

# A 10.1-ENOB, 6.2-fJ/conv.-step, 500-MS/s, Ringamp-Based Pipelined-SAR ADC With Background Calibration and Dynamic Reference Regulation in 16-nm CMOS

Jorge Lagos<sup>1</sup>, Senior Member, IEEE, Nereo Markulić<sup>2</sup>, Member, IEEE, Benjamin Hershberg<sup>3</sup>, Member, IEEE, Davide Dermit<sup>4</sup>, Member, IEEE, Mithlesh Shrivastava<sup>5</sup>, Member, IEEE, Ewout Martens<sup>6</sup>, Member, IEEE, and Jan Craninckx<sup>7</sup>, Fellow, IEEE

**Abstract**—This work presents a single-channel, fully dynamic pipelined-SAR ADC with relaxed architectural tradeoffs thanks to the use of ring amplification and background calibration. It leverages a novel SAR quantizer and narrowband dither injection to achieve fast and comprehensive background calibration of DAC mismatch, interstage gain, and ring amplifier (ringamp) linearity and bias optimality. The ADC also includes an on-chip, wide-range, fully dynamic reference regulation system. Implemented in 16-nm CMOS, it consumes 3.3 mW at 500 MS/s (including regulation) and achieves 10.1 ENOB and 75.5-dB SFDR, resulting in Schreier and Walden figure-of-merit (FoM) values of 171.1 dB and 6.2 fJ/conv.-step, respectively.

**Index Terms**—A/D, ADC, asynchronous, background calibration, dither, fully dynamic, low-power, narrowband dither, pipelined SAR, reference buffer, reference regulation, ringamp, ring amplification, ring amplifier, SAR quantizer, single-channel.

## I. INTRODUCTION

THE pipelined SAR ADC is a well-established architecture for applications requiring medium-to-large resolutions and bandwidths in the hundreds-of-MHz range, as it combines the good power efficiency and scalability of SAR conversion with the speed benefits of pipelined converters [1].

In this architecture, the topology choices for both the SAR quantizers and the residue amplifiers play a determinant role in the achievable speed, linearity, and power efficiency, with well-defined tradeoffs. On the quantizer side, solutions like the multi-comparator quantizer [2]–[4] and multi-bit-per-cycle conversion [5]–[7] can increase speed in exchange

Manuscript received August 26, 2021; revised November 2, 2021 and November 27, 2021; accepted November 27, 2021. Date of publication January 11, 2022; date of current version March 28, 2022. This article was approved by Associate Editor Yusuke Oike. (Jorge Lagos and Nereo Markulić contributed equally to this work.) (Corresponding author: Jorge Lagos.)

Jorge Lagos, Nereo Markulić, Mithlesh Shrivastava, Ewout Martens, and Jan Craninckx are with IMEC, 3001 Leuven, Belgium (e-mail: jorge.lagosbenites@imec.be).

Benjamin Hershberg was with IMEC, 3001 Leuven, Belgium. He is now with Hershberg Consulting Inc., Portland, OR 97225 USA.

Davide Dermit was with IMEC, 3001 Leuven, Belgium. He is now with arQana Technologies, 3001 Leuven, Belgium.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JSSC.2021.3133829>.

Digital Object Identifier 10.1109/JSSC.2021.3133829

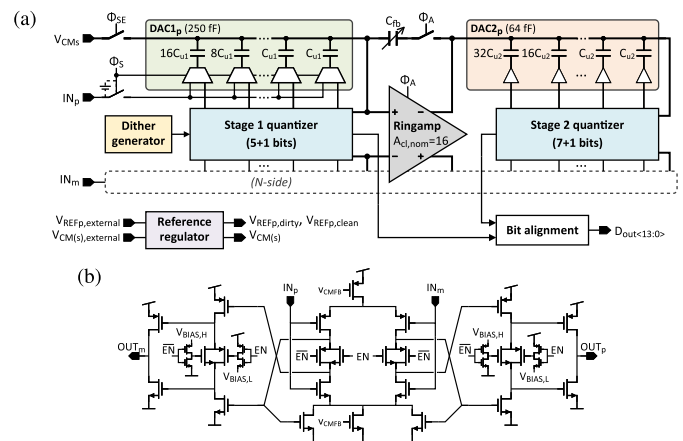


Fig. 1. (a) Proposed ADC architecture and (b) ringamp implementation [16].

for calibration and/or implementation complexity. On the residue amplifier side, the use of open-loop amplification [3], [8]–[10] has also been demonstrated as an effective means for enhancing speed and power efficiency, often at the cost of greater calibration complexity [10]–[12].

Ring amplification [13] is an alternative closed-loop amplification paradigm, and in recent years has broadly advanced the state-of-the-art in power efficiency for both pipelined [14]–[16] and pipelined-SAR ADCs [17]–[20]. Whereas recent ringamp-based pipelined ADCs rank at the forefront of the per-channel speed boundary, all previously reported ringamp-based pipelined SARs operate at speeds below the upper bound established by open-loop-amplifier-based designs [1]. This is to a large extent the result of two factors: target applications with high ENOB targets which in general lead to slower speeds regardless of the amplifier used, and a calibration-free approach where PVT robustness was ensured through extra design margin, at the expense of speed.

