Article

Extrema-Triggered Conversion for Non-Stationary Signal Acquisition in Wireless Sensor Nodes

Swagat Bhattacharyya and Jennifer O. Hasler *

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA; sbhattac8@gatech.edu
* Correspondence: jennifer.hasler@ece.gatech.edu

Abstract: While wireless sensor node (WSNs) have proliferated with the rise of the Internet of Things (IoT), uniformly sampled analog–digital converters (ADCs) have traditionally reigned paramount in the signal processing pipeline. The large volume of data generated by uniformly sampled ADCs while capturing most real-world signals, which are highly non-stationary and sparse in information content, considerably strains the power budget of WSNs during data transmission. Given the pressing need for intelligent sampling, this work proposes an extrema pulse generator devised to trigger ADCs at significant signal extrema, thereby curbing the volume of data points collected and transmitted, and mitigating transmission power draw. After providing a comprehensive signal-theoretic rationale, we construct and experimentally validate these circuits on a system-on-chip field-programmable analog array in a 350 nm complementary metal-oxide-semiconductor (MOS) process. Operating within a power range of 4.3–12.3 µW (contingent on the input bandwidth requirements), the extrema pulse generator has proven to be capable of effectively sampling both synthetic and natural signals, achieving significant reductions in data volume and signal reconstruction error. Using a nonideality-resilient reconstruction algorithm, that we develop in this work, experimental comparisons between extrema and uniform sampling show that extrema sampling achieves an 18-fold lower normalized root mean square reconstruction error for a quadratic chirp signal, despite requiring 5-fold fewer sample points. Similar improvements in both the reconstruction error and effective sampling rate objectives are found experimentally for an electrocardiogram signal. Using both theoretical and experimental methods, this work demonstrates the potential of extrema-triggered systems for extending Pareto frontiers in modern, resource-constrained sensing scenarios.

Keywords: analog–digital conversion; asynchronous; event detection; extrema; field-programmable analog array; non-stationary signal; nonuniform; reconfigurable; reconstruction; sampling; wireless sensor node

1. The Need for Intelligent Sampling Approaches

The expansion of the Internet of Things (IoT) has brought forth an explosion in the number of wireless sensor node (WSN) applications; the WSN market [1] is currently growing by 15.5% annually. The primary objective of a WSN (Figure 1a) is to efficiently encode observations of physical phenomena into digital symbols for wireless transmission to a base station. Nearly all application scenarios of WSNs are resource-constrained, and WSN designers must intelligently balance power usage, data integrity, and system adaptability, which are, in turn, dependent on both the input signal and the signal processing pipeline. Biomedical applications, such as wearables and implantable devices [2–5], are particularly challenging since they have stringent upper bounds on device size and power consumption.
To obtain approximate design constraints for implantables, we consider intracortical neural recorder arrays. The volume of packaged arrays must be a few mL or less. For typical die areas (roughly $30 \text{ mm}^2$), intracortical neural recorder arrays have a power limit of roughly $10 \text{ mW}$ to prevent damage to brain tissue [6]. The low-noise amplifiers (LNAs) required in the analog front end consume 20–30% of the total power in a typical neural recorder power budget, while wireless transmission and the analog–digital converters (ADCs) have a combined 40–50% contribution to the total power [6,7].

Modern ADCs perform at efficiencies exceeding 1 $\mu \text{W} / \text{Mbps}$ [8], yet simply connecting the ADC output to a standard digital pad, which can present up to 100 pF capacitance (at $V_{DD} = 1 \text{ V}$), would require 50 $\mu \text{W} / \text{Mbps}$. In fact, the power consumption of wireless transmission and ADCs can range from 100 $\mu \text{W} / \text{Mbps}$ for backscatter communication to 10 $\text{mW} / \text{Mbps}$ for short-range frequency-shift keying (FSK) transmission [6,7]. Bluetooth low energy 5 (BLE 5) and similar commercial standards need roughly 50 $\text{mW} / \text{Mbps}$ at 8 dBm transmit power [9].

Improvements in LNA power consumption and data transmission cost are fundamentally bottlenecked by gain and noise requirements [6]. Nevertheless, not all sensor data are relevant, and one can greatly benefit from reducing the transmitted data. However, data reduction approaches must not cause the excessive loss of information or require excessive resource overhead. Biological signals, epitomized by electrocardiograms (ECGs) (Figure 2a), are often non-stationary [10–12]. To accurately capture the rapid transients of the QRS complexes (marked in Figure 2a), which usually persist for less than 20% of the inter-heartbeat period, sampling rates of at least 250 Hz and resolutions of at least 8 bit are essential [13,14]. This non-stationary characteristic opens up the possibility of utilizing nonuniform sampling to significantly reduce the number of transmitted data points at the source without requiring complex data compression approaches.

