 \ln 2 V_{\text{eff}} C_C \alpha f_s A. \quad (25)$$

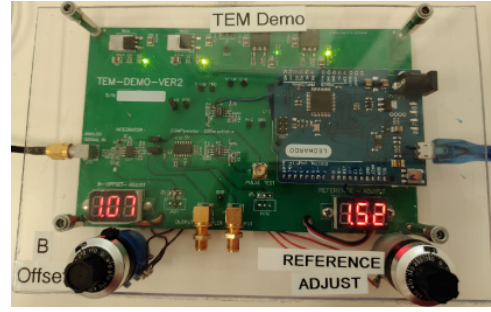


Fig. 17. A hardware prototype of IF-TEM Sampler [20].

The total estimated power of the IF-TEM is

$$\begin{aligned} P_{\text{IF-TEM}} &= P_{\text{INT}} + P_{C,\text{IF-TEM}}, \\ &= 96kT\alpha f_s + 8 \ln 2 V_{\text{eff}} C_C \alpha f_s A, \\ &= \frac{\alpha}{2^{2n-2}} P_S + \frac{2\alpha}{(2^n - 1)n} ((2^n - 1)P_C), \end{aligned} \quad (26)$$

where P_S and $(2^n - 1)P_C$ are power requirements of the S/H and quantizer parts (cf. (1)), respectively, in a conventional ADC. We assumed that the comparators were the same for both the conventional ADC and the IF-TEM.

From (26), we note that for $\alpha < \max(2^{2n-2}, (2^n - 1)n)$, where $n > 1$, $P_{\text{IF-TEM}}$ is smaller than the power of a conventional ADC. The power saving is largely due to the fact that there is only a single comparator in the IF-TEM compared to the conventional flash-ADC, where there are $2^n - 1$. Even if one uses a different quantizer than the flash-type one, the number of comparators can be reduced at the expense of other circuitry; still, the power saving of IF-TEM can be reduced if $\alpha < 1$. Hence, in the aforementioned analysis, α plays a key role, and next, we discuss its dependency on the class of signals and reconstruction algorithm.

First, we consider bandlimited signals whose Nyquist rate is f_{Nyq} , and hence the conventional ADC operates at the minimum rate $f_s = f_{\text{Nyq}}$. For these signals, as discussed earlier, reconstruction from IF-TEM samples is possible if $\text{FR}_{\text{min}} > f_{\text{Nyq}}$ [9]. If for a given signal's dynamic range A , we choose IF-TEM parameters $\{b, \kappa, \Delta\}$ such that $\text{FR}_{\text{min}} > f_{\text{Nyq}}$ is satisfied, then from (23) we know that the firing rate could be much higher than the Nyquist rate. Specifically, in $\alpha > \frac{b+A}{b-A} > 1$. This implies that IF-TEM always fires even if the input signal's amplitude does not change significantly. This fact is also depicted in Fig. 14. For bandlimited signals, a large value of α above one reduces the effective power saving even if $\alpha < \max(2^{2n-2}, (2^n - 1)n)$.

On the other hand, signals that are not necessarily bandlimited but are bursty in nature, where the signal's amplitude is above a certain level only in a few intervals, are much more suited for sampling by using an IF-TEM. Assume that the reconstruction is not a sinc-based interpolation as in the Shannon-Nyquist framework but uses a local interpolation such as a linear or a spline-based one. In such scenarios, FR_{min} can be set to zero, which implies there are no firings when the signal is not active, and the signal is reconstructed, up to a given accuracy, from the time encodings in the

