

Noise-Shaping Binary-to-Stochastic Converters for Reduced-Length Bit-Streams

Kleanthis Papachatzopoulos and Vassilis Paliouras, *Member, IEEE*

Abstract—Stochastic computations have attracted significant attention for applications with moderate fixed-point accuracy requirements, as they offer minimal complexity. In these systems, a stochastic bit-stream encodes a data sample. The derived bit-stream is used for processing. The bit-stream length determines the computation latency for bit-serial implementations and hardware complexity for bit-parallel ones. Noise shaping is a feedback technique that moves the quantization noise outside the bandwidth of interest of a signal. This paper proposes a technique that builds on noise shaping and reduces the length of the stochastic bit-stream required to achieve a specific Signal-to-Quantization-Noise Ratio (SQNR). The technique is realized by digital units that encode binary samples into stochastic streams, hereafter called as binary-to-stochastic converters. Furthermore, formulas are derived that relate the bit-stream length reduction to the signal bandwidth. First-order and second-order converters that implement the proposed technique are analyzed. Two architectures are introduced, distinguished by placing a stochastic converter either inside or outside of the noise-shaping loop. The proposed bit-stream length reduction is quantitatively compared to conventional binary-to-stochastic converters for the same signal quality level. Departing from conventional approaches, this paper employs bit-stream lengths that are not a power of two, and proposes a modified stochastic-to-binary conversion scheme as a part of the proposed binary-to-stochastic converter. Particularly, SQNR gains of 29.8 dB and 42.1 dB are achieved for the first-order and second-order converters compared to the conventional converters for equal-length bit-streams and low signal bandwidth. The investigated converters are designed and synthesized at a 28-nm FDSOI technology for a range of bit widths.

Index Terms—stochastic computing, Sigma-Delta modulators, noise shaping, stochastic bit-stream, digital filtering

1 INTRODUCTION

SIMPLICITY of Stochastic Computation (SC) systems renders them attractive for important contemporary applications, such as neural networks [1], [2], [3], digital filtering [4], and image processing [5], [6]. Moreover, for energy-demanding and hardware-costly applications, such as error-correction decoding [7] and MIMO detection [8], [9], stochastic solutions compete successfully with binary implementations. Although their randomness imposes challenges for applications requiring increased accuracy, SC systems, due to their inherent error tolerance, can outperform binary counterparts in terms of accuracy when subjected to aggressive voltage overscaling [10]. They also facilitate sub-threshold implementations for power-critical systems and reveal the potential for both power, area, and reliability benefits [11]. However, increasing precision by one bit doubles bit-stream length, incurring latency bottlenecks. The motivation for this paper is to reduce stochastic bit-stream length in SC systems taking into account the input signal bandwidth without increasing quantization in-band noise power level. Typical applications include digital filtering [4] and all-digital transmission schemes [12]. Essentially, the stream length reduction reduces SC latency and/or hardware complexity.

Binary-to-stochastic converters drive SC systems and determine the precision of the subsequent operations. Recent research focuses on randomization properties of the random

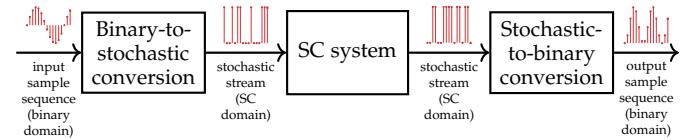


Fig. 1. Organization of a dynamic stochastic computing system.

number generators (RNGs) involved in these, so as to create bit-streams with the minimum required length. Liu and Han [13] introduce converters that rely on Sobol sequence generators. While favoring parallel implementations, these random number generators suffer from increased area overheads. Najafi *et al.* [14] propose the techniques of clock division, rotation, and relative prime stream length to improve the randomization of unary bit-streams. The aforementioned techniques are also adopted for converters that exploit Halton generators [15] as RNGs. Najafi and Lilja [16] propose the utilization of primitive linear feedback shift registers (LFSRs) for RNGs and for SC systems that employ pseudo-random bit-streams, as they generate bit-streams with lower mean absolute error than unary bit-streams, when operating for specific operation cycles. LFSR-based converters are advantageous as produced stream-bits are more randomized than in unary bit-streams. Salehi [17] develops an LFSR-sharing technique between multiple converters that minimizes the stochastic cross-correlation of generated streams. Architectures in [18] produce bit-streams by bit-permuting the input binary number.

SC systems can operate on bit-streams that encode sample sequences of time-varying signals [19], [20]. In the

Manuscript received July 17, 2023; accepted March XX, XXXX. Date of publication May XX, XXXX; date of current version July 17, 2023.

• The authors are with the Department of Electrical and Computer Engineering, University of Patras, 26504 Patras, Greece.
E-mail: papachatz@ece.upatras.gr; paliouras@ece.upatras.gr

context of dynamic stochastic computing, a discrete-time signal is encoded by a binary-to-stochastic converter, and the digital signal processing that follows is performed by stochastic elements, as shown in Fig. 1. The input signal is assumed to be sufficiently oversampled before its conversion to the SC domain to avoid aliasing. Following processing, the signal is reconstructed to the binary domain. In this context, except for the conventional conversion approaches discussed above [13]–[18], Sigma-Delta ($\Sigma\Delta$) modulators operate as binary-to-stochastic converters for dynamic SC systems as well. Furthermore, $\Sigma\Delta$ modulators serve as a low-cost conversion solution and demonstrate remarkable performance for finite and infinite impulse response filtering [21]–[24], where modulators drive the digital filter. Previous research investigates the utilization of conventional $\Sigma\Delta$ modulators as binary-to-stochastic converters [25], [26]. Gonzalez-Guerrero *et al.* [25] investigate the design of $\Sigma\Delta$ modulators for SC processing that employ asynchronous modulation. A recent approach in [3] exploits $\Sigma\Delta$ principles to facilitate multi-bit addition, and forms a multi-layer perceptron, with an inference accuracy similar to that of fixed-point implementations. Prior research efforts for SC systems consider only one-bit quantization and directly drive SC systems with the output of $\Sigma\Delta$ modulators and single-bit streams. Although multi-bit quantizers for $\Sigma\Delta$ modulation have been extensively discussed in the literature [27], they are not adopted in the SC framework as converters. In this paper, we combine a multi-bit quantizer as a part of the modulator, and a binary-to-stochastic converter that maps the quantized word to a stochastic stream. In contrast to a conventional $\Sigma\Delta$ modulator that produces a single bit for every k -bit binary input word, the converter's output in the proposed scheme is encoded as a multi-bit stream.

Typically, $\Sigma\Delta$ modulators process oversampled signals [27] that lie in a broad spectrum of frequencies [28], [29]. For applications targeting RF frequencies, the sampling rate is an essential metric for their operation. In this direction, efficient techniques have been proposed in the literature to accommodate all-digital modulators with high sampling rates in the gigahertz region [28], [30]. The methods in [28], [30] are applicable for the proposed converters. Furthermore, the introduced noise-shaping conversion technique for stochastic streams can be implemented with a great range of multi-bit $\Sigma\Delta$ and high-order modulation architectures [31]. In this work, we mainly target at converter designs with moderate $\Sigma\Delta$ orders and a low implementation overhead, suitable for low-cost SC applications.

Generalizing and extending prior work [32], in this paper, a novel binary-to-stochastic conversion technique is introduced that effectively transfers the quantization noise out of the signal bandwidth, thus, increasing SQNR. Closed formulas are derived for the SQNR at the output of proposed converters in terms of signal bandwidth and bit-stream length. The introduced architectures produce bit-streams with a length which is not a power of two. We also propose a scaling technique for stochastic-to-binary reconstruction in the employed noise-shaping feedback loops. The introduced technique is potentially employed for the binary signal reconstruction at the output of SC systems, as in Fig. 1. The following contributions are introduced here:

TABLE 1
Summary of Notation

Symbol	Description	Figure
B	signal bandwidth	–
L	bit-stream length of conv. converter	Fig. 2
L'	bit-stream length of prop. converter	Figs. 6 and 12
$x[n]$	n th binary input sample	Fig. 2, 6 and 12
k	precision of $x[n]$ in bits	Fig. 2, 6 and 12
a	converter's input (unipolar domain)	Fig. 2
a_q	quantized value of a	Fig. 2
y	converter's output	Fig. 2
$x_i[n]$	integrator's output	Figs. 6 and 12
$x'[n]$ ($x''[n]$)	accumulator's output	Fig. 6 (Fig. 12)
$y'[n]$ ($y''[n]$)	reconstructed output	Fig. 6 (Fig. 12)
f_{stream}	stochastic domain operation freq.	Figs. 6 and 12
f_{bin}	binary domain operation freq.	Figs. 6 and 12
c	feedback scaling factor	Figs. 6 and 12
c_q	quantized value of c	Figs. 6 and 12
m	$1 \leq \frac{2^m}{L'} < 2$	Figs. 6 and 12

- A conversion technique that offers SQNR gains of 28.9 dB and 42.1 dB (fixed stream length), or bit-stream length reductions of 87.2% and 92.18% (fixed SQNR) for the first- and second-order converters, respectively, compared to a conventional converter;
- Four Lemmas that quantify the bit-stream length reduction as a function of signal bandwidth;
- Two noise-shaping conversion architectures that place the converter's core either inside or outside the noise-shaping loop;
- A novel scaling technique for the binary reconstruction of reduced-length bit-streams; and
- A quantitative analysis of the hardware cost of the proposed fixed-point implementations at a commercial-grade 28-nm standard-cell library.

The notation used in this paper is summarized in Table 1. The remainder of this paper is structured as follows: Section 2 presents the background of stochastic computing. Sections 3, 4 and 5 introduce the architecture and SQNR performance at the output of conventional, introduced first-order and second-order converters, respectively. Section 6 presents the proposed stochastic-to-binary converter. Section 7 evaluates hardware complexity, and Section 8 presents the simulation-based SQNR for the investigated converters. Finally, Section 9 discusses conclusions.

2 STOCHASTIC COMPUTATIONS BASICS

SCs rely on logical operations performed on streams of data. A stochastic bit-stream encodes a value as the probability p of obtaining a bit of the stream with a value of one. Values in a digital signal processing system typically lie in the range $[0, 1]$ or in $[-1, 1]$, possibly derived after analog-to-digital conversion or as an output of a memory block [33]. Employing the unipolar representation [20], a number a in the range $[0, 1]$ is converted to a bit-stream A so as $a = p(s_i = 1)$, where $s_i \in \{0, 1\}$ is the i th bit of the bit-stream A . The bipolar representation allows SC systems to process signed numbers, lying in the range $[-1, 1]$. In the bipolar representation, conversion of x is realized by

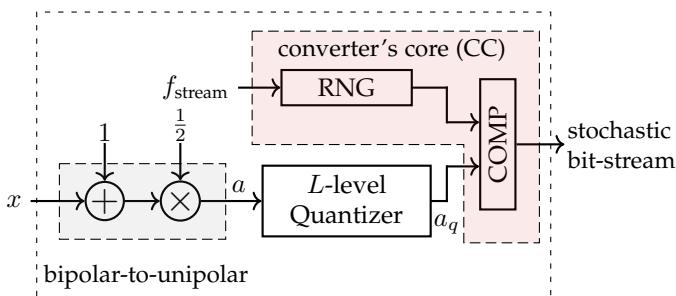


Fig. 2. Conventional bipolar binary-to-stochastic converter.