The relevance of nonuniform sampling stretches beyond its conventional applications in sensing into physical computation applications. Solutions derived from biologically inspired ordinary differential equations (ODEs) frequently exhibit non-stationary signals. Due to the analogous behavior of metal-oxide-semiconductor field-effect transistors (MOSFETs) and biological channels, these biologically inspired ODEs can often be computed physically with less power than is needed for sampling the solution at the necessary speed...
and precision [15]. Consequently, the same incentives for the nonuniform sampling of physiological signals are applicable to the digital readout of ODE solutions on analog hardware accelerators.

Figure 2. Illustration of an ECG (a) waveform and its wideband spectrogram showing the need for high resolution and high sampling rates during uniform sampling. Pareto fronts for uniform and nonuniform ECG sampling as determined by (b) an NRMSE minimization and (c) an AICc minimization problem. Both naive nonuniform sampling approaches show a set of special points, which includes extrema, that allow the corresponding nonuniform method to reconstruct ECGs using fewer sample points and achieve a lower reconstruction error (NRMSE) than uniform sampling.

This work, which builds on [16], proposes an extrema pulse generator which is capable of triggering an ADC (like an asynchronous successive approximation register (SAR)) at significant extrema and a timer to capture the corresponding timestamps (Figure 1b). While extrema sampling is a versatile, theoretically justified nonuniform sampling method, its practical use requires the development of more robust hardware and software than those currently available. In contrast to the limited number of previous hardware approaches [17,18] that have demonstrated low-power extrema detection circuits and signal reconstruction from extrema points, our work presents the following enhancements and technical contributions:

1. A comprehensive discussion justifying extrema sampling using signal-theoretic principles and two naive nonuniform sampling approaches.
2. A novel extrema pulse generator circuit design that is readily adaptable for scenarios with differing operating frequencies, power budgets, and signal-to-noise ratios (SNRs).
3. A reconstruction algorithm that is resilient to circuit nonideality and allows for a more relaxed set of assumptions about the interpolation function.

4. Experimental verification of the Pareto optimality of extrema sampling over uniform sampling for two test signals.

We construct the extrema pulse generator and experimentally demonstrate its performance using an system-on-chip (SoC) field-programmable analog array (FPAA) previously fabricated in a 350 nm complementary MOS process at Georgia Institute of Technology [19]. The rest of this work is structured as follows: Section 2 justifies the reasons for using extrema sampling as opposed to other nonuniform sampling approaches, Section 3 provides a brief overview of the SoC FPAA infrastructure, Section 4 presents an in-depth analysis of the subcircuits composing the extrema pulse generator, Section 5 explains the signal reconstruction process, Section 6 offers a comparative analysis of the system’s performance against other sampling methods, and Section 7 offers concluding remarks.

2. Nonuniform Sampling Approaches

A primary reason for the high energy efficiency of the human sensory nervous system is that it is only sensitive to novel events [20]. Nonuniform sampling approaches typically employ similar event-driven strategies to improve energy efficiency and reduce the need for energy-intensive compression algorithms and the data transmission cost in WSNs. Most nonuniform sampling approaches [21–27] leverage extra assumptions or information about signal features (besides the spectral support range) to sample more intelligently. Each nonuniform sampling approach comes with its own set of challenges, tradeoffs, and assumptions. As a result, nonuniform sampling approaches are often highly application-specific, posing a barrier to the commercial availability of nonuniform ADCs. Even in academic settings (e.g., clinical studies), a priori information about the signal features of interest can be limited, so a feature-specific sampling approach may cause the undesired loss of information. It is illuminating to analyze contemporary nonuniform sampling approaches before discussing extrema sampling.

Application-specific event-detectors usually wake up a microcontroller to sample an ADC when an interesting event is detected (such as acoustic spectra from a vehicle [27]). This approach performs well when interesting events are rare, since high-power components can be kept in sleep states for prolonged periods while an always-on, low-power classifier monitors for interesting events. Yet, as interesting events become more frequent, the benefits gleaned by application-specific event-detectors become less pronounced [27] and are eventually overcome by the additional hardware overhead. Low-power event classifiers are often heavily limited in scope and require meticulous retraining for a different event class.

Level-crossing ADCs [22], which generalize the principle of event detection, sample the input signal when it changes by some multiple of a least sensitive bit. Level-crossing ADCs map the problem of precision voltage measurement at uniform time steps to the often easier problem of precision time measurement as the input crosses (usually uniformly spaced) reference voltages. Level-crossing ADCs are expected to scale well since time–digital converter precision and energy efficiency improve with decreasing technology node [28]. Although there have been efforts to mitigate this [25,26], level-crossing ADCs tend to oversample many classes of signals, especially if good voltage resolution is required.