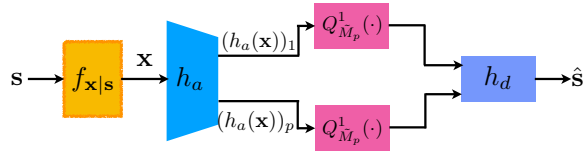


Fig. 18. A general framework of task-based quantizer

active period. In this case, $\alpha < 1$. An example of such a burst signal, which is similar to an FRI signal, is shown in Fig. 16, where the signal is active in certain time intervals and zero elsewhere. We assume that the signal is positive such that $0 \leq f(t) \leq A$ and hence bias is not added, that is, $b = 0$. Then, for $\kappa = 1$ and a particular value of Δ , we note that there are no firings during the inactive period of the signal. This implies that the integrator and the comparator were not activated, i.e., in sleep mode, in those intervals. However, if a conventional ADC is used to sample such signals, then they operate in both active and inactive modes and consume relatively higher power. Because of this reason, time-encoding-based ADCs, including IF-TEMs, are preferable in event detection applications such as bio-signal monitoring in wearable devices, voice-activity detection by voice-based virtual assistant platforms, and more. On another application front, even-based cameras are an emerging field where it is shown that they have higher resolution and lower power consumption than conventional cameras.

Several integrated circuit implementations of level-crossing-based systems are available [47]. However, there is very little work on IF-TEMs. In [20], a hardware prototype of an IF-TEM ADC was presented, where reconstruction of FRI signals from time encodings was demonstrated (see Fig. 17). The authors showed that the firing rate is ten times below the Nyquist rate of the signals.

VII. LOW-BIT QUANTIZATION

In the previous sampling frameworks, the focus was on reducing the sampling rate and the dynamic range of the signal to reduce power consumption. In this section, we discuss methods to reduce the number of quantization bits by exploring the task at hand. For signal reconstruction with acceptable accuracy, samples should be encoded with a sufficiently large number of bits. However, in many applications, the task is not signal reconstruction but extracting some information from the analog signal [13]. For example, in a MIMO communication system, it is necessary to estimate the channel, which is a low-dimensional task compared to underlying analog signal recovery. In such applications, task-based quantization was proposed where by using analog combiners and joint optimization of an analog front end and digital back end, the number of quantization bits can be significantly reduced [13].

A. Theoretical Framework of Task-Based Quantizer

To understand the task-based low-bit quantization system, let us consider a discrete model where we assume that the signals are already sampled [13]. The assumption helps us in focusing only on the quantization part. The general framework

of task-based quantization is shown in Fig. 18 where the goal is to estimate a low-dimensional vector or task $\mathbf{s} \in \mathbb{R}^K$ from measurements $\mathbf{x} \in \mathbb{R}^N$ where $K < N$. The information \mathbf{s} is statistically dependent on observations \mathbf{x} via a conditional distribution $f_{\mathbf{x}|\mathbf{s}}$. For example, \mathbf{x} represents directions of arrivals, and \mathbf{y} is a snapshot of N sensors or antennas for the case of radio receivers. Instead of individually quantizing entries of \mathbf{x} by using N scalar quantizers, the measurement \mathbf{y} is projected to a low-dimensional space by using a combiner $h_a: \mathbb{R}^N \rightarrow \mathbb{R}^P$. Each component of the combiner, $(h_a(\mathbf{x}))_p, 1 \leq p \leq P$ is independently quantized by a scalar quantizer $Q_{M_P}^1(\cdot)$ to get quantized measurements $Q_{M_P}^1((h_a(\mathbf{x}))_p), 1 \leq p \leq P$. Here, M_P represents the number of levels of each quantizer. The vector \mathbf{s} is estimated from the quantized measurements as

$$\hat{\mathbf{s}} = h_d \left(Q_{M_P}^1((h_a(\mathbf{x}))_1), \dots, Q_{M_P}^1((h_a(\mathbf{x}))_P) \right), \quad (27)$$

where $h_d(\cdot): \mathbb{R}^P \rightarrow \mathbb{R}^K$ is an estimator. The aforementioned system is hardware-limited by assuming that the number of quantization levels available, M , is fixed. This implies that, M_P should satisfy the inequality $M_P \leq M^{\frac{1}{P}}$.