In this work, we present a pipelined-SAR ADC (see Fig. 1) that exploits ring amplification and background calibration to simultaneously achieve high speed and power efficiency [21]. Compared to open-loop-amplifier-based pipelined SARs, a ringamp helps to relax several key architectural

tradeoffs. First, due to their closed-loop nature, ringamps are unconstrained in terms of the residue amplitude size they can process at their input. As a result, both the LSB size in the front-end stage (Stage 1) and the ADC full-scale can be of any size. This allows for a relatively small number of bits to be quantized by the Stage-1 SAR, minimizing the critical Stage-1 conversion time and maximizing the error tolerance provided by, e.g., 1 LSB of interstage redundancy, due to a larger LSB size. Second, ringamps can process nearly rail-to-rail output swings while maintaining high linearity. As a result, the LSB size in the back-end stage (Stage 2) can also be maximized due to a large interstage gain factor, thus relaxing the noise, error, and metastability requirements in this stage and allowing the use of a fast, power-efficient quantizer.

Several background calibration loops are used in the proposed architecture to correct errors, ensure robustness, optimize performance, and maximize power efficiency. First-order interstage-gain calibration corrects for finite-gain error (higher-order components of gain error do not require correction due to the highly-linear output characteristic of the ringamp [22]). DAC mismatch calibration is used in Stage 1 to allow for a minimum-area DAC sized only for thermal noise requirements. Finally, a novel approach is introduced for optimizing the ringamp biasing and settling behavior in the background to guarantee robust PVT operation. All these calibration techniques are implemented using narrowband dither injection at the Stage-1 quantizer, leveraging a new SAR quantizer developed to support this capability.

The core pipeline in this work exploits fully-asynchronous, event-based operation [23] to achieve fully-dynamic power consumption. In order to extend this desirable property to the overall ADC, we also introduce a fully-dynamic reference buffer based on Split-Reference Regulation [15]. Especially tailored for pipelined-SAR operation, the regulator improves upon the implementation in [15] by using charge-pumping to synthesize a wider range of reference voltages, with increased power efficiency and lower area requirements.

This article is organized as follows. Section II discusses the SAR quantizer introduced in this work. Section III presents the details of the proposed narrowband-dither injection technique and the background calibration loops that utilize it. The fully-dynamic on-chip reference regulator is described in Section IV. Experimental results are reported in Section V. Finally, Section VI summarizes the conclusions of this work.

## II. MULTI-LATCH SAR QUANTIZER

### A. Review of Traditional SAR Quantizer Topologies

SAR quantizers suffer from an intrinsic speed-resolution tradeoff due to the serial nature of their operation. This sets an upper limit on their achievable speed, and as a result topologies with minimal critical-path decision delay are essential when high resolution and conversion speed are required.

Moreover, in pipelined and pipelined-SAR ADCs the ability to inject a pseudorandom noise (PN) dither signal into the analog path is often desired, as it enables the background calibration of several analog non-idealities, including DAC mismatch and interstage gain [24]. PN modulation of the

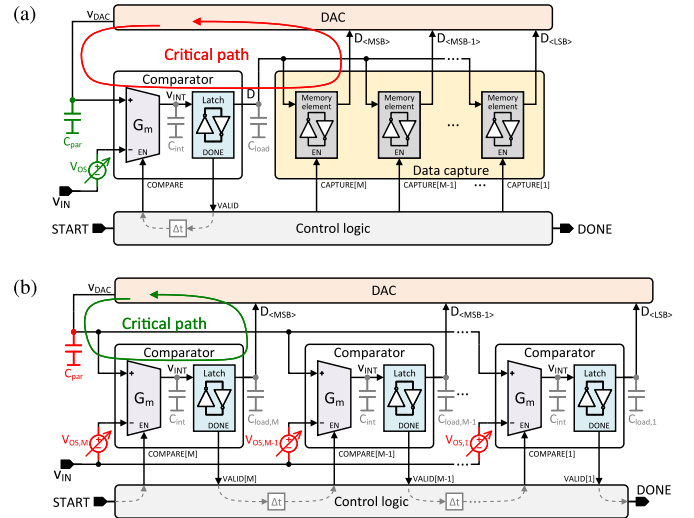


Fig. 2. Traditional SAR quantizer topologies. (a) Single-comparator. (b) Multi-comparator (“loop-unrolled”).

comparison threshold(s) of a stage sub-ADC is a particularly attractive way to inject dither because it does not rely on analog matching [24]. As such, a SAR quantizer able to easily vary its comparison thresholds on-the-fly with no performance loss is highly desirable [25], [26].

Existing single-bit-per-cycle SAR quantizers can be broadly classified into single-comparator and multi-comparator implementations, according to the number of comparators used. These architectures are now reviewed in terms of critical-path length, number of comparator offsets needed to calibrate, and amenability for dither injection.

1) *Single-Comparator SAR Quantizer*: Fig. 2(a) shows the structure of a single-comparator SAR quantizer [27]–[29]. As the name implies, it is implemented using a single comparator whose decisions are sequentially captured by several memory elements, one for each of the  $M$  bits resolved. The commonly-used comparator structure in [30], shown here, is composed of a memory-less preamplifier  $G_m$  followed by a regenerative latch. The preamplifier converts the input voltage into a current that is integrated on a capacitance  $C_{int}$  (usually a non-explicit, parasitic capacitor) and the resulting voltage  $v_{INT}$  is applied to the regenerative latch to produce the digital comparison result. Once the comparator resolves a bit (VALID signal is asserted) and its output is captured, it goes into reset, which usually happens within the time allocated for DAC settling, whose duration is set by ensuring an adequate delay  $\Delta t$  between the assertion of VALID and the launching of the next comparison.