If it is known a priori that the signal to be sampled has a sparse representation in some transform domain, then the signal can be acquired by random sampling approaches widely known as compressive sampling. Compressive sampling has been immensely successful in the recovery of many natural signals far below their Nyquist rate. Yet, a major demerit of compressive sampling lies in signal recovery, where a convex optimization problem (typically L1 norm minimization) must be solved [23,24]. The best case overhead for L1 minimization [29] tends to be higher than that of Lagrange [30,31] or cubic [32] interpolation. Furthermore, the convergence time of the recovery algorithm in compressive sensing, which
is uncertain (as opposed to interpolation-based approaches), can be problematic for real-time or closed-loop applications.

To summarize, the events identified by a nonuniform sampling methodology range in complexity from simple input value changes, as in level-crossing ADCs, to complex, signal-specific features identified by techniques like spectral template matching [27], as in application-specific event detectors. While the former method offers higher sensitivity with a lower component count (albeit with lower specificity), the latter approach provides higher specificity, but with lower sensitivity and an increased number of components. In contrast, our proposed approach, extrema sampling, is a broadly applicable solution that reduces energy usage with both a low component count and low reconstruction error.

Indeed, extrema sampling, in contrast to the previously mentioned methods, relaxes signal-specific assumptions, offering a framework applicable to a wide array of scenarios. Sampling at twice the mean frequency of the input signal, extrema sampling often samples considerably below the global Nyquist rate for non-stationary signals (even after accounting for the two-fold penalty associated with also acquiring sample timestamps). Moreover, extrema sampling does not necessitate costly signal reconstruction algorithms and is grounded [21] in theory since

1. Signal quantities of interest are often simply the time between extrema and the extrema values; interpolation may not be required in such use cases [33].
2. Since extrema are the zero crossings of the derivative of the signal, extrema essentially carry double the information of uniformly sampled points.
3. Extrema appear in excess of half the Nyquist rate in band-limited signals. It then follows from the second rationale that there is enough information to reconstruct band-limited signals perfectly from its extrema samples via variants of Lagrange interpolation [21].

In addition to the established reasons described above, in this work, we also show that extrema sampling arises naturally from the solution of a few types of optimization problems. The first problem we propose is the selection of time-domain points so as to minimize the polynomial reconstruction error subject to a constraint on the mean sampling rate ($F_{star}$). We formulate this first problem as the following constrained nonlinear integer programming problem, which we solve using a genetic algorithm:

$$\min_{\Phi} \text{NRMSE} \quad \text{s.t.} \quad \frac{k}{n} \leq F_{star}, \quad \forall \Phi_i \in \{0, 1\}, \quad \text{where} \quad \text{NRMSE} := \frac{\|F(X, \Phi) - X\|_2}{\|X - \langle X \rangle\|_2}. \quad (1)$$

Normalized root-mean-square error (NRMSE) is an estimate of the error between the original ECG data vector ($X$) and the reconstructed data vector ($F$), guided by a vector ($\Phi$) of length $n$ containing binary elements which decide which of the $n$ elements of $X$ to sample. It can be shown that the definition of NRMSE in this work, which was proposed previously in [15], is equivalent to the root-mean-square reconstruction error normalized by the standard deviation of the truth (original ECG data). While solving Equation (1), we ensure $X$ is highly oversampled so as to approximate a continuous-time signal, and we reconstruct $F$ using a piecewise cubic Hermite interpolating polynomial (PCHIP) function.

The Pareto fronts (please see [34] for a detailed discussion on Pareto fronts and optimality) obtained for uniform and nonuniform ECG sampling (as determined by solving Equation (1)) are shown in Figure 2b. As shown in Figure 2b, our naive nonuniform sampling approach has a tendency to prioritize significant signal extrema, which allows the nonuniform method to reconstruct ECGs using a fewer number of sample points and with a lower reconstruction error (NRMSE) than the uniform sampling approach (i.e., achieving a Pareto optimal tradeoff between NRMSE and the effective sampling rate $F_{eff}$ relative to uniform sampling). The optimality advantage of nonuniform sampling remains even if the extra overhead needed to acquire timestamps in the nonuniform case is considered.
Our second proposed problem is the selection of time-domain points so as to minimize a model-selective, information-theoretic criterion subject to a constraint on the mean sampling rate \(F_{\text{tar}}\). We formulate this second problem as the following constrained nonlinear integer programming problem, which is also solved using a genetic algorithm:

\[
\arg\min_{\Phi} \ AICc \quad \text{s.t.} \quad \frac{k}{n} \leq F_{\text{tar}} \quad \forall \Phi_i \in \{0, 1\}.
\]  