The goal of a hardware-limited task-based system is to jointly optimize the combiner $h_a(\cdot)$, quantizer Q_{M_P} , and estimator $h_d(\cdot)$. The problem is intractable without making explicit assumptions about the stochastic model $f_{\mathbf{x}|\mathbf{s}}$. For example, one can consider a linear model $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{w}$ where \mathbf{w} is noise [13]. Such models are common in many estimation tasks. Further, the combiner and estimator are assumed to be linear and represented by matrices. Within this linear setting, authors in [13] derived expressions for the combiner matrix and estimator that minimizes mean-squared error in the estimation of \mathbf{s} for a fixed M . The results are also extended to the measurements related to \mathbf{s} in a quadratic fashion [14]. In a nutshell, by reducing the dimension of the measurements from N to P , one can assign more bits during quantization when the overall number of bits is fixed. This improves the estimation accuracy.

Deep learning-based methods are also applied when the observations and unknowns, such as channel parameters, can not be represented linearly or quadratically [48]. Further, at the application front, task-based methods are applied to massive MIMO communications for estimation of the underlying channel from the high-dimensional received signals [49].

B. Task-Based ADC Hardware

A hardware prototype for MIMO channel reduction with the task of estimating the underlying channel is presented in [21].

One of the main goals of the low-bit quantization using a task-based (Task-Specific) framework is to reduce the overall system's power consumption. Implications of task-based receiver system design are considered in detail in this section. First of all, as previously mentioned, the power consumption of the ADCs is directly proportional to the number of bits, as suggested by Walden's FoM [2].

The task-based framework discussed in the earlier section was shown to reduce the quantization rate quite significantly (x2-6 times). However, additional analog processing

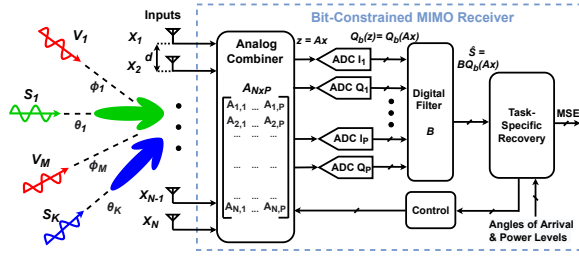


Fig. 19. Task-Specific hybrid MIMO receiver system utilizing low-quantization rate ADCs and power-efficient analog combiner [50].

is required for such systems, resulting in increased analog hardware complexity directly tied to power consumption, which is especially critical in MIMO receivers since analog processing overhead is proportional to the hardware processing channels. In fact, power efficiency, hardware complexity, and signal recovery accuracy are critical performance trade-offs for MIMO modern communication systems.

In [50], energy-efficient analog processing together with low-quantization ADCs and task-specific processing in application to MIMO receivers, was developed. The recent MIMO developments particularly focus on communication in congested environments, where both desired signals and undesired jammers arrive from various directions at the receiver.

Generally, the signal model at the MIMO receiver can be described as vector \mathbf{x} is an input observation vector consisting of a linear combination of K desired signals (*tasks*) arriving from corresponding angles θ_K . In addition, a set of M undesired signals (spatial blockers) arriving from respective angles ϕ_K , as shown in Fig. 19. Input observation signals are first processed in the analog combiner \mathbf{A} , performing linear matrix multiplication with signal dimensionality reduction, similar to the $h_a(\cdot)$ combiner in the task-based quantization system discussed earlier (Fig. 18). The output of the analog combiner is quantized using low-resolution ADCs (modeled as a scalar quantizer $Q_b(\cdot)$) and processed in the digital domain using matrix \mathbf{B} , similar to estimator $h_d(\cdot)$, obtaining desired signal (task) estimate $\hat{\mathbf{s}} = \mathbf{B}Q_b(\mathbf{A}\mathbf{x})$.