Due to its use of a single comparator, this topology benefits from simple offset calibration, small input capacitance ( $C_{par}$ ) and minimal area. However, it suffers from long critical-path length due to the memory elements required for data capture, which also increase the comparator load ( $C_{load}$ ). Also, although it is possible to perform dither injection in this topology, existing methods usually increase the capacitive loading at critical internal nodes (e.g.,  $C_{int}$ ) [31], thus penalizing speed.

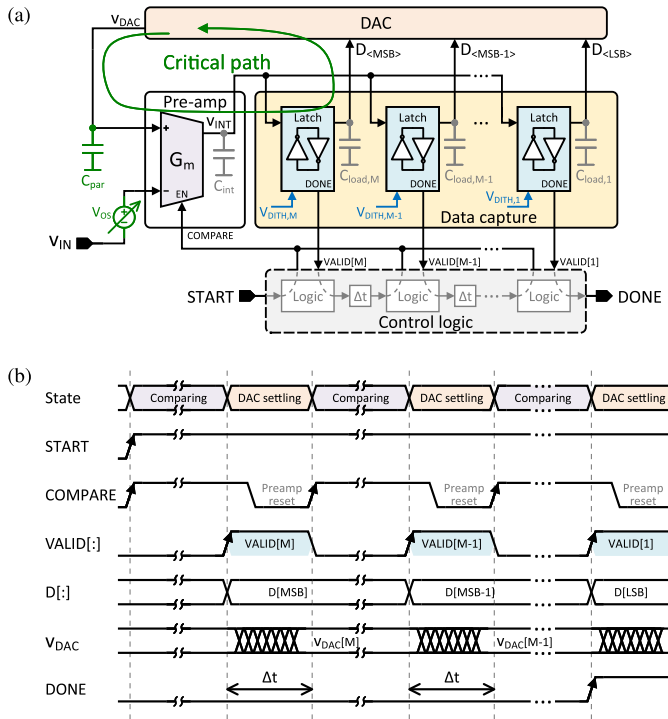


Fig. 3. (a) Proposed multi-latch SAR quantizer and (b) its timing diagram.

2) *Multi-Comparator SAR Quantizer*: The multi-comparator quantizer shown in Fig. 2(b) is an attempt to improve upon the classical single-comparator approach in several ways. Originally proposed in [2] under the name of “comparator-based SAR controller,” it is better known as the “loop-unrolled” SAR. In this architecture a dedicated comparator is used for each bit resolved, thereby eliminating the need for explicit data capture logic. Instead, the regenerative-latch output of each comparator is hardwired directly to a corresponding DAC control input. The control logic is also greatly simplified in this architecture and can be as simple as the  $\Delta t$  delays required to ensure adequate DAC settling. Finally, the use of dedicated comparators also enables the possibility of implementing per-bit dither injection in this architecture.

Due to its minimal critical-path length, this topology achieves the fastest conversion speed. However, it suffers from complex offset calibration (as all the individual offsets  $V_{OS,i}$  need to be compensated) and large input capacitance and area.

### B. Proposed Multi-Latch SAR Quantizer

Fig. 3 shows the proposed quantizer along with a simplified timing diagram. Conceptually lying halfway between the single- and multiple-comparator approaches, it is composed of a single preamplifier that sequentially drives multiple latches. Like in the multi-comparator quantizer, these latches serve as the data capture register and interface to the DAC directly without the need for additional memory elements.

This multi-latch architecture retains the low calibration complexity benefit of the single-comparator topology, because the shared preamplifier dominates the input-referred offset due

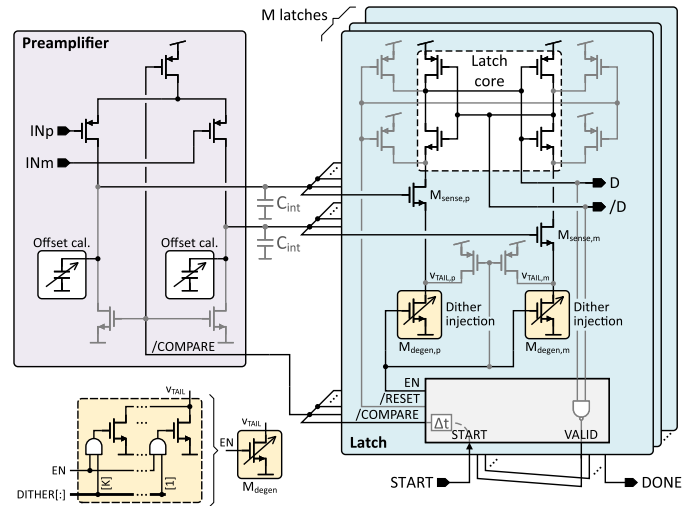


Fig. 4. Detailed implementation of the proposed quantizer.

to random mismatch. It also inherits the speed benefits of the loop-unrolled topology due to the direct connection of latch outputs to DAC inputs, while avoiding its drawbacks of increased offset calibration complexity and large input capacitance. In addition, like the multi-comparator quantizer, the multi-latch quantizer is capable of variable-amplitude, per-bit dither injection, which can be accomplished by embedding offset voltages  $V_{DITH,i}$  in the individual latches.