(2)

The number and selection of sample points on the time-domain waveform corresponds to the selection of the PCHIP model order and a choice of parameters. AICc denotes the Akaike Information Criterion corrected for small sample sizes, which is a widely used information-theoretic model selection criterion \([35,36]\) that rewards models with lower mean-squared error while appropriately penalizing models with a larger number of model parameters \((k)\) and adding a correction term for small sample sizes. In our mathematical formalism, AICc can be written as follows \([36]\):

\[
AICc := n \ln \left( \frac{1}{n} \| F(X, \Phi) - X \|_2^2 \right) + \frac{2kn}{n - k - 1}.
\]  

(3)

The optimization results of Equation (2) are shown in Figure 2c. We see similar results in Figure 2b, where the AICc minimization demonstrates that an intelligent choice of time-domain sample points leads to a precise, yet more parsimonious, representation of a signal than a uniformly sampled set of points. Additionally, like in Figure 2b, the optimal sample points also predominantly include extrema.

The efficacy of signal reconstruction from extrema samples is contingent on a proper selection of reconstruction basis. The discussion in the preceding paragraph suggests that, for certain smooth signals like ECGs, PCHIPs are a good choice of basis function because PCHIPs enforce continuity and smoothness conditions. However, for signals with abrupt changes or corners, such as a sawtooth wave, a different basis function (e.g., a linear interpolation method) could be more effective, offering accurate reconstructions without the need for complex algorithms. It should be noted that signals that are contaminated with high-frequency noise beyond signal frequencies should be prefiltered before extrema sampling to mitigate false positives. False positives do not constitute data loss and thus would not degrade reconstruction quality given an a posteriori assessment of which data points correspond to false positives (e.g., using a method we show in Section 5). However, the occurrence of false positives would diminish the data savings gained through extrema sampling, thereby increasing the mean effective sampling rate to a value that is closer to the Nyquist rate.

3. SoC FPAA Infrastructure

To construct our circuits, we use an SoC field-programmable analog array FPAA developed at Georgia Institute of Technology. The SoC FPAA is a highly versatile general-purpose analog computing platform in a 350 nm process \([19]\). The FPAA uses software tools that are openly available at hasler.ece.gatech.edu/FPAAtool/ and has 98 fully reconfigurable computational analog blocks (CABS) that are interconnected with a programmable, nonvolatile routing fabric comprising floating-gate (FG) transistors (Figure 3).

Each CABS contains a wide assortment of analog computational elements: operational transconductance amplifiers (OTAs), floating-gate OTAs (FGOTAs), discrete transistors, capacitors, current mirrors, and T-gates, which can be interconnected to synthesize larger circuits. Synthesized circuits can be either contained within a CABS or span multiple CABS.

The routing fabric of the FPAA is used to make flexible nonvolatile connections or to generate nonvolatile subcircuit current biases with 13-bit precision \([37]\) using roughly half-a-million FG transistors. An FG transistor is a MOSFET with only capacitors tied to the gate. In this way, the gate of a FG is floating, with no DC connection to any other node, which allows charge to be trapped on the gate. One of these gate capacitors
(typically a MOS capacitor) is used to remove electrons from the FG through Fowler–Nordheim tunneling across its insulator. The other capacitors, which typically have a higher coupling factor to the FG than the MOS capacitor, are control gates. The control gates are particularly useful during hot electron injection, a process by which electrons can be injected into the FG through the gate oxide of the FG transistor. FG transistors allow circuit designers a great level of control over transistor IV curves and can be used to mitigate manufacturing mismatch.

Figure 3. Block diagram and programming flow of the 350 nm SoC FPAA, which contains 98 fully reconfigurable CABs interconnected with FG routing fabric and programmed via open-source tools.

4. Extrema Pulse Generator

Comprising two subcircuits, the extrema detector and the edge detector, our proposed low-power extrema pulse generator is shown in Figure 4. The objective of the extrema detector is to change its output state at the input extrema. The edge detector then produces an active-low pulse given a state change on the extrema detector output. The hysteretic differentiator (HD) is pivotal to the overall extrema detector circuit; thus, it is imperative to first elucidate the HD operation.

Figure 4. Diagram of the extrema pulse generator. The capacitors drawn with gray lines are induced via routing parasitics.
4.1. Hysteretic Differentiator

In order to detect extrema, one must perform edge detection using some sort of differentiation operation. However, the noise immunity of linear differentiators is poor [20]. In a fundamental sense, differentiators are circuits whose outputs are insensitive to the absolute voltage level of the signal while remaining sensitive to the local signal derivative. The functionality of our extrema detector is contingent on the HD (Figure 5a), a nonlinear differentiator circuit [20] which is tolerant to noise.