Several hardware design considerations were accounted for

Algorithm 1: Analog combiner setting

Data: Fix $\mathbf{A}^{(0)}$

```

1 for  $k = 1, 2, \dots, k_{\max}$  do
2   Update  $\mathbf{A}^{(k)} \leftarrow \mathcal{O}_{\mathcal{L}}(\mathbf{A}^{(k-1)})$ 
3   if  $\text{mod}(k, k_{\text{proj}}) = 0$  then
4     Project via  $\mathbf{A}^{(k)} \leftarrow \mathcal{P}_{\mathcal{A}}(\mathbf{A}^{(k)})$ 
5   end
6 end
Output: Analog combiner  $\mathbf{A}^{(k_{\max})}$ .

```

in the solution algorithm. First, the system targets low-bit ADC *task* recovery without compromising recovery accuracy. Second, while MSE focuses on *task* recovery, spatial blocker impression in the analog domain before analog-to-digital conversion is embedded to avoid receiver desensitization and increased dynamic range requirement of the ADCs. Finally, the system targets a highly power-efficient operation of the analog combiner to reduce the overhead from analog pre-processing. The details of the Task-Specific MIMO optimization algorithm are provided in the box and algorithm *Task-Specific MIMO Recovery*.

The power consumption of a MIMO receiver can be estimated from the power consumption of the sub-blocks of the system. A conventional MIMO receiver hardware, shown in Fig. 20 (top left), consists of an analog signal acquisition block, such as low-noise amplifiers, variable gain amplifiers, and phase shifters that perform some form of signal processing in the analog domain, for example, for beamforming. For analysis simplicity, we considered an implementation using Vector Modulators (VMs), a circuit component that applies simultaneous complex gain and phase shift on the input signal. The next crucial component is a mixer, which performs signal downconversion from carrier to baseband frequencies. Signals are then amplified using baseband amplifiers and quantized using ADCs. With N input signals and N output streams, conventional MIMO system power consumption is expressed

Task-Specific MIMO Recovery Algorithm

Task-specific MIMO receiver design constraints can be accounted for in the convex optimization framework by formulating a loss measure:

$$\mathcal{L}(\mathbf{A}) = \text{MSE}(\mathbf{A}) + \gamma_I \text{IntRej}(\mathbf{A}) + \gamma_S \|\mathbf{A}\|_{1,1}. \quad (28)$$

Here $\|\cdot\|_{1,1}$ is the entry-wise ℓ_1 norm operator, while $\gamma_I, \gamma_S \geq 0$ are regularization coefficients, balancing the contribution of task recovery MSE, spatial interferer rejection, and sparsity level of the analog combiner in the overall loss measure. In addition, although (28) is expressed in the convex term, the final solution, accounting for finite VM resolution, is defined over a discrete domain space. The final objective is formulated as:

$$\mathbf{A} = \arg \min_{\mathbf{A} \in \mathcal{A}^{P \times N}} \mathcal{L}(\mathbf{A}) \quad (29)$$

The recovery algorithm performs k_{\max} rounds of a convex optimizer for minimizing $\mathcal{L}(\mathbf{A})$ over $\mathcal{C}^{P \times N}$, with periodic projections onto the discrete \mathcal{A} . Algorithm 1 summarizes the design procedure.