The defining features of the proposed architecture also introduce some practical implementation challenges. Due to its multiplexed use, the preamplifier should ideally be able to reset in the time  $\Delta t$  budgeted for DAC settling, as illustrated in Fig. 3(b). Also, in this topology the preamplifier load increases with the number of resolved bits. This, however, can be of no concern if the number of bits is kept low or if the preamplifier load needs to be high anyway due to noise constraints [32].

The implementation of the proposed quantizer is detailed in Fig. 4. The dynamic preamplifier and the regenerative latches are similar to the ones in [33], with appropriate modifications. The control logic block in each latch includes a tri-state-output buffer for driving the  $\text{/COMPARE}$  control line. The latches thus take turns sequentially to connect to this line in order to fire and reset the preamplifier for each comparison. Also, capacitive DACs are added at the preamplifier outputs to compensate for its offset.

In the latches, tail devices  $M_{degen,(p,m)}$  and local control logic are added to allow the selective enabling or disabling of the core latch by means of the control signals EN and  $\text{/RESET}$ . Their operation is as follows. When a given latch is active, both its EN and  $\text{/RESET}$  signals are high and its tail nodes  $v_{TAIL,(p,m)}$  see a low impedance to ground, allowing the sensing devices  $M_{sense,(p,m)}$  to sense the preamplifier output polarity and transfer it to the latch core. Conversely, when a latch is inactive its EN and  $\text{/RESET}$  signals are low so that its tail nodes are pulled-up to the positive supply along with all the latch-core nodes, thereby shutting down the sensing devices and effectively disconnecting the latch from the preamplifier. Once a latch resolves a bit, its VALID signal is fed to

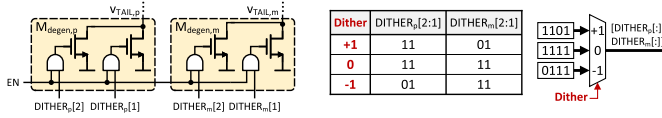


Fig. 5. Example of dither injection in a latch using source-degeneration arrays with  $K = 2$  devices per side.

the subsequent latch and used to fire the next comparison after a delay  $\Delta t$  allocated for DAC settling. The latches retain their decisions until the full quantizer is reset.

As depicted in Fig. 4, the per-bit dither injection is accomplished by implementing the latch tail devices  $M_{\text{degen},(p,m)}$  as programmable source-degenerating arrays, each composed of  $K$  NMOS transistors that can be independently enabled according to the value of the  $\text{DITHER}[\cdot]$  codeword. Thus, offset voltages can be introduced by using different  $\text{DITHER}[\cdot]$  values on each half of the latch, effectively shifting the comparison threshold by a corresponding  $V_{\text{DITH},i}$  amount. As an example, Fig. 5 shows a minimalistic implementation with  $K = 2$ , illustrating how the three-level logic dither values can be translated into suitable codewords via a MUX whose output controls how many source-degenerating devices are enabled on each latch side. The implementation of Fig. 4 has the benefit of supporting variable-amplitude dither injection natively, as the magnitude of the introduced offset can be made variable by having  $K \geq 2$  unary devices or by using non-unary devices. The dither codewords  $\text{DITHER}[\cdot]$  are programmed into the latch before the quantizer begins operating (during input signal tracking), which significantly relaxes the speed requirements and the power spent for dither injection in comparison to a single-comparator architecture.

In this work, the frontend quantizer is sized for speed, achieving a 1-mV-input resolution time of 90 ps, an input-referred noise of 1 mV<sub>rms</sub>, and consuming 140 fJ/bit. Its input capacitance is 3 fF and the residual offset after calibration is  $<5$  mV<sub>rms</sub>, which together with the injected dither (5 mV nominal) is absorbed by the interstage redundancy. The preamplifier load  $C_{\text{int}}$  is 4.8 fF, composed of 3.4 fF from the input capacitance of the six latches and 1.4 fF from routing parasitics.

### III. BACKGROUND CALIBRATION

Stage-1 DAC mismatch and interstage gain are digitally corrected in this design, leveraging on-chip, noninvasive, narrowband-dither calibration mechanisms that operate in the background during normal ADC operation. In this section, we discuss the principles and implementation aspects of these techniques. Moreover, we show how the injected dither is also leveraged for linearity and bias optimization of the residue amplifier, ensuring robust, PVT-resilient operation.

#### A. Traditional Broadband-Dither-Based Calibration

We start by conceptualizing the general dither-based background calibration principle, considering its application to interstage-gain correction. Let us assume the canonical two-stage pipelined architecture shown in Fig. 6(a) where