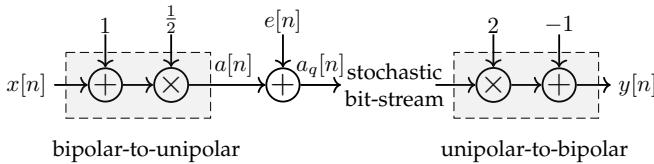


Fig. 3. Linear model of conventional binary-to-stochastic converter and reconstruction to bipolar domain.

initially mapping x to a number a , $0 \leq a \leq 1$, using the transformation

$$a = \frac{x + 1}{2}, \quad (1)$$

and subsequently mapping a to a unipolar stochastic bit-stream. The inverse stochastic-to-binary conversion is realized by initially deriving a from the bits s_i of the stochastic bit-stream, and subsequently transforming a to x , as

$$x = 2a - 1, \quad (2)$$

$$a = \frac{1}{L} \sum_{i=1}^L s_i, \quad (3)$$

where L is the bit-stream length.

SC systems offer massive parallelism [13], [14] since operations can be realized with minimal-complexity independent computational units. In particular, each parallel unit processes a certain bit of the bit-stream, as they are of equal weight [20]. Intermediate approaches process a sub-stream instead of a single stream-bit, trading-off area complexity for latency reduction [8], [34]. Furthermore, complicated arithmetic functions are implemented using simple gates only, achieving significant area savings. Stochastic-based approximations can be possibly implemented as truncated Taylor expansions [35], [36] or Bernstein polynomials [37]. Due to inherent stream randomness, the length of the bit-stream should be sufficiently longer than the expected precision of the results in order to maintain sufficient accuracy. Thus, reduced power dissipation compared to the binary counterparts can be achieved when accuracy is not the primary objective [5].

3 SQNR IN CONVENTIONAL BIPOLAR CONVERTERS

For the investigated architectures, it is assumed that the input binary word x has a fixed-point precision of k bits,

where the output stochastic bit-stream has a precision $\lceil \log_2 L \rceil$ bits, where $k > \lceil \log_2 L \rceil$. Fig. 2 illustrates the architecture of a conventional binary-to-stochastic converter (BSC) that assumes the bipolar format [16], [38]. The particular structure is used as a benchmark circuit for the evaluation of introduced converters. The input binary word is mapped to the unipolar domain, $[0, 1]$, and subsequently quantized by an L -level quantizer. As an RNG we employ a structure that produces only once every number in $[0, L - 1]$ with $\lceil \log_2 L \rceil$ -bit precision. The evolving output of the RNG is compared with the quantized word a_q and a stream-bit is obtained in every clock cycle. An L -bit stochastic stream is produced in L clock cycles. The only noise source in this system is the L -level quantizer. Mapping of the quantized word to a bit-stream does not introduce further noise. Hereafter, the concepts of bandwidth, oversampling, and quantization noise refer to the converter input and output sample sequences, shown in Fig. 1, and they are not related with the SC representation, employed for the encoding of discrete samples.

It is important to note that the proposed architecture is general enough to support any type of converter's core (CC), shown in Fig. 2, consisting here of an RNG and a comparator. Potentially, the CC can be implemented as an LFSR and a comparator [16], a Sobol-based [13], or a Halton-based converter [15]. The SQNR models introduced in this paper are valid, for any CC implementation that does not impose noise other than the quantization noise.

A linear model is derived by replacing the quantizer with an additive noise source, $e[n]$. We assume that a reconstructed signal, $y[n]$, is obtained at the system output as in (2). The linear model of conventional bipolar converter in Fig. 3, has two inputs, namely $x[n]$ and $e[n]$, while for the reconstructed signal it holds that

$$y[n] = 2\left(\frac{1}{2}(x[n] + 1) + e[n]\right) - 1 = x[n] + 2e[n]. \quad (4)$$

Based on (4), transfer functions for each input to the output can be derived. The output $y[n]$ comprises two components; i.e.,

$$y[n] = y_e[n] + y_x[n], \quad (5)$$

where $y_e[n]$ and $y_x[n]$ are the components due to $e[n]$ and $x[n]$, respectively. The transfer function $H_e(z)$ which maps the z -transform $E(z)$ of the noise $e[n]$ to the z -transform $Y_e(z)$ of the output component $y_e[n]$ is derived from (4) assuming $x[n] = 0$, as

$$y_e[n] = 2e[n] \Rightarrow H_e(z) = \frac{Y_e(z)}{E(z)} = 2. \quad (6)$$

The transfer function $H_x(z)$, derived from (4) for $e[n] = 0$, i.e.,

$$y_x[n] = x[n] \Rightarrow H_x(z) = \frac{Y_x(z)}{X(z)} = 1. \quad (7)$$

Denoting the power density spectrum of quantizer as $S(\omega)$ and for white noise $e[n]$, the power density spectrum of $y_e[n]$ is

$$\Phi(e^{j\omega}) = S(\omega) \|H_e(z = e^{j\omega})\|^2. \quad (8)$$

For the case that $\mathcal{E}(e[n]) = 0$, it holds that $S(\omega) = \sigma_e^2$, simplifying (8) to

$$\Phi(e^{j\omega}) = \sigma_e^2 \|H_e(z = e^{j\omega})\|^2. \quad (9)$$

In order to derive the SQNR, we assume a band-limited input signal $x[n]$ and integrate $\Phi(e^{j\omega})$ in the signal bandwidth. Assume that the band-limited input signal $x[n]$ has a bandwidth of $2B$ and power P_x . The power of the noise, P_{2B} , in the bandwidth following a low-pass filter with a cut-off frequency at $\omega_c = B$ is

$$P_{2B} = \frac{1}{2\pi} \int_{-B}^B \Phi(e^{j\omega}) d\omega. \quad (10)$$

Furthermore, the SQNR for the bipolar converter is

$$\text{SQNR} = \frac{P_x}{P_{2B}}. \quad (11)$$

In the following, we derive closed formulas for the SQNR at the output of a conventional bipolar binary-to-stochastic converter. In order to compute the quantizer in-band noise power, two cases are distinguished, namely rounding and truncation.

3.1 Noise power and SQNR for rounding

For a rounding quantizer, we assume that the error follows a uniform distribution and lies in $-\frac{1}{2L} \leq e[n] \leq \frac{1}{2L}$. This statistical noise model is valid for a sufficiently complex input signal and sufficiently small quantization steps, so as its amplitude can traverse many steps from sample to sample (cf. [33, p. 194]). From a practical perspective, rounding a number can be realized by adding 0.5 and then by keeping only the $\lceil \log_2 L \rceil$ most significant bits. Furthermore, the noise variance is [33], [39]

$$\sigma_e^2 = \mathcal{E}(e^2[n]) = \frac{1}{12L^2}. \quad (12)$$

From (6) and (9), the in-band noise power for rounding is

$$P_{2B,\text{round, conv}} = \frac{1}{2\pi} \int_{-B}^B 4\sigma_e^2 d\omega = \frac{B}{3\pi L^2}. \quad (13)$$

From (11) and (13), it is obtained that the SQNR for a conventional bipolar converter with a rounding quantizer is

$$\text{SQNR}_{\text{round, conv}} = \frac{3\pi L^2 P_x}{B}. \quad (14)$$

3.2 Noise power and SQNR for truncation

In case of a truncation quantizer, the error lies in $0 \leq e[n] \leq \frac{1}{L}$ [33], [39]. Practically, it is realized by keeping only the $\lceil \log_2 L \rceil$ most significant bits and is simpler than rounding since the addition of 0.5 is avoided. The noise spectral density is

$$S(\omega) = \frac{\pi}{2L^2} \delta(\omega) + \frac{1}{12L^2}, \quad (15)$$

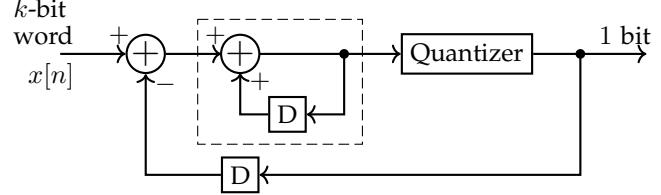


Fig. 4. $\Sigma\Delta$ modulator with 1-bit quantizer.

with $\omega \in [-\pi, \pi]$ and δ the Dirac delta function. For this case, the in-band noise power is

$$P_{2B,\text{trunc, conv}} = \frac{4}{2\pi} \int_{-B}^B S(\omega) d\omega = \frac{3\pi + B}{3\pi L^2}. \quad (16)$$

Therefore, from (11) and (16), the SQNR for a conventional bipolar converter with a truncating quantizer is

$$\text{SQNR}_{\text{trunc, conv}} = \frac{3\pi L^2 P_x}{3\pi + B}. \quad (17)$$

4 PROPOSED FIRST-ORDER SYSTEM

This section initially introduces two architectures for the proposed noise-shaping binary-to-stochastic conversion technique; the first one consists of the CC inside the noise-shaping loop, referred to as CC-I, and the second one consists of the CC outside the noise-shaping loop, denoted as CC-O. Then, the SQNR at the output of the introduced converter is derived, and formulas that quantify the bit-stream length reduction compared to a conventional converter are formulated.

4.1 Architectures of first-order noise-shaping bipolar converter

Both of the proposed architectures rely on a baseline $\Sigma\Delta$ modulator, indicatively depicted in Fig. 4. The block-level designs for the proposed CC-I and CC-O architectures are depicted in Fig. 5, described in a more detail in the remainder of this section.

4.1.1 CC-I architecture

The CC-I architecture for the proposed first-order noise-shaping bipolar converter is shown in Fig. 6. The feed-forward path of the first-order converter comprises a subtractor and an integrator (dashed box). A conventional binary-to-stochastic converter (BSC), shown in Fig. 2, follows the output of the integrator, $x_i[n]$. Resembling a $\Sigma\Delta$ modulator [27], the proposed architecture comprises a feedback path through which the output bit-stream returns and is reconstructed in order to be subtracted from the input. Subtraction is performed as a binary operation; hence, a modified stochastic-to-binary converter (SBC) is used to reconstruct the quantizer output from the stochastic bit-stream. The reconstructed sample is denoted as $y'[n]$ in Fig. 6. A proposed SBC accumulates the stream bits and performs a modified unipolar-to-bipolar conversion, described by $y'[n] = \frac{2c_q}{2^m} x'[n] - 1$, where $x'[n]$ is the accumulator

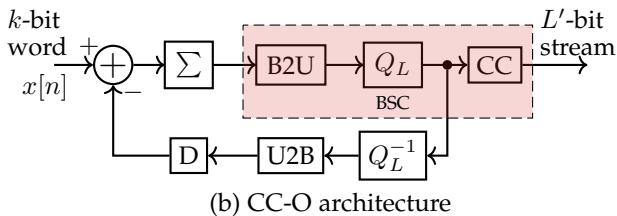
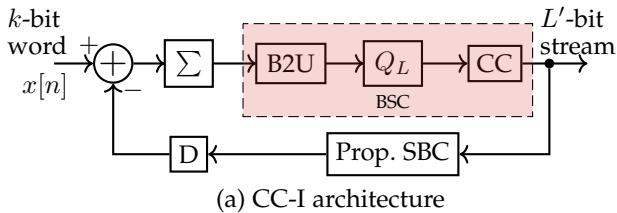


Fig. 5. Proposed first-order (a) CC-I and (b) CC-O noise-shaping converters. B2U block denotes bipolar-to-unipolar conversion, U2B is unipolar-to-bipolar conversion, \sum block is an integrator, Q_L is an L -level quantizer, and Q_L^{-1} is a mapper from L -level values to a binary representation.