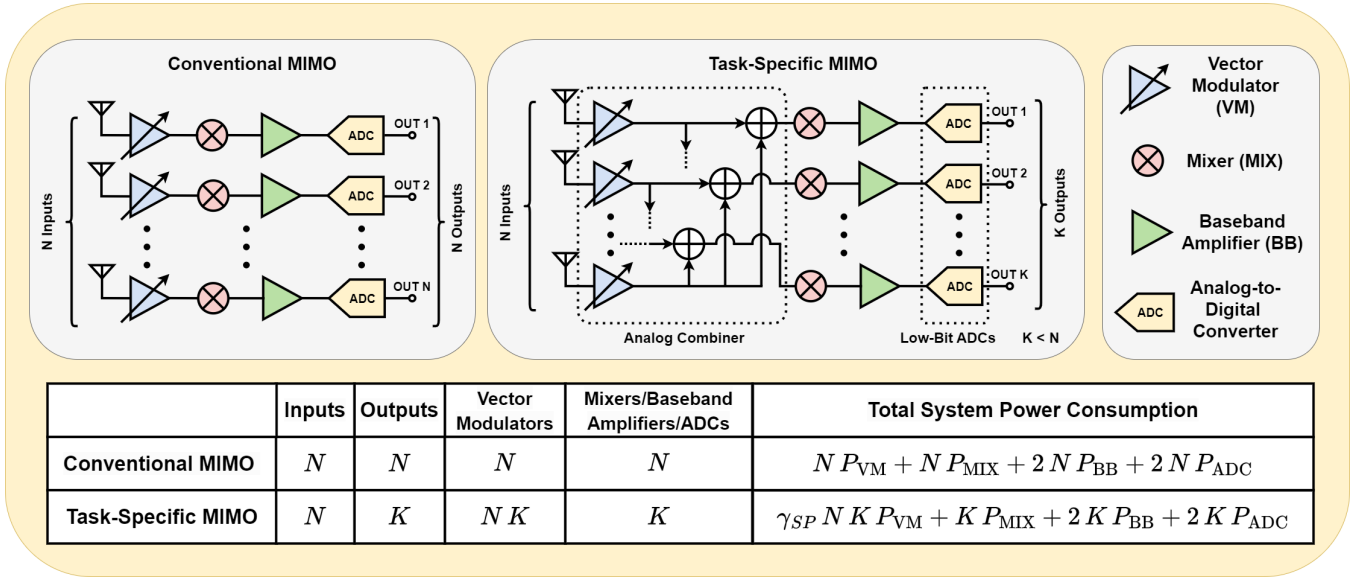


Fig. 20. Architecture and power consumption estimation comparison between conventional MIMO and Task-Specific MIMO.

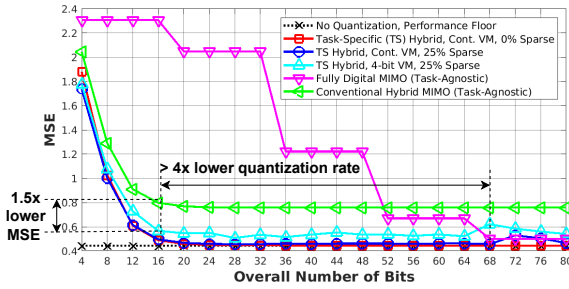


Fig. 21. MSE vs. the total number of quantization bits for recovering 2 desired signals in the presence of 2 spatial interferers [50]. Task-specific solutions are compared to conventional hybrid and fully digital recovery approaches.

as:

$$P_{\text{Conv MIMO}} = N P_{VM} + N P_{MIX} + 2 N P_{BB} + 2 N P_{ADC}, \quad (30)$$

where P_{VM} is the power consumed by the vector modulator, P_{MIX} is the downconversion mixer consumption, and P_{BB} and P_{ADC} are baseband amplifiers and ADCs power consumption, respectively, each doubled to account for the quadrature I and Q paths.

The Task-Specific MIMO receiver, on the other hand, first processes input signals in the analog combiner, which performs analog pre-processing with dimensionality reduction. For a system with N input signals and K output streams, where $K < N$, the number of required VMs is $N \times K$. However, due to the lower number of output signals, only K baseband processing chains are needed. In addition to that, one has to consider optimal analog pre-processing circuit design, such that the power overhead from the pre-quantizing analog combiner would not exceed power savings enabled by the low quantization rate. Overall analog combiner power consumption depends on the number of active VM elements in \mathbf{A} and individual VM power consumption, which is proportional to

the VM resolution. By utilizing sparsification of \mathbf{A} (zeroing particular elements of matrix \mathbf{A}), implemented as keeping the corresponding VM in sleep mode, and incorporating fine and coarsely quantized VMs, additional power savings in the analog combiner can be achieved. Hardware non-idealities, in this case, are accounted for in the back-end convex optimization algorithm.