![HD schematic diagram](image)

\[V_{in} \rightarrow G_{HD} \rightarrow V_{hd,1}, V_{hd,2} \rightarrow \text{Output}\]

**Figure 5.** HD (a) schematic diagram and (b) experimental measurements of \(V_{lpf}\) and \(V_{hd,2}\) in response to 500 Hz sinusoidal inputs with different signal amplitudes (increasing left to right).

In essence, the structure of the HD is analogous to a voltage follower comprising a nonlinear buffer stage that is driven by OTA \(G_{HD}\), which is biased via an FG pFET. The output of the HD (\(V_{hd}\)) corresponds to the output of \(G_{HD}\). For a small HD input (\(V_c\)), the output swing of the buffer stage is also small, and the buffer can be approximated as a linear system; thus, \(V_{hd}\) tracks \(V_c\) closely given a small \(V_c\). In contrast, \(V_{hd}\) is sensitive to \(\text{sgn}(\partial V_c/\partial t)\) when \(V_c\) is large, transitioning sharply at signal extrema since the dominant FET swaps; given the unconventional positioning of the transistors (pFET drain at ground and nFET drain at Vdd), the swap of the dominant FET requires a large change in the common gate control voltage to source or sink an appreciable amount of current. Note that the pFET is dominant for decreasing \(V_c\), and the nFET is dominant for increasing \(V_c\).

4.2. Extrema Detector Circuit

While the HD can generate sharp transitions upon observing extrema, the output transients of an HD are slow if a large neighborhood surrounding the extrema is flat, and the swing of an HD is not rail-to-rail. Therefore, a single HD cannot fulfill the key objective of the extrema detector: to produce a digital output (\(V_{comp}\)) that flips its state when it observes a significant \(V_{in}\) extrema. Therefore, to produce the extrema detector, we cascade two HDs (labeled HD1 and HD2 in Figure 4). This approach allows HD1 to sharpen \(V_{in}\) extrema so that HD2 can have faster transients. Therefore, since the input of HD2 is already sharp, HD1 dominates latency; thus, the use of two HDs does not appreciably increase the overall system power draw. We then compare the output of HD2 to the output of HD1 to generate the digital classification \(V_{comp}\).

Directly cascading HDs leads to nonidealities, as each HD contributes output noise, has an input feedthrough component, and overall gain. In order to address these nonidealities, we introduce other subcircuits into the system: the noise filter, scaler, integrator, and Schmitt trigger. It is undesirable to pass the output noise of HD1 far beyond input signal frequencies into HD2, since this high-frequency noise would be amplified. To this end, we use a low-pass OTA-capacitor noise filter to mitigate high-frequency noise from HD1.

The input offset and input swing of HD2 can be transformed via the voltage bias \(V_{TRIM,1}\) and resistors \(R_H\) and \(R_L\) present in the scaler circuit, respectively. \(V_{TRIM,1}, R_H,\) and \(R_L\) are intelligently chosen to satisfy two criteria:
1. For typical $V_{in}$ swings, the input voltage swing of HD2 does not saturate $V_{hd2}$ near the supply rails.

2. The HD1 output offset is larger than the HD2 output offset by the HD2 output noise swing. This condition mitigates spurious comparisons.

To better illustrate the circuit operation, we show the response of the cascaded HDs to sinusoids of successively larger amplitudes in Figure 5b.

We cascade an integrator with a Schmitt trigger to produce a noise-immune comparator. We tune the time constant of the integrator so the switching period of the comparator is well below the period of $V_{in}$ but higher than the period of undesired noise components. As shown in Figure 6a, our Schmitt trigger is composed of two current-starved inverters arranged in a topology inspired by [38]. The second current-starved inverter in the cascade uses an FG pFET bias to limit the short-circuit current. In the first starved inverter, the bias is directly set by the input pFET, which itself is an FG transistor. In fact, the FG pFET in the first inverter has two control gates, where the first control gate ($C_{IN}$) corresponds to the Schmitt trigger input ($V_{int}$), and the second control gate ($C_{FB}$) is tied to the Schmitt trigger output ($V_{comp}$) so as to create a positive feedback loop.