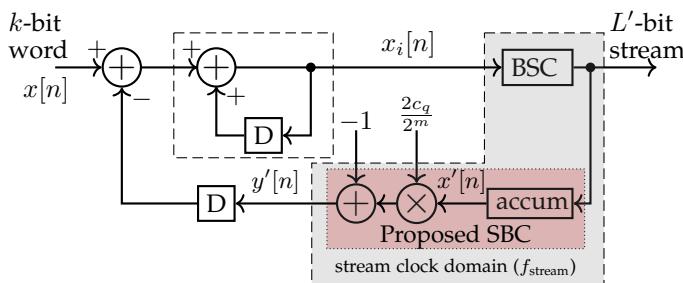


Fig. 6. First-order CC-I architecture of proposed binary-to-stochastic converter. Hardware units in gray region belong to stream clock domain, operating with a clock frequency f_{stream} .

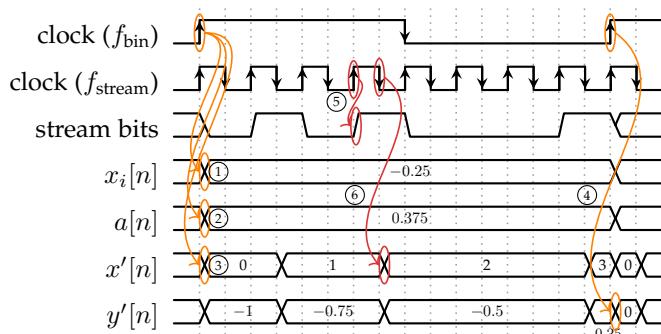


Fig. 7. Timing diagram of proposed first-order noise-shaping converter for $L' = 8$ bits.

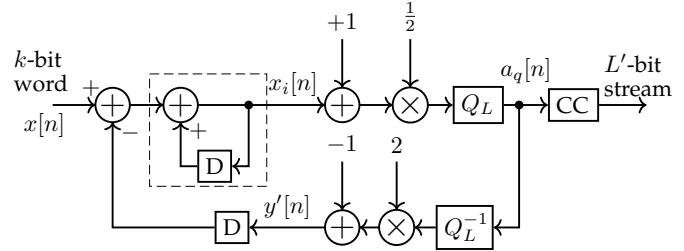


Fig. 8. First-order CC-O architecture of proposed binary-to-stochastic converter. Q_L denotes an L -level quantizer and Q_L^{-1} a mapper from L -level values to a binary representation.

output, and the constants c_q and m are defined in Section 6. While a conventional SBC reconstructs a binary value from a bit-stream by simple accumulation and a shift operation, the introduced modified SBC also scales its output by a factor c_q to compensate for the fact that the reduced bit-stream length is possibly not a power of two, as detailed in Section 6.

In the proposed BSC, for every k -bit binary word in the input of the converter, an L' -length bit-stream is produced. Binary additions at the feed-forward path are performed with a k -bit precision. Assume that the units of gray-shaded region in Fig. 6 lie in the stream clock domain, operating at a frequency f_{stream} , while the remainder units operate at a frequency f_{bin} . The bits composing the stream can be produced sequentially in L' clock cycles, or in parallel in a single cycle, or in N parallel (L'/N)-bit sub-streams. The proposed converter principles are straightforward to extend to parallel stream (or sub-stream) generation using a number of parallel BSC units, within which a RNG with a different initial state exists. Henceforth, we focus on sequential bit-stream generation. Bit-stream length of a noise-shaping converter determines operation frequencies by the ratio $f_{\text{stream}}/f_{\text{bin}} = L'/N$. The same principle also holds for a conventional BSC. Clock domains are separated by registers controlled by f_{bin} . Provided that the k -bit binary word in the input of the converter corresponds to a sample of a band-limited signal in $[0, B]$, the clock frequency of the converter should be $f_{\text{bin}} = \text{OSR} \cdot 2B$ to avoid aliasing, where OSR is the oversampling ratio.

To clarify the cycle-to-cycle operation, a timing diagram is presented in Fig. 7 that refers to a converter with $L' = 8$, where stream-bits are produced sequentially, and, thus, $f_{\text{stream}} = 8f_{\text{bin}}$. Specifically, signals $x_i[n]$ and $a[n]$ change with respect to f_{bin} clock, indicatively shown by transitions ① and ②, respectively, and $y'[n]$ is captured by f_{bin} clock, shown by transition ④, when stochastic-to-binary conversion is completed. Furthermore, stream bits are generated at the rising edge of f_{stream} clock, i.e., transition ⑤, the accumulator changes state at the falling edge of f_{stream} , i.e., transition ⑥, while it zeros at the rising edge of f_{bin} clock, i.e., transition ③.

4.1.2 CC-O architecture

The second proposed architecture for the introduced noise-shaping technique is demonstrated in Fig. 8. In contrast to CC-I, in the second one, the quantized word a_q returns through the feedback loop in order to be subtracted from the input. The CC in this case is moved outside the loop in order

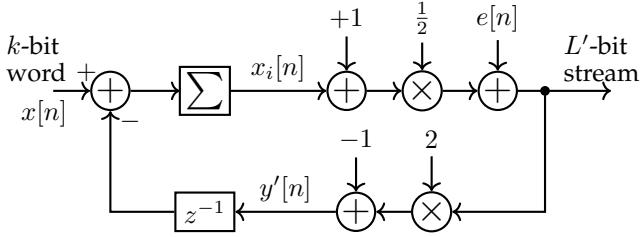


Fig. 9. Linear model for first-order noise-shaping system. Blocks marked with \sum are integrators, described by the transfer function $\frac{1}{1-z^{-1}}$.

to map the quantized binary word a_q to an L' -bit stochastic stream. Furthermore, a mapper block, denoted as Q_L^{-1} , is placed in the feedback loop. This block is necessary for the cases that L' is not a power of two, while it is trivial when L' is a power of two. Furthermore, the relation between clock frequencies of the two domains for this architecture is also determined by the expression $f_{\text{stream}}/f_{\text{bin}} = L'/N$. In essence, CC-O refers to a multi-bit $\Sigma\Delta$ modulator that produces a quantized binary value a_q , which drives a CC unit. Although CC-O is less elaborate than the first one, CC-I can shape any additional noise source introduced by the mapping of the quantized word to a bit-stream. This is possible when the RNG produces truly random numbers that may not follow a uniform distribution due to the limited length of the stream sequence [40], [41].

Stability analysis of both proposed noise-shaping architectures does not differ from that of a conventional multi-bit $\Sigma\Delta$ modulator [29], [42], [43].

4.2 Derivation of noise power and SQNR

Fig. 9 illustrates an additive noise model for the first-order noise-shaping converter. Since the binary signal is reconstructed in the feedback of both CC-I and CC-O, before subtracted from the current input sample, the noise model of Fig. 9 is the same for both architectures. Following the reconstruction to binary domain according to (2), the output $y'[n]$ comprises the components $y'_e[n]$ and $y'_x[n]$, induced by the noise $e[n]$ and the input $x[n]$, respectively, as

$$y'[n] = y'_e[n] + y'_x[n]. \quad (18)$$

Assuming zero input,

$$y'_e[n] = 2\left(\frac{1}{2}(x_i[n] + 1) + e[n]\right) - 1 = x_i[n] + 2e[n]. \quad (19)$$

A noise transfer function $H'_e(z)$ is derived by evaluating the transfer function of (19). Since

$$X_i(z) = -\frac{z^{-1}}{1-z^{-1}}Y'_e(z), \quad (20)$$

$$Y'_e(z) = X_i(z) + 2E(z), \quad (21)$$

it follows that

$$H'_e(z) = \frac{Y'_e(z)}{E(z)} = 2(1 - z^{-1}). \quad (22)$$

The noise power density spectrum is computed using (22),

$$\Phi'(e^{j\omega}) = S(\omega)\|H'_e(z = e^{j\omega})\|^2 = 16\sin^2\frac{\omega}{2}S(\omega), \quad (23)$$

and the in-band noise power for a truncating quantizer, using (10) and (23), is

$$P'_{2B} = \frac{16}{2\pi} \int_{-B}^B \left(\frac{\pi\delta(\omega)}{2L^2} + \frac{1}{12L^2} \right) \sin^2\frac{\omega}{2} d\omega \quad (24)$$

$$= \frac{2(B - \sin B)}{3\pi L^2}. \quad (25)$$

Notably, the in-band noise power for a rounding quantizer, computed using (10), (12) and (22), leads also to (25). From (25), the SQNR at the output of the proposed first-order bipolar converter, for both types of the quantizer, is

$$\text{SQNR}' = \frac{\pi}{2} \frac{3L^2 P_x}{(B - \sin B)}. \quad (26)$$

Furthermore, from (18) and assuming $e[n] = 0$, it holds that

$$y'_x[n] = x_i[n]. \quad (27)$$

Since

$$X_i(z) = X(z) - \frac{z^{-1}}{1-z^{-1}}Y'_x(z), \quad (28)$$

$$Y'_x(z) = X_i(z) \quad (29)$$

it follows that

$$H'_x(z) = \frac{Y'_x(z)}{X(z)} = 1, \quad (30)$$

where $Y'_x(z)$ and $X(z)$ are the z -transforms of time-domain signals $y'_x[n]$ and $x[n]$, respectively.

4.3 Bit-stream length reduction of first-order converter

The introduced first-order converter efficiently reduces the in-band quantization noise of an L' -length bit-stream. We seek for the length, L' , of a bit-stream produced by the introduced first-order noise-shaping converter that has equal SQNR as that of an L -length bit-stream of a conventional converter.

Lemma 1. A first-order converter with an L' -level quantizer, i.e., generating L' -length bit-streams, achieves the same SQNR as a conventional converter with an L -level rounding quantizer when

$$L' = L \sqrt{\frac{2(B - \sin B)}{B}}, \quad (31)$$

where B is the bandwidth of the input signal.