The resulting system power consumption expression is:

$$P_{\text{TS MIMO}} = \gamma_{SP} N K P_{VM} + K P_{MIX} + 2 K P_{BB} + 2 K P_{ADC}, \quad (31)$$

Here, γ_{SP} is the analog combiner sparsity coefficient: $\gamma_{SP} = 1$ denotes a non-sparse \mathbf{A} , while, a practical value of $\gamma_{SP} = 0.75$, which corresponds to 25% sparsity, was used in the system. Fig. 20 features Task-Specific MIMO receiver diagram (top right) and system parameters with power consumption estimation expressions. It is important to note that Task-Specific MIMO incorporates additional power savings from low-resolution VMs and ADCs.

The proposed solution implementation was simulated in Matlab in [50]. The numerical results in Fig. 21 show that the quantization rate can be reduced by a factor of 4 compared to conventional fully digital MIMO for the same accuracy. In addition, power consumption estimation based on state-of-the-art integrated components shows that more than 58% power reduction can be achieved compared to the task-agnostic solution by incorporating sparsity of the analog combiner, variable-resolution VMs, and low-resolution ADCs while suppressing interferers by more than 36 dB [50].

Regarding the ADC power consumption estimation, we recall from the discussion above that typical Flash ADC power is estimated as $P_{SH} + (2^n - 1)P_{CADC}$, where n is the number of bits. Although theoretical power-saving gain should be exponentially proportional to the reduction of the number of bits, practical ADC implementations exhibit power consumption overhead coming from other peripheral circuits

(digital processing and biasing network), especially at a low resolution when consumption of those sub-blocks becomes more dominant. This creates a power consumption floor for low-resolution ADCs, as demonstrated in the ADC performance survey [23].

VIII. ACKNOWLEDGEMENT

We would like to acknowledge several authors' publications and their research works that are relevant to this review. We could not cite them due to a strict restriction on the number of references.

IX. CONCLUSION

Low-power ADCs are key for emerging applications. In this review, we discussed four approaches that can be applied to reduce the power consumption of ADCs. Among these techniques, three directly reduce power by reducing the sampling rate, dynamic range, and resolution of an ADC, whereas the fourth method relies on a time-based discrete representation of the signal. Any two or more of these methods can be combined together to make the sampling process even more power efficient. A common feature of the four frameworks is to have a preprocessing analog front-end, a low-power ADC, and a digital signal processor. For all these four frameworks, we discussed power consumption and hardware prototypes that are developed in our labs. These power-efficient ADCs play a crucial role in designing compact, portable medical and communication devices.