![Figure 6. Schmitt trigger (a) schematic diagram and (b) hysteresis curve measured from an SoC FPAA implementation. (c) Measured output of the extrema detector given sinusoidal inputs of increasing frequencies (increasing left to right). The capacitors with gray lines are induced via routing parasitics.](image)

The hysteresis curve of the Schmitt trigger is mostly shaped by the parameters of the two-input FG pFET in the first inverter. Mainly, the ratio of $C_{FB}$ to the total capacitance on the FG node ($C_{T}$) controls the spacing between the low–high ($V_{TH}$) and high–low ($V_{TL}$) transitions, while the charge trapped on the FG ($Q_{FG}$) sets the low–high output transition level ($V_{TH}$). We can derive compact (approximate) expressions for $V_{TH}$ and $V_{TL}$ by solving for the transition voltage of the first inverter in the cascade (i.e., the DC point where the output of the first inverter $V_p = V_{int}$) using the square-law models for MOSFETs operating in above-threshold saturation [38]:

$$V_{TH} \approx \frac{C_T \cdot V_{DD}}{C_T + C_{IN}} - \frac{Q_{FG}}{C_T + C_{IN}}; \text{ where, } C_T \approx C_{IN} + C_{FB} \text{ and}$$

$$V_{TH} - V_{TL} \approx \left(\frac{C_{FB} \cdot V_{DD}}{C_T + C_{IN}}\right).$$

In our derivation of Equation (5), we have assumed $\mu_p (W/L)_p \approx \mu_n (W/L)_n$, which was ensured during the design of the SoC FPAA, and that $C_{IN}$ and $C_{FB}$ are large compared to any other miscellaneous capacitances on the floating gate node. We tune the Schmitt trigger to symmetrize the hysteresis curve around the mean value of $V_{int}$, resulting in the response shown in Figure 6b. After biasing, we observe that the extrema detector response has some latency, which can be characterized by using sinusoidal test inputs with increasing frequencies, as shown in Figure 6c. We find that the latency has a component that is invariant to the input and a component that scales proportionately to the input signal period.
4.3. Edge Detector Circuit

Our edge detector circuit, shown in Figure 7a, generates a negative pulse when its input $V_{\text{comp}}$ has a falling or rising edge, as shown in Figure 7b. An OTA integrator and a current-starved inverter are used within the edge detector to produce a delayed and inverted copy of the input ($V_d$). $V_d$ is then compared to $V_{\text{comp}}$ using an ‘XOR’ operation to generate the edge detector output $V_{\text{event}}$.

We induce integration capacitance $C_{P,3}$ using routing parasitics. The clock pulse width is tuned using the integrator bias $G_{\text{TIM}}$. We then trim any mismatch between the OTA slew rates on the negative and positive edges by setting the common mode voltage of the OTA in the integrator using its reference $V_{\text{TRIM}}$. Our tuning approach results in a symmetric clock pulse width (20 µs here) on maxima and minima.

![Figure 7. Edge detector (a) schematic diagram and (b) experimental wavefronts.](image)

5. Reconstruction Algorithm

In the context of this work, extrema sampling corresponds to the sampling of $V_{\text{in}}$ and the recording of the corresponding timestamps on each falling edge of $V_{\text{event}}$ using an 8-bit oscilloscope. The reconstruction of the input signal from extrema samples is performed through two successive processes: (1) sample extrapolation and (2) polynomial interpolation. Algorithm 1 and Figure 8 show the details of the extrapolation algorithm, which first infers if a sample corresponds to a local extremum from the values of the surrounding points. At each identified extrema, the extrapolation algorithm compensates sample timestamps by estimating the extrema pulse generator latency. Latency is estimated via a linear model since, as mentioned previously, the latency has an input-invariant component and a component that grows proportionately to the period of the input signal.

The latency model depends on process–voltage–temperature (PVT) conditions and is estimated through sinusoidal input test signals. During sample extrapolation, the input period is estimated from a local sinusoidal assumption. To summarize, given an extremum, the following sequence of events occurs:

1. Timestamps of the two neighboring points to the sample are leveraged for the estimation of the local period of the signal.
2. A linear model is used to estimate the delay from the true extrema location to the clock pulse produced by the extrema pulse generator.
3. Extrema voltage values are estimated from the sampled voltage using a parabolic approximation of the waveform and the delay estimate; the form of the parabolic approximation is found by computing a Taylor series expansion of the local sinusoid.

We reconstruct the input from the extrapolated sample points using polynomial interpolation. While theoretically ideal [21,30], Lagrange interpolation variants can be unreliable if nonidealities, such as a slight misalignment of the extrapolated extrema points with the true extrema locations or a few false negatives/positives from the extrema pulse generator are present. This intolerance of Lagrange interpolation to nonidealities makes polynomial interpolation approaches often more preferable in practice. Previous studies used a family of Bézier curves with concavity restrictions [17,18]; however, the concavity
assumptions implicate that this family of functions would perform poorly with certain signals (e.g., triangle/sawtooth waves). We use PCHIPs [32] in this work because PCHIPs make much more general assumptions about the characteristics of the input signal and are much more well-behaved given nonidealities in timing or false positives/negatives. PCHIPs also have minimal overshoot.