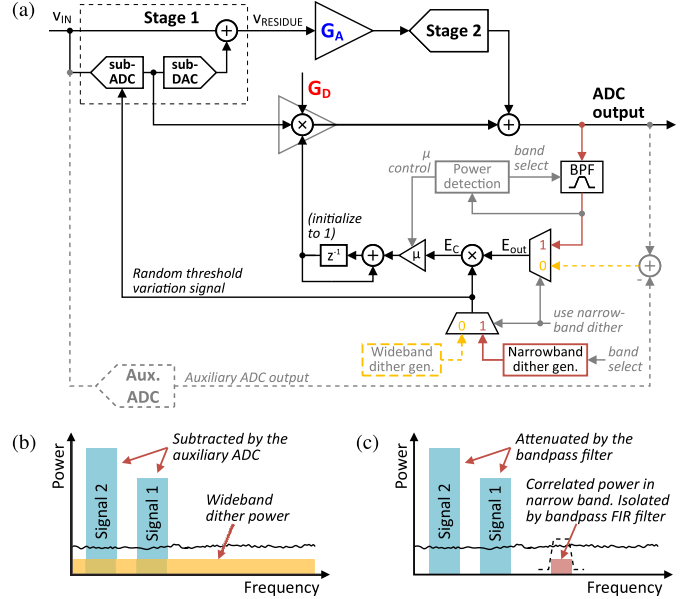


Fig. 6. (a) Conceptual dither-based interstage-gain calibration and associated dither injection (b) wideband and (c) narrowband dither injection.

the sub-ADC within the first stage contains a programmable input-referred offset that can be updated from cycle to cycle. Moreover, let us assume that the interstage amplification  $G_A$  has 1-bit redundancy, such that under ideal operating conditions the amplified residue never exceeds half of the Stage-2 quantization range. When a PN dither sequence is used to modulate the sub-ADC offset, a proportional increase in quantization noise will appear in the analog output residue of Stage 1. However, thanks to redundancy, the error is fully corrected during the digital output reconstruction [24], and as long as the residue  $v_{\text{RESIDUE}}$  (after amplification by  $G_A$ ) falls within the linear quantization range of Stage 2, the impact of the “error” induced by the sub-ADC offset variation is nulled. However, for accurate (corrective) reconstruction, it is imperative that the analog interstage gain  $G_A$  is known, i.e., that the digital reconstruction gain  $G_D$  matches  $G_A$  to sufficient accuracy. Under these conditions, the overall ADC digital output remains unaffected by the injected dither, accurately representing the quantized input signal  $v_{\text{IN}}$ . In contrast, when  $G_A \neq G_D$  the ADC output becomes corrupted. It inaccurately represents  $v_{\text{IN}}$ , and carries some non-zero power correlated with the injected dither. This dither-correlated energy can be tracked and minimized in the background by appropriate readjustment of  $G_D$ , and an LMS-based correction loop [see Fig. 6(a)] can be built in the digital domain to achieve this, as described below.

For the sake of explanation, let us initially assume that the injected dither is wideband, as depicted by the spectrum in Fig. 6(b) [and flag use narrowband dither = “0” in Fig. 6(a)]. Moreover, let us assume there exists an ideal auxiliary ADC which converts the same input signal  $v_{\text{IN}}$  as the main ADC, but without the presence of dither [see Fig. 6(a)]. The dither-correlated energy at the main ADC output can then be reliably tracked, as  $v_{\text{IN}}$  can be easily removed from the

main ADC output by subtracting the auxiliary ADC output to produce a difference  $E_{OUT}$  that holds information pertaining only to errors induced within the main ADC. To estimate the dither-correlated energy,  $E_{OUT}$  is multiplied with the injected dither sequence, obtaining a product  $E_C$  that will have a positive or negative dc component depending on whether  $G_A$  is larger or smaller than  $G_D$ , respectively. This signal is then digitally integrated (with an integration gain determined by a constant  $\mu$ ), and the output of the integrator is used to scale the digital reconstruction gain  $G_D$ . The integrator will then gradually drift toward an optimum value that ensures  $G_D = G_A$ . Under this condition the reconstruction is again accurate,  $E_C$  becomes a zero-mean signal, and the integrator output varies only slightly around its optimum value due to the presence of other noise and error sources in the main ADC output that the auxiliary ADC does not cancel. Notably, the convergence speed of this algorithm is proportional to the integration constant  $\mu$  but the variation around the final estimate (i.e., noise in the estimation) is inversely proportional to it. In other words, the integration constant  $\mu$  trades off convergence speed for estimation noise.

Obviously, an ideal auxiliary ADC does not exist in practice, so the input signal cannot be subtracted from the ADC output before correlation with the dither sequence. Alternatively, the algorithm can also operate without the removal of  $v_{IN}$ , but then the small dither-correlated power present in  $E_{OUT}$  is swamped by the much larger signal power now also present in  $E_{OUT}$ , which injects a large amount of noise into  $E_C$ . To circumvent this, the integration constant  $\mu$  needs to be extremely small, and although this filters out the estimation noise, it also slows down the estimation process, rendering the algorithm impractical for real-time PVT tracking.

To address the problem of limited estimation speed and accuracy, [24] makes use of a split-ADC architecture where two parallel, half-sized main ADC paths operate with the same input signal but with independent dither. Although the signal components can then be successfully removed from the ADC output to isolate the dither information, this architecture comes at the cost of additional analog overhead, increased parasitics, and frontend-sampler matching issues.

### B. Proposed Narrowband-Dither-Based Calibration

To overcome the aforementioned issues, we propose an alternative method based on narrowband dither injection. This is depicted by the spectrum in Fig. 6(c) [and flag use narrowband dither = “1” in Fig. 6(a)]. As shown in this figure, the dither is injected in a narrow band that does not overlap with the input signal band, and is isolated at the ADC output using a digital bandpass filter (BPF) centered at the injection band. This allows for reliable tracking of dither-correlated energy and permits the use of an integration constant  $\mu$  that provides fast estimation, while avoiding the complications of the auxiliary-ADC and split-ADC implementations.