REFERENCES

- [1] T. Sundstrom, B. Murmann, and C. Svensson, "Power dissipation bounds for high-speed nyquist analog-to-digital converters," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 3, pp. 509–518, 2009.
- [2] R. H. Walden, "Performance trends for analog to digital converters," *IEEE Comm. Mag.*, vol. 37, no. 2, pp. 96–101, Feb. 1999.
- [3] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge University Press, 2015.
- [4] A. Bhandari, F. Krahmer, and R. Raskar, "On unlimited sampling and reconstruction," *IEEE Trans. Signal Process.*, vol. 69, pp. 3827–3839, 2020.
- [5] A. Bhandari, F. Krahmer, and T. Poskitt, "Unlimited sampling from theory to practice: Fourier-Prony recovery and prototype ADC," *IEEE Trans. Signal Process.*, vol. 70, pp. 1131–1141, 2022.
- [6] E. Azar, S. Mulleti, and Y. C. Eldar, "Robust unlimited sampling beyond modulo," *arXiv preprint arXiv:2206.14656*, 2022.
- [7] S. Mulleti and Y. C. Eldar, "Modulo sampling of FRI signals," *arXiv preprint arXiv:2207.08774*, 2022.
- [8] B. F. Logan Jr, "Information in the zero crossings of bandpass signals," *Bell Sys. Technical J.*, vol. 56, no. 4, pp. 487–510, 1977.
- [9] A. A. Lazar and L. T. Tóth, "Perfect recovery and sensitivity analysis of time encoded bandlimited signals," *IEEE Trans Circuits and Syst. I: Regular Papers*, vol. 51, no. 10, pp. 2060–2073, 2004.
- [10] D. Gontier and M. Vetterli, "Sampling based on timing: Time encoding machines on shift-invariant subspaces," *Applied and Comput. Harmonic Anal.*, vol. 36, no. 1, pp. 63–78, 2014.
- [11] R. Alexandru and P. L. Dragotti, "Reconstructing classes of non-bandlimited signals from time encoded information," *IEEE Trans. Signal Process.*, vol. 68, pp. 747–763, 2019.
- [12] H. Naaman, S. Mulleti, and Y. C. Eldar, "FRI-TEM: Time encoding sampling of finite-rate-of-innovation signals," *IEEE Trans. Signal Process.*, vol. 70, pp. 2267–2279, 2022.
- [13] N. Shlezinger, Y. C. Eldar, and M. R. Rodrigues, "Hardware-limited task-based quantization," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5223–5238, Aug. 2019.
- [14] S. Salamatian, N. Shlezinger, Y. C. Eldar, and M. Médard, "Task-based quantization for recovering quadratic functions using principal inertia components," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 390–394.
- [15] K. V. Mishra, Y. C. Eldar, E. Shoshan, M. Namer, and M. Meltsin, "A cognitive sub-Nyquist MIMO radar prototype," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 937–955, 2020.
- [16] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," *IET Circ. Devices & Sys.*, vol. 5, no. 1, pp. 8–20, 2011.
- [17] M. Mishali and Y. C. Eldar, "Sub-Nyquist sampling: Bridging theory and practice," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 98–124, 2011.
- [18] A. Bhandari, "Back in the US-SR: Unlimited sampling and sparse super-resolution with its hardware validation," *IEEE Signal Process. Lett.*, vol. 29, pp. 1047–1051, 2022.
- [19] S. Mulleti, E. Reznitskiy, S. Savariego, M. Namer, N. Glazer, and Y. C. Eldar, "A hardware prototype of wideband high-dynamic range ADC," (to appear) *IET Circuits Devices Syst.*, 2023.
- [20] H. Naaman, N. Glazer, M. Namer, D. Bilik, S. Savariego, and Y. C. Eldar, "Hardware prototype of a time-encoding sub-Nyquist ADC," *arXiv preprint arXiv:2301.02012*, 2023.
- [21] T. Gong, N. Shlezinger, S. S. Ioushua, M. Namer, Z. Yang, and Y. C. Eldar, "RF chain reduction for MIMO systems: A hardware prototype," *IEEE Sys. J.*, vol. 14, no. 4, pp. 5296–5307, 2020.
- [22] R. T. Yazicigil, T. Haque, P. R. Kinget, and J. Wright, "Taking compressive sensing to the hardware level: Breaking fundamental radio-frequency hardware performance tradeoffs," *IEEE Signal Process. Mag.*, vol. 36, no. 2, pp. 81–100, 2019.
- [23] B. Murmann, "ADC Performance Survey 1997-2022," [Online]. Available: <https://github.com/bmurmann/ADC-survey>.
- [24] M. Mishali and Y. C. Eldar, "Wideband spectrum sensing at sub-Nyquist rates," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 102–135, 2011.
- [25] Y. C. Eldar and M. Mishali, "Beyond bandlimited sampling," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 48–68, 2009.
- [26] T. Zirtiloglu, N. Shlezinger, Y. C. Eldar, and R. Tugce Yazicigil, "Power-efficient hybrid MIMO receiver with task-specific beamforming using low-resolution ADCs," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, 2022, pp. 5338–5342.
- [27] 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.
- [28] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1741–1757, May 2007.
- [29] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 31–40, Mar. 2008.
- [30] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1827–1842, Apr. 2011.
- [31] S. Mulleti and C. S. Seelamantula, "Paley-Wiener characterization of kernels for finite-rate-of-innovation sampling," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5860–5872, Nov. 2017.
- [32] Y.-P. Lin and P. Vaidyanathan, "Periodically nonuniform sampling of bandpass signals," *IEEE Trans. Circuits Syst. II*, vol. 45, no. 3, pp. 340–351, 1998.
- [33] C. Herley and P. W. Wong, "Minimum rate sampling and reconstruction of signals with arbitrary frequency support," *IEEE Trans. Info. Theory*, vol. 45, no. 5, pp. 1555–1564, 1999.
- [34] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.
- [35] —, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. Sel. Topics in Signal Process.*, vol. 4, no. 2, pp. 375–391, Apr. 2010.
- [36] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice Hall, 1997.
- [37] T. Blu, P. Thevenaz, and M. Unser, "MOMS: Maximal-order interpolation of minimal support," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1069–1080, Jul. 2001.
- [38] H. J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, no. 1, pp. 37–52, 1967.
- [39] R. T. Yazicigil and et al., "Wideband rapid interferer detector exploiting compressed sampling with a quadrature analog-to-information