Algorithm 1: Extrapolation of samples \((t_i, x_i), \forall i \leq n\)

Data: \(n \geq 2, t_i, x_i\)
Result: \(t_{ci}, x_{ci} = x_i + \sigma_i\Delta_i\)

\[
\begin{align*}
\text{i } & \leftarrow 2; \\
\text{while } i \leq n \text{ do} \\
& a_{i-1} \leftarrow (|x_{i-2} - x_{i-1}| + |x_{i} - x_{i-1}|)/2; \\
& f_{i-1} \leftarrow (1/(t_{i-1} - t_{i-2}) + 1/(t_{i} - t_{i-1}))/2; \\
& \text{if } \text{sgn}(x_{i-1} - x_{i-2}) = -\text{sgn}(x_{i} - x_{i-1}) \neq 0 \text{ then} \\
& \quad \sigma_{i-1} \leftarrow \text{sgn}(x_{i-1} - x_{i-2}); \\
& \quad \tau_{i-1} \leftarrow \alpha_{i-1}\sigma_{i-1} + \beta_{i-1}\sigma_{i-1}/f_{i-1}; \\
& \text{else} \\
& \quad \sigma_{i-1}, \tau_{i-1} \leftarrow 0; \\
& \text{end} \\
& t_{ci-1} \leftarrow t_{i-1} - \tau_{i-1}, \Delta_{i-1} \leftarrow \pi a_{i-1}f_{i-1}\tau_{i-1}^2; \\
& i \leftarrow i + 1; \\
\end{align*}
\]

Figure 8. Annotated diagram depicting the sample extrapolation method detailed in Algorithm 1 and relevant indexed quantities.

6. Results and Discussion

As shown in Figure 4, with the exception of the scaler circuit, we construct all other extrema pulse generator circuits on the 350 nm SoC FPAA. During characterization, we supply dynamic voltages from a function generator (Digilent Analog Discovery 2), supply static voltages from a power supply (Agilent E3620A), and acquire extrema samples from an 8-bit oscilloscope (Tektronix TDS5034B) on the falling edge of \(V_{\text{event}}\). An annotated photograph of our experimental test setup is shown in Figure 9. We optimize and demonstrate our circuit first for a quadratic chirp and then for an ECG signal. The majority of the power
draw of the extrema pulse generator stems from use of OTA circuits. The OTA-capacitor inter-
actions on each OTA output node determine the bandwidth of the overall extrema pulse
generator. Since the OTAs are biased in subthreshold saturation, the transconductance of
each OTA scales proportionally to its bias current. Consequently, the overall power scales
proportionally to the input bandwidth. In this work, the power-bandwidth scaling factor is
100 nW/Hz, which translates to a 12.3 μW draw for the quadratic chirp and a 4.3 μW draw
for the ECG.

From the reconstructed waveforms in Figure 10a,b, we visually observe that the
ECG and the quadratic chirp can be reconstructed quite well. Numerically, the NRMSE
corresponding to the reconstructions shown in Figure 10a,b are 0.044 (quadratic chirp)
and 0.261 (ECG). These NRMSEs are substantially less than the NRMSEs observed if these
same signals are uniformly sampled at the same mean rate. Typically, uniform sampling
would need to sample a couple of times faster than our proposed nonuniform sampling
approach in order to achieve a similar reconstruction error. In a quadratic chirp, the average
signal frequency ($F_{\text{avg}}$) is less than a quarter of the global Nyquist rate of the signal. We
therefore experimentally find that extrema sampling, which samples at a rate of roughly
$2F_{\text{avg}}$, results in an effective sampling rate of roughly $F_{\text{Nyquist}}/2.5$; even after accounting for
the sampling of timestamps, extrema sampling reduces the amount of data recorded during
the acquisition of a quadratic chirp while achieving an 18-fold lower reconstruction error
than uniform sampling. If uniform sampling is used, the ECG signal must be sampled three
times faster on average in order to obtain the NRMSE as the extrema sampling approach.
Furthermore, the NRMSE in uniform sampling is four-fold higher given the same $F_{\text{eff}}$
as the extrema sampling approach. In both ECG and quadratic chirp sampling, extrema
sampling remains a Pareto improvement over uniform sampling, which means that it is
an improvement in both NRMSE and $F_{\text{eff}}$ compared to the Pareto front of the uniform
sampling approach (even after accounting for any additional overhead associated with
timestamp acquisition). Nevertheless, implementation nonidealities (mostly false positives),
mean that solutions lying on the ideal Pareto front in Figure 2c cannot be attained with our
extrema pulse generator.