Proof: For the case of an L -level rounding quantizer of a conventional converter, from (14) and (26), it is obtained that

$$\text{SQNR}' = \text{SQNR}_{\text{round, conv}} \quad (32)$$

$$\frac{\pi}{2} \frac{3(L')^2 P_x}{B - \sin B} = \frac{3\pi L^2 P_x}{B} \Rightarrow$$

$$\frac{L'}{L} = \sqrt{\frac{2(B - \sin B)}{B}}. \quad (34)$$

□

Lemma 2. A first-order converter with an L' -level quantizer, i.e., generating L' -length bit-streams, achieves the same

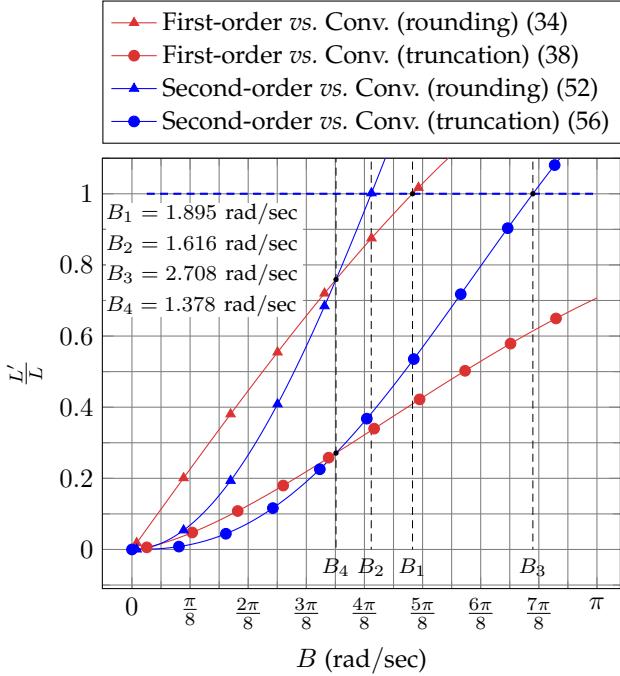


Fig. 10. Ratios of bipolar bit-stream lengths using a first-order and a second-order noise-shaping converter over a stream length of a conventional converter for the same SQNR. Conventional converter relies either on a rounding or on a truncating quantizer.

SQNR as a conventional converter with an L -level truncating quantizer when

$$L' = L \sqrt{\frac{2(B - \sin B)}{3\pi + B}}, \quad (35)$$

where B is the bandwidth of the input signal.

Proof: For the case of an L -level truncating quantizer of a conventional converter, from (17) and (26), it is obtained that

$$\text{SQNR}' = \text{SQNR}_{\text{trunc, conv}} \quad (36)$$

$$\frac{\pi}{2} \frac{3(L')^2 P_x}{B - \sin B} = \frac{3\pi(L)^2 P_x}{3\pi + B} \Rightarrow \quad (37)$$

$$\frac{L'}{L} = \sqrt{\frac{2(B - \sin B)}{3\pi + B}}. \quad (38)$$

□

Bit-stream length reduction ratios, (34) and (38), are plotted in Fig. 10 as a function of the signal bandwidth B . It can be seen that substantial bit-stream length savings, referring to $L' < L$, can be achieved as signal bandwidth decreases. Savings are observed for $B < B_1$, $B_1 = 1.895$ rad/sec, when the first-order system is compared to a conventional converter with rounding quantizer, while it is $L' < L$ within $[0, \pi]$ for a conventional converter with a truncating quantizer. The behavior of the introduced converter remains identical under either rounding or truncation; however, conventional converters are clearly affected.

4.4 SQNR bounds due to integer stream lengths

Lemmas 1 to 2 derive bit-stream lengths L' which maintain the SQNR provided by a conventional L -bit stream. As

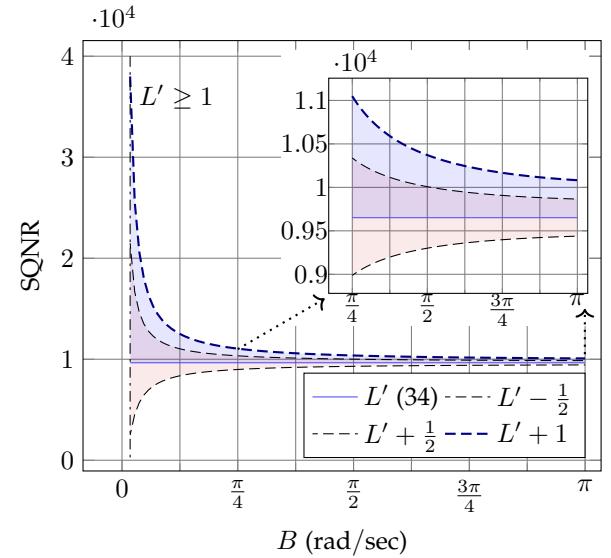


Fig. 11. Upper and lower bounds on SQNR due to using integer L'_i , instead of L' . Rounding of L' gives results in the red shaded area, while ceiling gives results in the blue shaded area. For this plot, $L = 32$ and $P_x = B$.

the stream length has to be an integer, denoted as L'_i , and assuming that L'_i is obtained by rounding L' to the nearest integer, it holds that

$$L' - \frac{1}{2} < L'_i < L' + \frac{1}{2}. \quad (39)$$

Hence, the SQNR achieved by an L'_i -bit stream, when L'_i is given by (31), over a conventional converter with a rounding quantizer, is bounded by

$$\frac{\pi}{2} \frac{3 \left(L \sqrt{\frac{2(B - \sin B)}{B}} - \frac{1}{2} \right)^2 P_x}{B - \sin B} < \text{SQNR}(L'_i) < \frac{\pi}{2} \frac{3 \left(L \sqrt{\frac{2(B - \sin B)}{B}} + \frac{1}{2} \right)^2 P_x}{B - \sin B}. \quad (40)$$

For maximum SQNR, instead of rounding L' , the integer L'_i can be selected as

$$L'_i = \left\lceil L \sqrt{\frac{2(B - \sin B)}{B}} \right\rceil. \quad (41)$$

Due to the ceiling function, $L' \leq L'_i < L' + 1$, and, hence,

$$\frac{\pi}{2} \frac{3 \left(L \sqrt{\frac{2(B - \sin B)}{B}} \right)^2 P_x}{B - \sin B} \leq \text{SQNR}(L'_i) < \frac{\pi}{2} \frac{3 \left(L \sqrt{\frac{2(B - \sin B)}{B}} + 1 \right)^2 P_x}{B - \sin B}. \quad (42)$$

Indicatively, Fig. 11 displays the SQNR that is derived from (40) and (42) for $L = 32$ and $P_x = B$. Furthermore, by requiring $L' \geq 1$, for $L = 32$, gives that $B \geq 0.054$ rad/sec. Similarly, SQNR, achieved by the integer bit-stream length L'_i , over a truncating quantizer for the conventional converter can be derived using (35) for L' . In this case, $L' \geq 1$

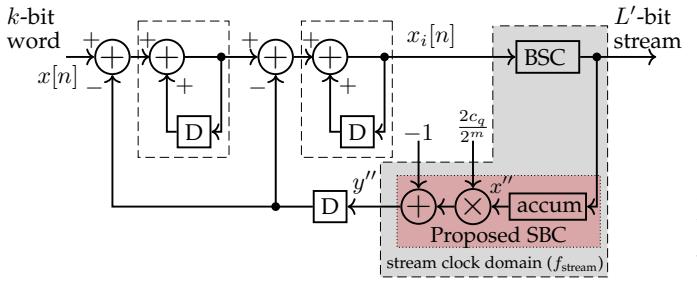


Fig. 12. Second-order CC-I architecture of proposed binary-to-stochastic converter.

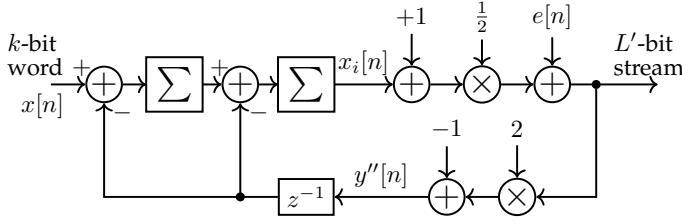


Fig. 13. Linear model for second-order noise-shaping system.

for $B \geq 0.550$ rad/sec. It is also observed that, as the signal bandwidth increases, estimated SQNR bounds become tighter.

5 PROPOSED SECOND-ORDER SYSTEM

5.1 Architecture of second-order noise-shaping bipolar converter

The architecture of the proposed second-order converter is presented in Fig. 12 with the CC in the loop. The second-order CC-O follows directly, resembling the first-order structure of Fig. 8. The second-order architecture comprises an additional integrator and subtractor, while the remainder structure is identical to the first-order architecture.

5.2 Derivation of noise power and SQNR

The noise transfer function $H_e''(z)$ for a second-order system with an additive noise source, as shown in Fig. 13, is

$$H_e''(z) = 2(1 - z^{-1})^2 = 2(1 + z^{-2} - 2z^{-1}). \quad (43)$$

We derive

$$\|H_e''(z = e^{j\omega})\|^2 = 4(1 + e^{-2j\omega} - 2e^{-j\omega}) (1 + e^{2j\omega} - 2e^{j\omega}) \quad (44)$$

$$= 8(3 + \cos(2\omega) - 4 \cos(\omega)). \quad (45)$$

From (10), (15) and (45), the in-band noise for the second-order system that uses a truncating quantizer, is derived as

$$P_{2B}'' = \frac{1}{2\pi} \int_{-B}^B \left(\frac{\pi\delta(\omega)}{2L^2} + \frac{1}{12L^2} \right) \quad (46)$$

$$= \frac{2(3 + \cos(2\omega) - 4 \cos(\omega))d\omega}{\pi 3L^2}. \quad (47)$$

The in-band noise power for a rounding quantizer, computed by (10), (12), and (45), leads also to (47). From (11)

and (47), the SQNR at the output the second-order bipolar converter, for both types of the quantizer, is

$$\text{SQNR}'' = \frac{\pi}{2} \frac{3L^2 P_x}{3B + \sin B(\cos B - 4)}. \quad (48)$$

5.3 Bit-stream length reduction of second-order converter

We seek the length, L' , of a bit-stream produced by the introduced second-order noise-shaping converter that has equal SQNR as that of an L -length bit-stream of a conventional converter. We distinguish two cases, over truncating and rounding quantizer for the conventional bipolar converter.

Lemma 3. A second-order converter with an L' -level quantizer, i.e., generating L' -length bit-streams, achieves the same SQNR as a conventional converter with an L -level rounding quantizer when

$$L' = L \sqrt{\frac{2(3B + \sin B(\cos B - 4))}{B}}, \quad (49)$$

where B is the bandwidth of the input signal.

Proof: For the case of an L -level rounding quantizer of a conventional converter, from (14) and (48), it is obtained that

$$\text{SQNR}_{\text{round, conv}} = \text{SQNR}'' \quad (50)$$

$$\frac{3\pi L^2 P_x}{B} = \frac{\pi}{2} \frac{3(L')^2 P_x}{3B + \sin B(\cos B - 4)} \Rightarrow \quad (51)$$

$$\frac{L'}{L} = \sqrt{\frac{2(3B + \sin B(\cos B - 4))}{B}}. \quad (52)$$

□

Lemma 4. A second-order converter with an L' -level quantizer, i.e., generating L' -length bit-streams, achieves the same SQNR as a conventional converter with an L -level truncating quantizer when

$$L' = L \sqrt{\frac{2(3B + \sin B(\cos B - 4))}{3\pi + B}}, \quad (53)$$

where B is the bandwidth of the input signal.