- converter," *IEEE J. Solid-State Circuits*, vol. 50, no. 12, pp. 3047–3064, Dec. 2015.
- [40] —, "A compressed-sampling time-segmented quadrature analog-to-information converter for wideband rapid detection of up to 6 interferers with adaptive thresholding," in *IEEE Radio Frequency Integrated Circuits Symposium*, May 2016, pp. 282–285.
 - [41] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.
 - [42] T. Haque and et al., "Theory and design of a quadrature analog-to-information converter for energy-efficient wideband spectrum sensing," *IEEE Trans. Circuits Syst. I*, vol. 62, no. 2, pp. 527–535, Feb. 2015.
 - [43] M. Saberi, R. Lotfi, K. Mafinezhad, and W. A. Serdijn, "Analysis of power consumption and linearity in capacitive digital-to-analog converters used in successive approximation adcs," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 8, pp. 1736–1748, 2011.
 - [44] W. Tang, A. Osman, D. Kim, B. Goldstein, C. Huang, B. Martini, V. A. Pieribone, and E. Culurciello, "Continuous time level crossing sampling adc for bio-potential recording systems," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 6, pp. 1407–1418, 2013.
 - [45] J. Szyduczyński, D. Kościelnik, and M. Miśkiewicz, "Time-to-digital conversion techniques: A survey of recent developments," *Measurement*, p. 112762, 2023.
 - [46] R. van de Plassche, *Sample-and-hold amplifiers*. Boston, MA: Springer US, 2003, pp. 313–347. [Online]. Available: https://doi.org/10.1007/978-1-4757-3768-4_7
 - [47] Z. Wang, H. Zhang, Y. Zhang, L. Shen, J. Ru, H. Fan, Z. Tan, Y. Wang, L. Ye, and R. Huang, "A software-defined always-on system with 57–75-nw wake-up function using asynchronous clock-free pipelined event-driven architecture and time-shielding level-crossing adc," *IEEE Journal of Solid-State Circuits*, vol. 56, no. 9, pp. 2804–2816, 2021.
 - [48] N. Shlezinger and Y. C. Eldar, "Deep task-based quantization," *Entropy*, vol. 23, no. 104, pp. 1–18, Jan. 2021.
 - [49] N. Shlezinger, Y. C. Eldar, and M. R. D. Rodrigues, "Asymptotic task-based quantization with application to massive MIMO," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 3995–4012, Aug. 2019.
 - [50] T. Zirtiloglu, N. Shlezinger, Y. C. Eldar, and R. Tugce Yazicigil, "Power-efficient hybrid mimo receiver with task-specific beamforming using low-resolution adcs," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5338–5342.