![Figure 9. Experimental setup for individually characterizing extrema pulse generator circuits and
demonstrating extrema sampling with the full system.](image-url)
Figure 10. Experimental results of extrema pulse generator sampling and reconstruction for (a) a quadratic chip and (b) an ECG. The ECG is filtered with a 60 Hz notch before input into the extrema pulse generator. A comparison between uniform sampling and extrema sampling performance for (c) the quadratic chirp and (d) the ECG. The inner points corresponding to the uniform samples form a Pareto front for the uniform approach; extrema sampling is a Pareto improvement.

We compare the performance of previous nonuniform sampling methodologies; specifically, audio-range FPAA or application-specific integrated circuit (ASIC) approaches in similar technology nodes are compared with our extrema pulse generator in Table 1. Notably, reference [39] is a recent demonstration of extrema sampling for a task other than data rate reduction [39], the authors use maxima sampling for envelope signal estimation in a resource-constrained voltage-controlled oscillator (VCO) with automatic gain control. We find that the proposed extrema pulse generator is more energy-efficient than other nonuniform sampling methodologies on FPAs [17,18,27], drawing less power for the same input bandwidth. A better performance can be attained on ASICs [25] since routing parasitics can be made lower than an FPAA implementation.

As demonstrated, PCHIP reconstruction works well; however, our PCHIPs do not leverage information pertaining to the classification of the sample points (maximum, minimum, or false positive) explicitly. In subsequent discussions, the performance can be improved further by extracting information from sample point classification. The extrema pulse generator also has a tradeoff between noise immunity and output delay, which should be decoupled in future architectural designs. Nevertheless, the great potential of extrema sampling for data reduction in ADCs is clearly demonstrated by the results of this work.
Table 1. Comparison of nonuniform sampling approaches.

<table>
<thead>
<tr>
<th>Application</th>
<th>Proposed</th>
<th>[17,18]</th>
<th>[27]</th>
<th>[39]</th>
<th>[25]</th>
<th>[40]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>FPAA</td>
<td>FPAA</td>
<td>FPAA</td>
<td>FPAA</td>
<td>ASIC</td>
<td>ASIC</td>
</tr>
<tr>
<td>Process (nm)</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>130</td>
<td>180</td>
</tr>
<tr>
<td>Bandwidth (Hz)</td>
<td>60, 1000</td>
<td>60</td>
<td>1000</td>
<td>30–40,000</td>
<td>4000</td>
<td>250</td>
</tr>
<tr>
<td>Power (µW)</td>
<td>4.3, 12.3</td>
<td>4.95</td>
<td>43</td>
<td>46–50</td>
<td>6.5</td>
<td>109</td>
</tr>
</tbody>
</table>

7. Conclusions

This work presented theoretical and experimental evidence in support of extrema sampling as a conversion paradigm in ADCs. Our hardware implementation featured a noise-robust extrema pulse generator to nonuniformly trigger an ADC at the extrema of the input signal. In a commercial implementation, the conversion clock for the ADC can be toggled between the extrema pulse generator and a uniform clock source. We also presented an algorithm for reconstructing the input signal from extrema samples. Using our approaches, we show the system performance for two test signals: a quadratic chirp and an ECG. Compared to uniform sampling, our approach achieves a Pareto-optimal tradeoff between the reconstruction error and ADC power consumption for both the quadratic chirp and the ECG. Our results suggest extrema sampling is a strong candidate for the nonuniform acquisition of a wide class of non-stationary signals.

Author Contributions: Both S.B. and J.O.H. contributed to the circuit design and writing. S.B. was also responsible for the derivations and experimental measurements. All authors have read and agreed to the published version of the manuscript.

Funding: This material is based on work partially supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-2039655.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors thank Pranav O. Mathews for their helpful discussions.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

- ADC: Analog–Digital Converter
- AIc: Akaike Information Criterion corrected (for small sample sizes)
- ASIC: Application-Specific Integrated Circuit
- BLE: Bluetooth Low Energy
- CAB: Computational Analog Block
- ECG: Electrocardiogram
- FET: Field-Effect Transistor
- FG: Floating-Gate
- FGOTA: Floating-Gate Operational Transconductance Amplifier
- FPAA: Field-Programmable Analog Array
- FSK: Frequency-Shift Keying
- HD: Hysteretic Differentiator
- IoT: Internet of Things
- LNA: Low-Noise Amplifier
References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.