The convergence speed of the proposed algorithm depends on the ratio of the dither-correlated power to other uncorrelated power (including the signal power) at the integrator input. This has to be taken into account when considering the location

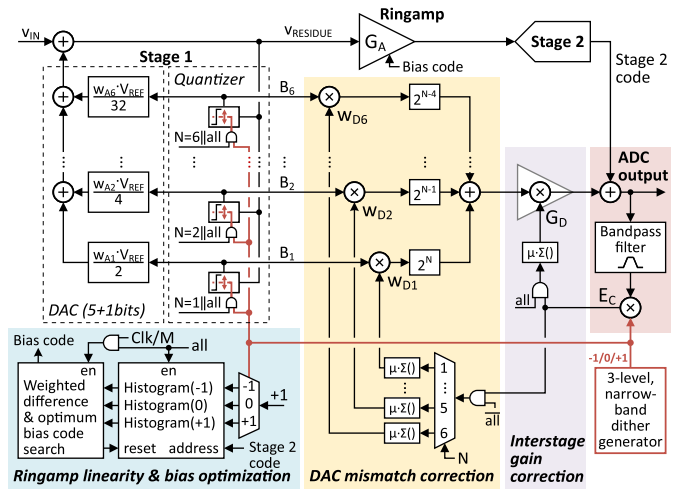


Fig. 7. Implemented calibration engines (color-shaded blocks implemented off-chip; power-detector-based control of Fig. 6(a) omitted for simplicity).

of the dither band and its bandwidth. Under the assumptions of ideal bandpass filtering, uniformly-distributed dither power (non-overlapping with the signal) and white noise, this “dither-correlated-power-to-other-power” ratio is independent of the dither bandwidth size. In practice, however, the largest contributor to the estimation noise is the input signal itself through its leakage to the integrator input due to non-ideal filtering. Hence, a sharp and narrow BPF which cancels as much signal leakage as possible is preferred. From another perspective, a narrow dither bandwidth is also preferred so that most the spectrum can be used for the input signal. This naturally comes at the cost of digital power and area (i.e., more taps are necessary to realize a sharper filter profile).

An important consideration for the practical implementation of the proposed algorithm is the potential overlap between the input signal and the dither band due to variation of the signal spectrum across time. As shown in Fig. 6(a), this can be addressed by means of a digital power detector at the BPF output, which measures the power in the dither band (based on an estimation of the RMS value of the digital BPF output). This power is ideally composed exclusively of dither, and its approximate expected level is known (and can be simulated). Any excess power must come from the input signal, and achieves a maximum when the signal and dither bands fully overlap. Based on the information provided by the power detector, the estimation error (i.e.,  $\mu$ ) can be adjusted automatically. Alternatively, the dither band can be moved to a different location (after a programmable time-out), effectively implementing a search for a band without signal power that can be used for dither injection. Critically, any non-ideal overlap of signal and dither bands will affect the convergence speed of the estimator, but convergence is ultimately always guaranteed independent of signal characteristics and/or filter configuration. Likewise, any mismatch between the dither injection and extraction filters affects only convergence speed due to lost dither power, but not convergence itself.

The proposed approach is used in this work to implement the interstage-gain and DAC-mismatch calibration engines shown in Fig. 7, as described in Sections III-C–III-E.

### C. Interstage-Gain Calibration

As described in Section II, this work exploits a quantizer capable of per-bit dither injection. Based on a three-level PN dither sequence, the Stage-1 quantization thresholds can be either fixed to a nominal position (dither code 0) or offset to some positive or negative value (dither codes +1 and -1, respectively), as shown in Fig. 7. For interstage-gain calibration, a common dither value must be used in all thresholds within a single conversion cycle (control signal all = “1”). For DAC mismatch calibration only one of the thresholds within a conversion cycle is dithered (signaled by the control signal  $N$ ), while all the other remain in their nominal position (all = “0”).

In the presence of a gain error, i.e., when  $G_A \neq G_D$  (we initially assume no DAC mismatch:  $w_{A,i} = w_{D,i} = 1$ ), the injected quantizer-threshold dithering signal induces an error during the Stage-1 SAR conversion. This error is tracked in the narrow band of interest, which is isolated at the ADC output by a digital bandpass FIR filter, and used in an LMS-based correction loop (implemented off-chip in MATLAB). Power-detector-based adjustment of the convergence rate (scaling factor  $\mu$ ) is continuously active in the background, to ensure reliable estimation independently of the relationship between the signal and the dither bandpass positions (the power detector is omitted in Fig. 7 for simplicity).

In this design, the BPF passband is set to roughly 5% of the available Nyquist band. A 32-tap FIR filter is used both in the dither injection (implemented on-chip) and in the dither extraction at the ADC output (implemented off-chip). With a value for  $\mu$  chosen such that estimation noise has no significant impact on the results, an initial uncorrected 5% interstage-gain error can be calibrated within  $\sim 350$  kcycle, as shown in Fig. 8(a). A convergence-speed comparison with the prior-art wideband dither technique is also shown in Fig. 8(b). More than two orders of magnitude in speed enhancement are achieved with the proposed narrowband technique, when the signal is well isolated from the dither band. Finally, Fig. 8(c) shows a sweep of the BPF center frequency for a fixed input signal (for multiple ADC output-data captures). The algorithm reliably estimates the interstage gain thanks to the power-detector-based scaling of  $\mu$ , which is adaptively adjusted to provide sufficient attenuation of the signal-induced error in the gain estimation when the signal and dither bands overlap.