Proof: For the case of an L -level truncating quantizer of a conventional converter, from (17) and (48), it is obtained that

$$\text{SQNR}_{\text{trunc, conv}} = \text{SQNR}'' \quad (54)$$

$$\frac{\pi(L)^2 P_x}{3\pi + B} = \frac{\pi}{2} \frac{3(L')^2 P_x}{3B + \sin B(\cos B - 4)} \Rightarrow \quad (55)$$

$$\frac{L'}{L} = \sqrt{\frac{2(3B + \sin B(\cos B - 4))}{3\pi + B}}. \quad (56)$$

□

Fig. 10 depicts bit-stream length ratios computed by (52) and (56). The proposed second-order converter achieves a bit-stream length reduction when $B < B_2$, $B_2 = 1.616$ rad/sec, and $B < B_3$, $B_3 = 2.708$ rad/sec, over a conventional converter with a rounding and a truncating quantizer, respectively. Bit-stream length reductions are possible for a wider bandwidth when the first-order or

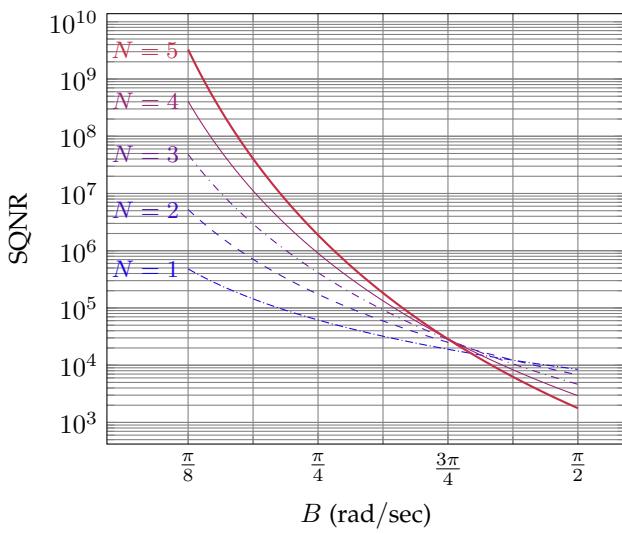


Fig. 14. SQNR at the output of proposed N th-order noise-shaping converter when $L' = 32$ and $P_x = 1$.

the second-order noise-shaping converter is compared to a conventional converter with a truncating quantizer.

A first-order converter leads also to stream reduction gains within a wider bandwidth than a second-order converter. However, by comparing the in-band noise power for the first-order and second-order noise-shaping converter, *i.e.*, (25) and (47), it follows that, when $B < B_4$, $B_4 = 1.378$ rad/sec, the second-order converter leaves less in-band noise power than the first-order system. Consequently, as shown in Fig. 10, bit-stream length reduction is greater when utilizing the second-order noise-shaping converter over the first-order converter for $B < B_4$.

5.4 General case of an N th-order system

The noise transfer function $H_e^N(z)$ for an N th-order system [31, Chapter 4] is

$$H_e^N(z) = 2(1 - z^{-1})^N. \quad (57)$$

Following (57), we derive

$$\begin{aligned} \|H_e^N(z = e^{j\omega})\|^2 &= 4 \left((1 - e^{j\omega})(1 - e^{-j\omega}) \right)^N \\ &= 2^{2N+2} \sin^2 N \left(\frac{\omega}{2} \right). \end{aligned} \quad (58)$$

From (10) and (58), the in-band noise power for an N th-order system P_{2B}^N that employs a quantizer with a noise spectral density $S(\omega)$ is

$$P_{2B}^N = \frac{1}{\pi} \int_{-B}^B S(\omega) 2^{2N+1} \sin^2 N \left(\frac{\omega}{2} \right) d\omega. \quad (59)$$

Hence, the SQNR achieved by an N th-order noise-shaping converter is derived by (11) and (59). Indicatively, in Fig. 14 the SQNR of a proposed N th-order noise-shaping converter is depicted for certain orders of the proposed converter. For low-bandwidth cases, as the order of the proposed converter increases, SQNR increases, while, for high bandwidths, SQNR increases with a decrease of the converter order. A discussion of stability in related high-order systems is offered by Pavan *et al.* in [31].

6 PROPOSED STOCHASTIC-TO-BINARY CONVERTER FOR REDUCED-LENGTH BIT-STREAMS

6.1 Architecture of proposed stochastic-to-binary converter

Typical implementations of SC systems process bit-streams with a length that is a power of two. In these cases, binary reconstruction according to (3), requires, from a hardware perspective, accumulation of stream-bits and a scaling operation by $\frac{1}{L}$. Since L is a power of two, scaling is realized by simply right-shifting by $\log_2 L$ positions the output of accumulator.

In this paper, the proposed noise-shaping technique leads to bit-stream length reduction and bit-stream lengths which are possibly not a power of two. Thus, the reconstruction to binary domain requires an actual scaling by $\frac{1}{L'}$ to be performed that burdens hardware complexity. Except for the binary reconstruction required in the feedback loop of CC-I, the binary reconstruction is necessary at the SC system output, as shown in Fig. 1. Following a straight-forward fixed-point implementation of (3) and an n -bit quantization of $\frac{1}{L'}$, *i.e.*, $\frac{1}{2^n} \lfloor 2^n \frac{1}{L'} \rfloor$, leads to a pure accuracy for the binary reconstruction. In order to improve representation efficiency, we propose the stochastic-to-binary conversion to be performed as

$$y' = 2 \frac{x'}{L'} - 1 = 2 \frac{x'}{2^m} \frac{2^m}{L'} - 1 = 2^{1-m} c x' - 1, \quad (60)$$

where y' is the reconstructed binary value in $[-1, 1]$, $x' = \sum_{i=1}^{L'} s'_i$ is the accumulator output as shown in Figs. 6 and 12, s'_i is the i th bit of the L' -length bit-stream and $c = \frac{2^m}{L'}$. The value of integer m is selected so as $1 \leq \frac{2^m}{L'} < 2$. For fixed-point implementations of the proposed architectures, c is quantized with n -bit precision taking the form of

$$c_q = \frac{1}{2^n} \lfloor 2^n c \rfloor = \frac{1}{2^n} \left\lfloor 2^{n+m} \frac{1}{L'} \right\rfloor. \quad (61)$$

Hereafter, for the investigated fixed-point implementations of the proposed binary-to-stochastic converter, we employ a stochastic-to-binary conversion based on (60) and (61). In order to keep the hardware overhead minimal, the term $2^{1-m} c_q x'$ in Figs. 6 and 12 is implemented as the addition of shifted versions of x' , computed by the proposed stochastic-to-binary converter. As the feed-forward path employs k -bit additions, the reconstructed binary value $y'[n]$ of Figs. 6 and 12 is restricted to k bits before it is subtracted from $x[n]$.

6.2 Experimental evaluation of proposed SBC

We assess the SQNR at the output of the proposed first-order binary-to-stochastic converter for certain implementations of the stochastic-to-binary converter. Fig. 15 depicts the SQNR when the stochastic-to-binary converter relies on a floating-point implementation and two fixed-point implementations of stochastic-to-binary conversion. The fixed-point implementations are based either on (3) and a fixed-point value of $\frac{1}{L'}$, or on (60) and (61). Furthermore, it holds that $k = 15$, $m = 7$ and certain values of n are examined for $L' = 100$. It is observed that the simulation-based SQNR with a reconstruction that relies on (60) and (61) is in close agreement with the model SQNR for $n \geq 8$. However, the

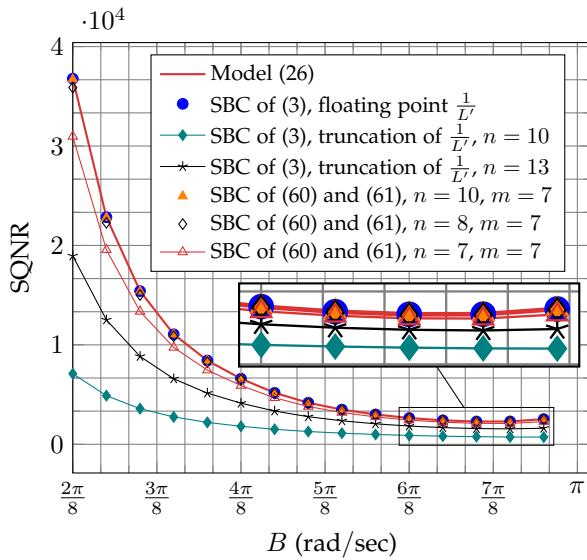


Fig. 15. Simulation-based and model SQNR for a first-order noise-shaping converter and certain implementations of stochastic-to-binary converter. A random number sequence is used as an input. For this plot, $L' = 100$.

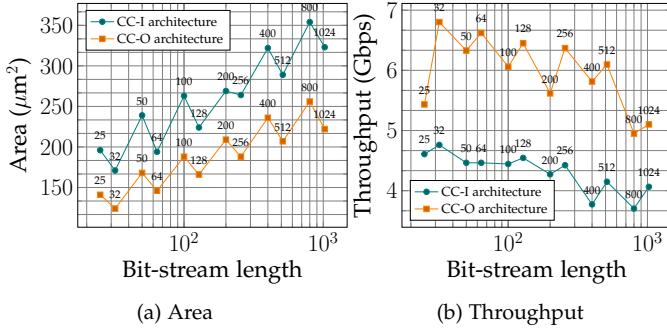


Fig. 16. Area and throughput comparison of the proposed first-order CC-I and CC-O architectures at a 28-nm FDSOI technology.

simulation-based SQNR with a reconstruction that relies on (3) and $\frac{1}{2^{10}} \lfloor 2^{10} \frac{1}{L} \rfloor$ deviates significantly from the model SQNR. Although a reconstruction that relies on (3) and $\frac{1}{2^{13}} \lfloor 2^{13} \frac{1}{L} \rfloor$ increases significantly the SQNR compared to the previous case, it also presents substantial deviation from the model SQNR. For high bandwidth cases, it is observed that SQNR error decreases.

7 HARDWARE COMPLEXITY

The investigated binary-to-stochastic converters are designed and synthesized with Cadence Genus at a 28-nm FDSOI technology. Serial implementations are assumed for all the investigated converters, which use one RNG per converter. Fixed-point operations are performed with one integral and $k - 1$ fractional bits for the integrator and the subtractor. Here, it is assumed that $k = \lceil \log_2 L \rceil + 3$. In all cases, synthesis targets the maximum performance.