### D. DAC Mismatch Calibration

When random capacitor mismatch is present in the Stage-1 DAC, the corresponding bit weights do not follow binary scaling anymore (i.e.,  $w_{A,i} \neq 1$  in Fig. 7). Just as with the interstage-gain error, DAC mismatch effects are nulled during digital reconstruction when the reconstruction coefficients accurately match their analog counterparts,  $w_{A,i} = w_{D,i}$  (assuming  $G_D = G_A$ ). As shown in Fig. 7, this design also uses narrowband-dither-based calibration to estimate the optimal reconstruction coefficients  $w_{D,i}$ . The DAC bit weights are calibrated sequentially, starting from the MSB. Only one of the comparisons within a quantization cycle is dithered at a time (e.g.,  $N = 1$  for calibration of the MSB), while all others remain in their nominal position (all = “0”). In this

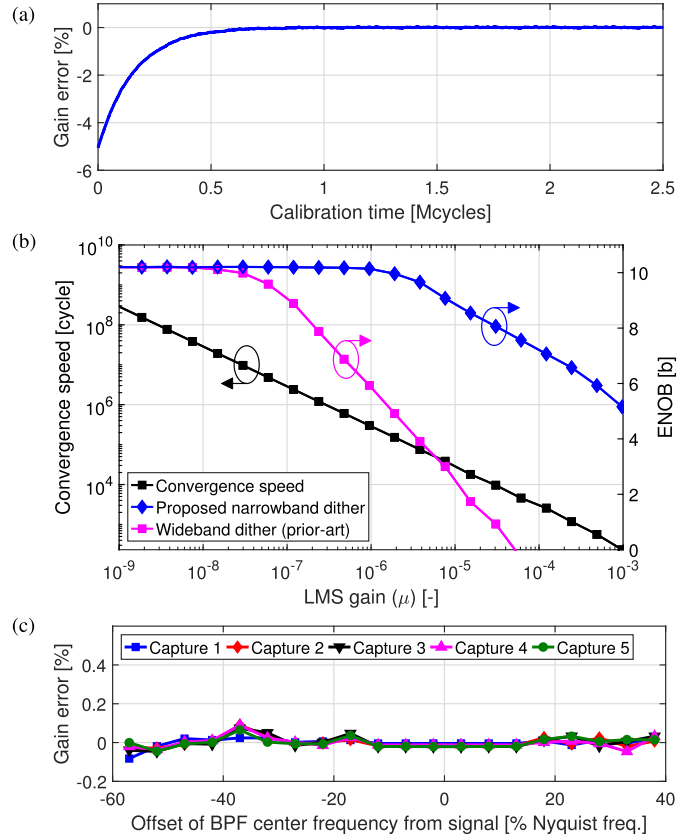


Fig. 8. Measured interstage-gain estimation. (a) Convergence speed. (b) Convergence speed comparison with wideband dither injection. (c) Gain estimation error for variable BPF center frequency (and fixed  $v_{IN}$  location).

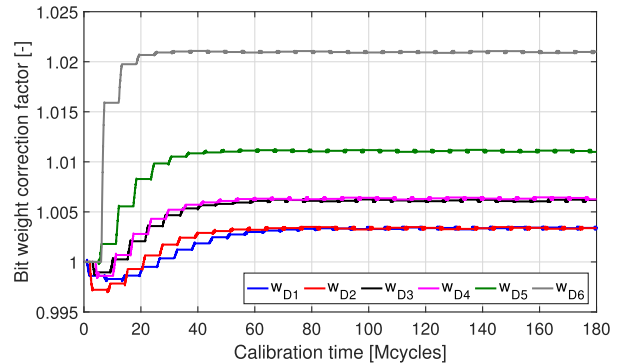


Fig. 9. Measured estimation of the optimal reconstruction coefficients  $w_{D,i}$ .

way, it is possible to independently track the dither-correlated energy present at the ADC output for each DAC capacitor. The correction mechanism and its convergence speed are similar to that of the interstage-gain calibration, as shown<sup>1</sup> in Fig. 9.

Note that the DAC mismatch relates to MOM capacitor mismatch, which is insensitive to environmental variations (e.g., changes in temperature and supply voltage). Therefore,

<sup>1</sup>The “staircase” shape of the convergence curves in Fig. 9 results from the particular implementation of the calibration algorithm, in which the bit weights are not estimated using single attempts but following an iterative procedure: each weight is estimated during a fixed number of cycles and then held, and after all the weights have been estimated, the process is repeated.

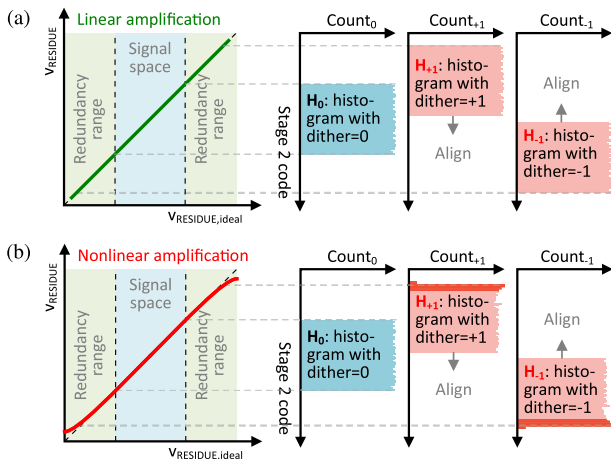


Fig. 10. Nonlinearity estimation based on backend-code histogram analysis. (a) Linear and (b) nonlinear amplification of a uniformly-distributed input.

the mismatch calibration engine estimating  $w_{D,i}$  needs to operate only at chip startup (with interstage-gain calibration initially disabled). After all the  $w_{D,i}$  are estimated, they remain fixed and the interstage-gain calibration takes over.

As discussed in Section II-B, the proposed dither injection method requires negligible analog overhead. Varying the quantizer thresholds is simple, and there is no direct critical-timing limitation (quantizer offsets can be programmed while the stage is tracking its input). Another important advantage is robustness to analog inaccuracies. This stems from the fact that the dither is injected into the analog signal path in the form of quantization error [24]. When the optimal digital reconstruction coefficients are known, the digital redundancy of the system cancels this quantization error. Thus, the absolute accuracy of the injected dither is irrelevant for the convergence or accuracy of the proposed calibration algorithms, and as long as redundancy bounds are not exceeded and as long as the offsets introduced in the quantization thresholds vary monotonically with the logic dither values ( $-1, 0, +1$ ), there are no matching or analog precision requirements for the calibration hardware to operate robustly across PVT variations.

### E. Residue-Amplifier Linearity and Bias Optimization

By exploiting the proposed quantizer and the injected three-level dither, linearity and bias optimization of the residue amplifier can also be performed in the background at no extra analog hardware cost. The method, illustrated in Fig. 10, is based on the analysis of the backend code histograms that result for each applied dither value. For now, let us assume that the statistical distribution of the analog residue generated by Stage 1 is uniform and bounded to 1 LSB. In the case of perfectly-linear amplification, the amplified residue and the Stage-2 code histogram will also be uniformly distributed. Injecting dither through the Stage-1 quantizer will shift the Stage-2 histogram by a fixed value, using some of the interstage redundancy range. Critically, this does not change the uniform shape of the Stage-2 histogram, only its offset [see Fig. 10(a)]. Now, let us consider the practical case of nonlinear amplification. Since each dither code induces a different

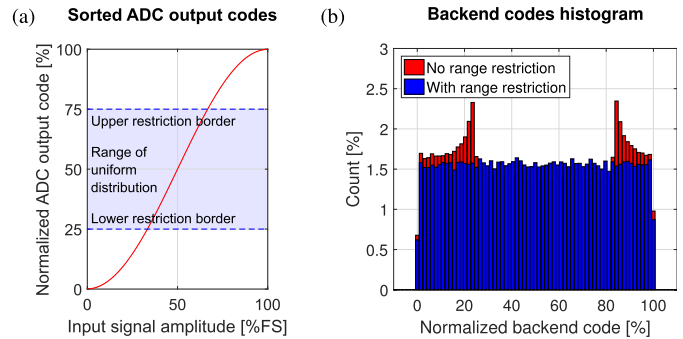


Fig. 11. (a) Range restriction concept for uniformization of the backend histograms under sinusoidal-input excitation and (b) backend code histograms.

offset in the amplified residue, the latter experiences different nonlinear transformations, and this distortion is captured in the Stage-2 histogram. For example, in the case of compressive nonlinearity [see Fig. 10(b)], the compression at the edges of the amplifier output range will cause the histograms for the  $+1$  and  $-1$  dither cases to be more nonuniform than the zero-dither case that lies in the central, less-compressed region. From these thought experiments, the main principle of the proposed technique emerges: the more linear the amplifier, the more invariant the Stage-2 histogram shape will be to dither.

Practical implementation involves injecting a three-level PN dither sequence into the Stage-1 quantizer and then sorting the Stage-2 output codes into three histograms according to the injected dither values. To calculate an estimate of the amplifier nonlinearity, the histograms are first aligned for maximum correlation. Then we compute the metric

$$NL = \sum (H_0 \times |H_{+1} - H_0| + H_0 \times |H_{-1} - H_0|) \quad (1)$$

which involves: 1) calculating the absolute differences between histograms, bin-by-bin (these differences are smaller with linear amplification); 2) scaling the differences by the non-dithered-histogram bin values; and 3) summing all the products to create a single nonlinearity metric: NL. This value serves as an indicator of the degree of amplification nonlinearity present and can be used as an objective function to optimize the residue amplifier biasing so that  $NL \rightarrow 0$ .

For reliable NL estimation, the input signal must preferably be uniformly distributed, which is typically true for realistic communication channels. In the presence of a deterministic sinusoidal excitation (typically used for testing purposes in the laboratory), this is no longer the case. An example of a backend histogram for such a case is depicted on Fig. 11(b) (red histogram). Codes corresponding to the peaks of the sine wave are excited more frequently, resulting in a characteristic spiking in the histogram.<sup>2</sup> This nonuniform histogram profile can degrade the accuracy of the NL estimate. To overcome this issue, we restrict the range of ADC output codes for which backend histogram information is collected. As shown in Fig. 11(a), the histogram is updated only for the center half of the ADC output range, where the codes distribute

<sup>2</sup>Note that since Stage 1 operates with a redundant LSB decision [34], the spiking “folds” within the histogram itself and it does not appear at its edges.