Synthesis results regarding CC-I and CC-O for the proposed first-order noise-shaping conversion technique are reported subsequently. Specifically, Fig. 16 presents the estimated maximum throughput rate and area for the two

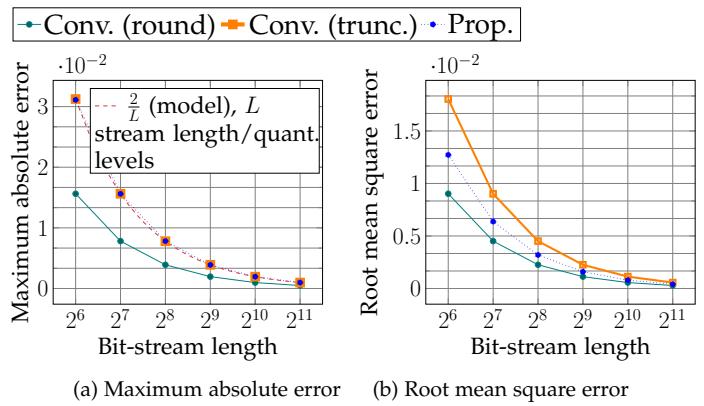


Fig. 17. Simulation-based error at the output of conventional and proposed noise-shaping converters for a chirp input signal.

architectures. Points are annotated by the bit-stream length. It is shown that CC-O outperforms CC-I in terms of maximum achievable throughput rate, and occupies the least area for all the considered stream lengths. Furthermore, it is shown that the area of both architectures is minimized for a constant k when bits-stream length is a power of two, *i.e.*, an implementation that corresponds to $L' = 128$ has less hardware cost than that of $L' = 100$, where both use a fixed-point precision of $k = 10$ bits. This happens as the cost for binary reconstruction on the feedback loop of both architectures is significantly simplified.

Table 2 also reports the occupied area and the estimation of power dissipation by the synthesis tool for the investigated converters. Here, both first-order and second-order noise-shaping structures are investigated, relying CC-I. The designed converters employ a k -bit fixed-point precision and an L -level truncating quantizer. Furthermore, the fixed-point scaling in the modified SBC on the feedback path is performed with k -bit precision. Table 2 also shows the performance, area, and power metrics of conventional BSCs when the constituent RNGs rely either on counters or on primitive LFSRs [16]. The maximum throughput and the corresponding area requirements of a $\Sigma\Delta$ modulator are also quantified, using k -bit precision and an 1-bit quantizer. The proposed first-order and second-order converters show an average $2.548\times$ and $3.548\times$ area increase with respect to a conventional converter that generates streams of equal length, while power increases by $1.271\times$ and $1.223\times$, respectively.

8 EXPERIMENTAL EVALUATION

This section provides fixed-point simulations on the bit-stream length reduction and SQNR benefits, derived in Sections 3 to 5, for the proposed converters.

8.1 Conversion error

We initially estimate the error between the input and output introduced by the conventional and proposed first-order binary-to-stochastic converters for certain bit-stream lengths. In the examined scenario, samples of a chirp signal are converted into bit-streams of certain length by the investigated converters. As an RNG of the CC, a structure that

TABLE 2
Hardware Complexity and Performance at a 28-nm FDSOI Technology¹

L	k	Throughput (Gbps)				Area (μm^2)				Total Power (μW) ⁵			
		Conv. ²	1 st -order ³	2 nd -order ³	$\Sigma\Delta$ ⁴	Conv.	1 st -order	2 nd -order	$\Sigma\Delta$	Conv.	1 st -order	2 nd -order	$\Sigma\Delta$
32	8	6.896-7.575	4.761	4.761	5.128	44.390-38.406	171.686	216.838	33.730	244-205	580	575	624
64	9	6.451-6.944	4.464	4.504	4.901	53.094-47.437	194.861	256.224	87.802	246-215	628	643	676
128	10	6.535-6.896	4.545	4.464	5.000	64.410-56.576	224.128	294.086	104.122	363-241	720	728	772
256	11	6.451-6.711	4.424	4.219	4.854	77.357-63.866	264.058	331.840	121.530	362-262	843	791	829
512	12	6.172-6.667	4.149	3.937	4.716	85.843-70.067	289.626	368.614	116.307	391-289	875	806	834
1024	13	5.102-6.578	4.065	3.906	4.587	92.698-76.922	323.245	410.938	125.882	439-317	944	930	872

¹Synthesis targets maximum performance. Typical delay corner and nominal supply voltage (1 V). ²Conventional BSC with a counter-based and LFSR-based RNGs reported in (counter-based)-(LFSR-based) format [16]. ³Proposed CC-I noise-shaping converter. ⁴ $\Sigma\Delta$ modulator with 1-bit quantizer, shown in Fig. 4. ⁵Synthesis estimation.

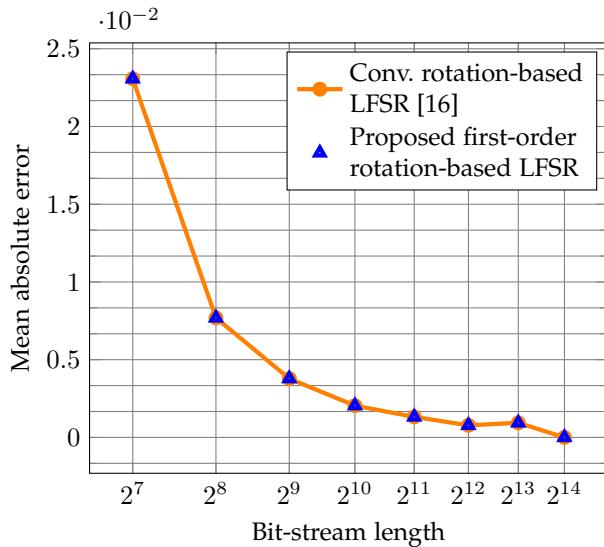


Fig. 18. Mean absolute error when multiplying two bipolar stochastic bit-streams produced by the first-order and conventional binary-to-stochastic converters.

produces every number in $[0, L - 1]$ only once is employed. Specifically, the simulation-based maximum absolute and the root mean square errors of the investigated converters are displayed in Fig. 17 for power-of-two bit-stream lengths. The error that the examined converters present is solely attributed in the quantization procedure; the proposed and conventional converter with a truncating quantizer present a maximum absolute error equal to $\frac{2}{L}$, which is twice the step of quantizer, shown with a dashed red line in Fig. 17(a), while the maximum absolute error is $\frac{1}{L}$ for the conventional converter with a rounding quantizer. For the conventional converter, the maximum error is derived by the difference $y[n] - x[n]$ exploiting (4). Furthermore, the root mean square error of the proposed converter lies between that of the conventional converter with a rounding and a truncating quantizer.

Regarding the randomization of bits inside a stream encoding a binary sample, the distribution of ones and zeros is determined only by the randomization properties of the RNG of CC. Thus, randomization properties of a stream produced by the proposed noise-shaping converter are the same as that of a conventional BSC. In order to show that the randomization of bits of the introduced binary-to-stochastic

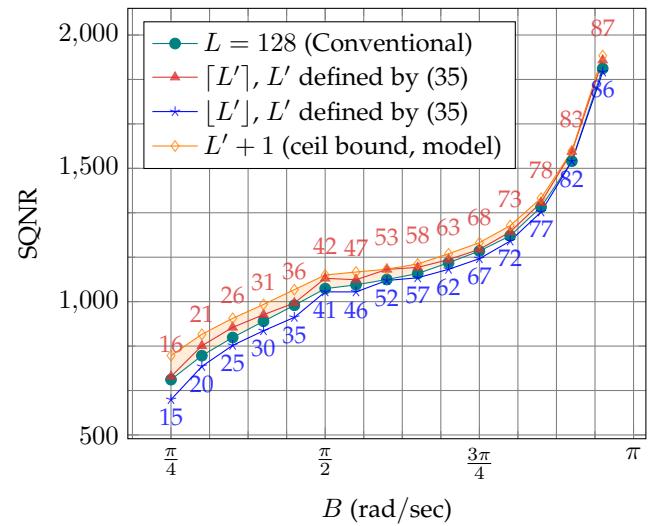


Fig. 19. Simulation-based SQNR of first-order converter and conventional bipolar converter. A random number sequence is used as an input. Fixed-point scaling factor with 10-bit precision for the proposed SBC on feedback loop.

converter does not impose any restriction on the accuracy of the following SC operations, we evaluate the mean absolute error in the case of multiplication. Specifically, we multiply all 7-bit precision numbers in the stochastic domain when using the introduced converters and conventional LFSR-based converters. We also employ the technique of rotation of one stream with respect to the other to surpass correlation issues [16]. Specifically, Fig. 18 displays the mean absolute error for certain stream lengths. It is shown that the full precision of multiplication is maintained in case of a 2^{14} -bit stream length, as expected. Furthermore, for smaller stream bit-stream lengths, the introduced and conventional converters achieve the same mean absolute error.

8.2 Experimental evaluation for the proposed BSCs

We assume that a truncating quantizer of seven bits is used by all converters. Hence, the conventional converter generates a bit-stream of $L = 128$ bits per input binary word. The proposed noise-shaping converters produce bit-streams of reduced length L' , derived from (35) and (53), for the first-order and second-order systems, respectively. The SQNR is evaluated by Matlab simulations and is depicted

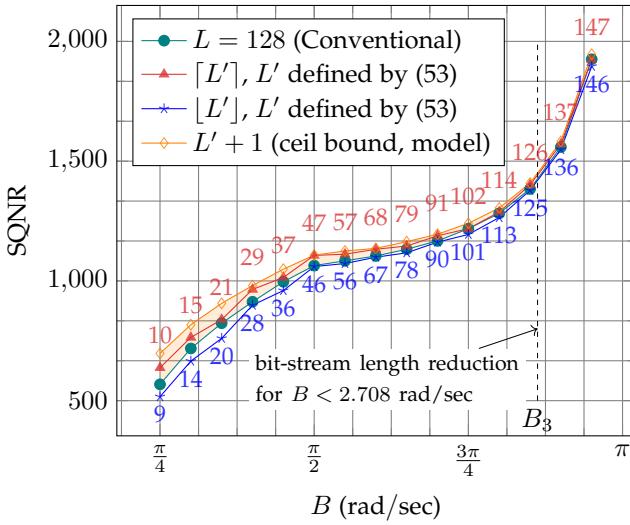


Fig. 20. Simulation-based SQNR of second-order converter and conventional bipolar converter. A random number sequence is used as an input. Fixed-point scaling factor with 10-bit precision for the proposed SBC on feedback loop.

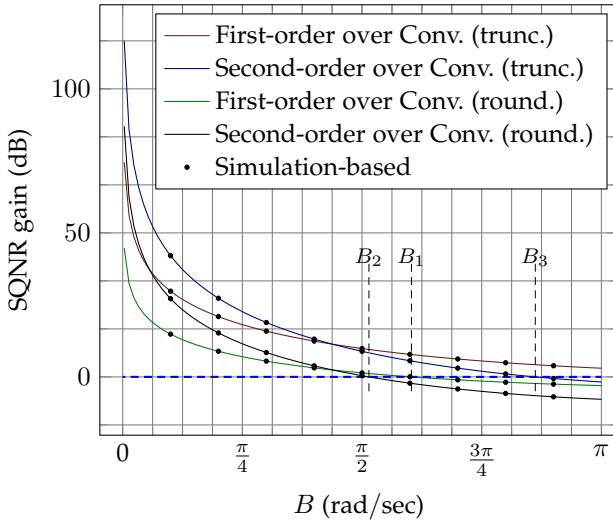


Fig. 21. SQNR gain in dB for a stochastic bit-stream due to noise shaping, as a function of the bandwidth. Gains are evaluated with respect to a conventional converter with a rounding or a truncating quantizer and the same bit-stream length as the proposed noise-shaping one.

TABLE 3

SQNR Gain due to Noise Shaping over Conventional Converters as a Function of Bandwidth

Proposed		Conventional with Truncating Quantizer	
	Achieved SQNR Gain	Range for Gain ≥ 1 (rad/sec)	
First-order	$\frac{3\pi+B}{2(B-\sin B)}$	$[0, \pi]$	
Second-order	$\frac{3\pi+B}{2(3B+\sin B(\cos B-4))}$	$[0, 2.708]$	
Proposed		Conventional with Rounding Quantizer	
	Achieved SQNR Gain	Range for Gain ≥ 1 (rad/sec)	
First-order	$\frac{B}{2(B-\sin B)}$	$[0, 1.895]$	
Second-order	$\frac{B}{2(3B+\sin B(\cos B-4))}$	$[0, 1.616]$	

TABLE 4
Probability Density Function of Biased RNG

Random number	0.125	0.375	0.625	0.875
Probability	0.1	0.1	0.4	0.4

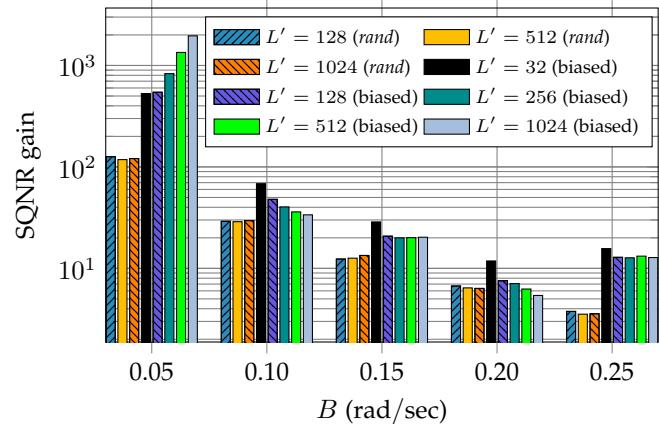


Fig. 22. SQNR gain of CC-I over CC-O for a sine wave input at B rad/sec. Matlab's `rand` function and a biased RNG are utilized.

in Figs. 19 and 20. For each bandwidth point, two cases are simulated, corresponding to reduced lengths $[L']$ and $[L']$, also shown in Figs. 19 and 20.

The obtained values are found in close agreement with the analytical models, and lie inside the ceiling bound estimated by the SQNR model for $L' + 1$, provided by (42) for the first-order converter. Furthermore, with an increase of bandwidth in Figs. 19 and 20, the SQNR distance for $[L']/[L']$ and the L decreases, as qualitatively also shown by the SQNR model values of Fig. 11. In essence, bit-stream length reduction increases as the bandwidth of input signal decreases. Indicatively, for a signal bandwidth of $\pi/4$ rad/sec, a bit-stream length reduction of 87.5% and 92.18% is offered when the first-order and second-order converter are utilized, respectively. It is noted that the derivation of (35) and (53) does not consider the quantization of factor c in the proposed SBCs. When the factor c is quantized to a sufficient precision, the scaling operation does not impose notable impact. Furthermore, as signal bandwidth increases, the margin between the SQNR achieved by the proposed converters and a conventional converter reduces, as indicated also by SQNR models in Fig. 11.

We also evaluate the SQNR performance of the proposed binary-to-stochastic converters for fixed-point implementations. Specifically, we evaluate SQNR gains when the introduced converters produce equal-length bit-streams as a conventional bipolar converter. We assume that an L -level quantizer is included in all conversion structures. The fixed-point scaling factor in the proposed stochastic-to-binary converter is implemented with a precision of 10 bits. SQNR gains are assessed by the ratio of SQNR of the introduced over the SQNR of the conventional converter, as described in Table 3, and depend only on signal bandwidth. Fig. 21 depicts the simulation-based SQNR gains for $L = 32$ and the model ones. As signal bandwidth increases, the achieved

gains reduce. Indicatively, the introduced first-order and second-order converters offer 29.8 dB and 42.1 dB SQNR gains compared to a conventional converter with a truncating quantizer and equal-length streams for 0.1 π rad/sec, respectively.

8.3 Comparison of CC-I and CC-O architectures

A perfectly uniform RNG does not introduce any additional noise in the CC other than the quantization noise. For the same stream length and a perfectly uniform RNG, the proposed CC-I and CC-O architectures demonstrate the same SQNR at the converter's output, given by (26) and (48), for a first-order and a second-order system, respectively. In contrast, a non-perfectly-uniform RNG introduces additional noise, related to the mapping of the quantized word a_q of Fig. 2 to a bit-stream. The CC-I architecture shapes also this noise, as it is introduced inside the noise-shaping loop. The particular advantage compensates the additional hardware cost of CC-I over CC-O, quantified in Fig. 16. The specific characteristics of the non-perfectly-uniform RNG impact on the output SQNR. Certain cases are investigated below that highlight the advantage of CC-I over CC-O. The SQNR at the output of the proposed converters in the presence of a non-perfectly-uniform RNG depends also on the noise spectrum of the RNG. Deriving closed-form expressions for the SQNR in this case is an open problem, not addressed in this paper.

We assume random numbers in [0, 1]. In the first scenario, the *rand* function of Matlab is exploited, where the probability density function derived from a random number sequence deviates from a uniform distribution for a finite sequence length. In the second scenario, a biased RNG is exploited, which favors the generation of numbers greater than 0.5. Table 4 describes the probability density function of the employed biased RNG. The analysis focuses on a sine wave input signal at B rad/sec and equal bit-stream lengths. The related SQNR gains, evaluated at the output of CC-I and CC-O architectures, are depicted in Fig. 22, for the two RNGs and certain bit-stream lengths, L' . The noise of the particular RNGs surpasses any shaped quantization noise, initially observed at a_q of Fig. 8. Therefore, the spectrum at the output of the CC-O architecture is flat, resembling that of a conventional BSC in Fig. 2. The SQNR benefits of CC-I architecture is greater in low spectrum, as in case of a perfectly uniform RNG.

8.4 Evaluation in stochastic digital filtering

SC low-precision digital filters have gained extensive interest in recent years owing to low hardware cost requirements. They can be also designed in conjunction with $\Sigma\Delta$ principles to shape the quantization noise. In [24], Sotiriadis and Temenos propose such an architecture, where one-bit length streams are utilized and a binary adder is exploited to countermeasure the inherent scaling of SC adders, deriving the output in binary representation. The filter architecture by Yuan and Wang in [44], while surpasses the accuracy degradation problem at the output of filter taps by more elaborate SC adders, it exploits a two-stream representation. The introduced converters, as they do not impose any restriction on the subsequent operations, can be potentially

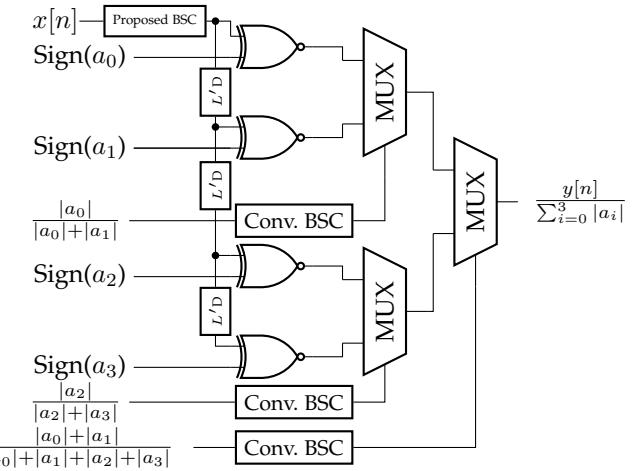


Fig. 23. A fourth-order finite impulse response stochastic filter with uneven-weighted adders.

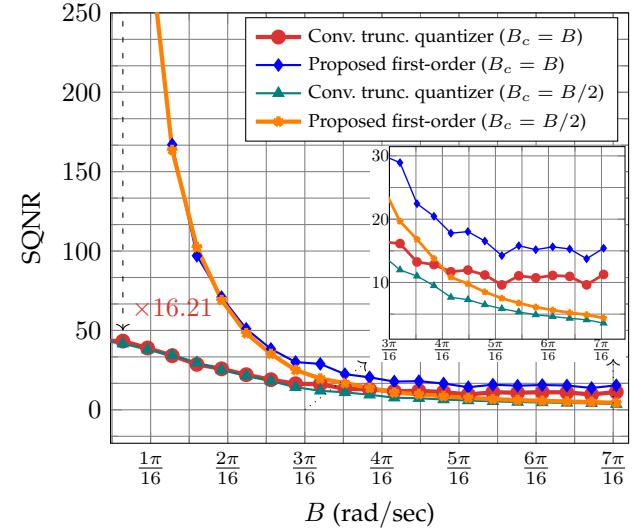


Fig. 24. SQNR at the output of an eighth-order stochastic low-pass filter with a cut-off frequency B_c . For this plot, a random number sequence is used as an input and $L' = 32$.

TABLE 5
SQNR at the Output of a Low-Pass Fourth-Order Filter with a Cut-Off Frequency at B_c

B	SQNR					
	$B_c = B/2$			$B_c = B$		
	Conv.	1 st -order*	Gain	Conv.	1 st -order*	Gain
0.05	44.7	1808.9	40.4	45.7	2106.7	46.0
0.10	42.7	527.7	12.3	42.5	527.0	12.4
0.15	38.8	231.1	6.0	37.1	201.1	5.4
0.20	34.3	127.0	3.7	34.2	127.8	3.7
0.25	29.6	78.7	2.7	29.7	80.0	2.7
0.30	24.8	52.8	2.1	22.5	47.9	2.1
0.35	20.9	37.6	1.8	21.7	39.5	2.1
0.40	17.6	27.9	1.6	18.6	30.0	1.6
0.45	14.8	21.5	1.4	16.0	23.6	1.5

* Proposed noise-shaping converter

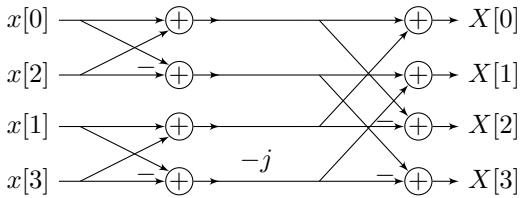


Fig. 25. Four-point FFT composed of radix-2 butterflies.

adopted in filter designs in [24], [44]. Here, we adopt a filter architecture that relies on uneven weighted adders on the technique by Liu and Parhi (cf. [4, Fig. 6(a)]), due to its lower hardware complexity, with an L' -bit stream, and the output in the bipolar representation. Indicatively, Fig. 23 displays the architecture of a fourth-order finite impulse response stochastic filter, *i.e.*, $y[n] = \sum_{i=0}^3 a_{ix}[n-i]$. Specifically, a proposed first-order binary-to-stochastic converter generates the input stochastic bit-stream that refers to $x[n]$. A delayed sample of $x[n], x[n-k]$, results by placing kL' delay elements at the output of the introduced binary-to-stochastic converter. Furthermore, the stochastic bit-streams of select signals for the adders are generated by conventional binary-to-stochastic converters. Given the cut-off frequency and the filter order, the filter coefficients are estimated by *fir1* function of Matlab [45]. In the filter architecture, all converters produce L' -length bit-streams per input binary sample.

The analysis considers a fourth and an eighth-order low-pass stochastic filter with a cut-off frequency of B_c rad/sec. The eighth-order structure is a straight-forward extension of the architecture in Fig. 23. Fig. 24 displays the SQNR for equal-length bit-streams of input $x[n]$ generated by either the introduced first-order or a conventional converter for certain cut-off frequencies; both converters generate 32-bit streams. The SQNR gains for the fourth-order SC filter are displayed in Table 5, and are as high as $\times 46$ for low bandwidths. It is shown that the benefits of the proposed converters compared to the conventional one are preserved at the output of SC digital filters and they are more pronounced at low bandwidths, reaching a $\times 56.89$ SQNR gain for $B=0.02\pi$ for the eighth-order filter. The particular observation is important. A target SQNR at a filter output can be achieved with a reduced bit-stream length when the proposed converters are utilized. This further leads to latency and power reduction for bit-serial implementations or area reduction for bit-parallel implementations of SC systems.

8.5 Evaluation in stochastic FFT

The case of a four-point radix-2 FFT is initially studied, the signal flow graph of which is shown in Fig. 25. A straightforward stochastic architecture is derived by directly mapping all operations to their stochastic equivalents. Inputs to the FFT are consecutive non-overlapping sets of four consecutive data samples, *i.e.*, the $(4i)$ th, $(4i+1)$ th, $(4i+2)$ th, and $(4i+3)$ th samples. We compare conventional and first-order noise-shaping converters producing stochastic streams of equal length at the FFT inputs. A converter per input is used, so as the shaping process can be applied on each input

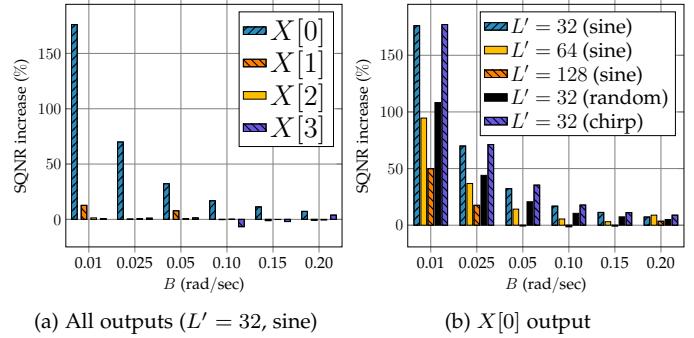


Fig. 26. SQNR increase at the outputs of four-point radix-2 FFT.

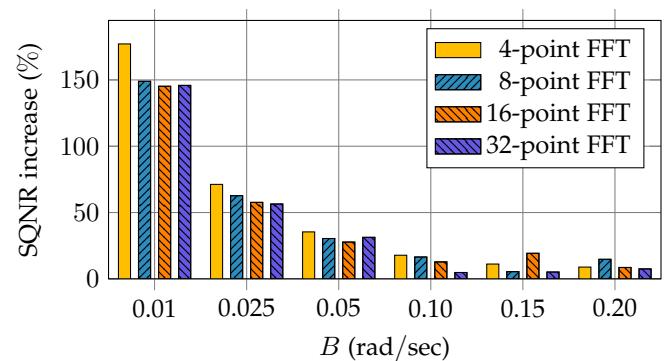


Fig. 27. SQNR increase at the $X[0]$ output of certain radix-2 FFTs for a chirp input signal and $L' = 32$.

separately. Exploiting a single noise-shaping converter for the data sequence and subsequently selecting the streams that correspond to the $(4i)$ th, $(4i+1)$ th, $(4i+2)$ th, and $(4i+3)$ th samples does not produce a noise-shaped signal for each FFT input. Fig. 26(a) depicts the SQNR percentage increase, observed at the outputs of the FFT, when using a proposed first-order converter with respect to a conventional one. The proposed converters achieve substantial SQNR benefits owing to the noise-shaping process, for the $X[0]$ output. The combinational path to $X[0]$ involves in-series addition operations and the noise-shaping benefits are propagated at the particular output. The remainder of the combinational paths involve also subtraction operations, that degrade the SQNR in case of both converters, and the SQNR increase at the outputs is almost negligible. We further investigate the case of $X[0]$ output and estimate the SQNR benefits for certain stream lengths L' . A chirp signal and a random input sequence with a maximum bandwidth at B rad/sec are also tested. Fig. 26(b) presents the relevant metric, revealing significant SQNR benefits that increase as the bandwidth of input signal decreases. We extend the analysis to eight-point, sixteen-point, and thirty-two-point radix-2 FFTs and evaluate the SQNR increase at the $X[0]$ output when using a conventional and a noise-shaping converter. The particular output refers to the sum of all input sample sequences for $L'=32$, computed by a binary addition tree. Specifically, Fig. 27 demonstrates the relevant SQNR increase. It is observed that the SQNR benefits for FFTs with a greater complexity are in the same scale as that of a four-point FFT, experimentally proving the consistency of the benefits by the introduced noise-shaping technique as

complexity increases.

9 CONCLUSIONS

This paper introduces novel architectures for binary-to-stochastic conversion that rely on noise-shaping principles. The introduced converters reduce bit-stream length for a target SQNR compared to conventional binary-to-stochastic converters by effectively shaping the in-band quantization noise. We also provide closed formulas for the SQNR at the reconstructed output of the investigated converters, as well as the required bit-stream length generated by the proposed converters for a target SQNR. Among the broad spectrum of high-order and more involved $\Sigma\Delta$ modulators on which the proposed conversion can be based, two block-level architectures are proposed based on plain $\Sigma\Delta$ structures that implement the proposed noise-shaping conversion technique. The analysis for fixed-point implementations of proposed converters reveal significant bit-stream length reduction benefits compared to conventional structures. Furthermore, we introduce a stochastic-to-binary converter for the reconstruction of truncated bit-streams to binary numbers, required on the feedback path of the proposed binary-to-stochastic converters. Synthesis results of the investigated fixed-point architectures are provided at a 28-nm FDSOI technology. A quantitative analysis also reveals that the SQNR gains, offered by the introduced converters, can be observed at the output of a SC system, such as a low-pass filter or an FFT structure.

REFERENCES

- [1] M. Koo, G. Srinivasan, Y. Shim, and K. Roy, "SBSNN: Stochastic-Bits Enabled Binary Spiking Neural Network with On-Chip Learning for Energy Efficient Neuromorphic Computing at the Edge," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 8, pp. 2546–2555, 2020.
- [2] A. Ardkani, F. Leduc-Primeau, N. Onizawa, T. Hanyu, and W. J. Gross, "VLSI Implementation of Deep Neural Network Using Integral Stochastic Computing," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 10, pp. 2688–2699, 2017.
- [3] N. Temenos and P. P. Sotiriadis, "A Stochastic Computing Sigma-Delta Adder Architecture for Efficient Neural Network Design," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2023.
- [4] Y. Liu and K. K. Parhi, "Architectures for Recursive Digital Filters Using Stochastic Computing," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3705–3718, 2016.
- [5] B. Moons and M. Verhelst, "Energy-Efficiency and Accuracy of Stochastic Computing Circuits in Emerging Technologies," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 4, no. 4, pp. 475–486, 2014.
- [6] P. Li, D. J. Lilja, W. Qian, K. Bazargan, and M. D. Riedel, "Computation on Stochastic Bit Streams Digital Image Processing Case Studies," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 3, pp. 449–462, 2013.
- [7] A. Naderi, S. Mannor, M. Sawan, and W. J. Gross, "Delayed Stochastic Decoding of LDPC Codes," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5617–5626, 2011.
- [8] J. Chen, J. Hu, and J. Zhou, "Hardware and Energy-Efficient Stochastic LU Decomposition Scheme for MIMO Receivers," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 24, no. 4, pp. 1391–1401, 2015.
- [9] K. Han, J. Hu, J. Chen, and H. Lu, "A Low Complexity Sparse Code Multiple Access Detector Based on Stochastic Computing," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 2, pp. 769–782, 2017.
- [10] A. Alaghi, W.-T. J. Chan, J. P. Hayes, A. B. Kahng, and J. Li, "Trading Accuracy for Energy in Stochastic Circuit Design," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 13, no. 3, pp. 1–30, 2017.
- [11] Z. Zhang, R. Wang, Z. Zhang, Y. Zhang, S. Guo, and R. Huang, "Circuit Reliability Comparison Between Stochastic Computing and Binary Computing," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 12, pp. 3342–3346, 2020.
- [12] C. Andriakopoulos, K. Papachatzopoulos, and V. Palouras, "A Novel Stochastic Polar Architecture for All-Digital Transmission," in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2021, pp. 1–5.
- [13] S. Liu and J. Han, "Toward Energy-Efficient Stochastic Circuits Using Parallel Sobol Sequences," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 7, pp. 1326–1339, 2018.
- [14] M. H. Najafi, D. Jenson, D. J. Lilja, and M. D. Riedel, "Performing Stochastic Computation Deterministically," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 27, no. 12, pp. 2925–2938, 2019.
- [15] Z. Lin, G. Xie, W. Xu, J. Han, and Y. Zhang, "Accelerating Stochastic Computing Using Deterministic Halton Sequences," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 10, pp. 3351–3355, 2021.
- [16] M. H. Najafi and D. Lilja, "High Quality Down-Sampling for Deterministic Approaches to Stochastic Computing," *IEEE Transactions on Emerging Topics in Computing*, 2018.
- [17] S. A. Salehi, "Low-Cost Stochastic Number Generators for Stochastic Computing," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 28, no. 4, pp. 992–1001, 2020.
- [18] V. Sehwag, N. Prasad, and I. Chakrabarti, "A Parallel Stochastic Number Generator with Bit Permutation Networks," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 2, pp. 231–235, 2017.
- [19] S. Liu and J. Han, "Dynamic Stochastic Computing for Digital Signal Processing Applications," in *2020 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2020, pp. 604–609.
- [20] S. Liu, W. J. Gross, and J. Han, "Introduction to Dynamic Stochastic Computing," *IEEE Circuits and Systems Magazine*, vol. 20, no. 3, pp. 19–33, 2020.
- [21] P. W. Wong, "Fully Sigma-Delta Modulation Encoded FIR Filters," *IEEE Transactions on Signal Processing*, vol. 40, no. 6, pp. 1605–1610, 1992.
- [22] D. A. Johns and D. M. Lewis, "Design and Analysis of Delta-Sigma Based IIR Filters," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 40, no. 4, pp. 233–240, 1993.
- [23] N. Saraf, K. Bazargan, D. J. Lilja, and M. D. Riedel, "IIR Filters Using Stochastic Arithmetic," in *2014 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2014, pp. 1–6.
- [24] N. Temenos, A. Vlachos, and P. P. Sotiriadis, "Efficient stochastic computing fir filtering using sigma-delta modulated signals," *Technologies*, vol. 10, no. 1, p. 14, 2022.
- [25] P. Gonzalez-Guerrero, S. G. Wilson, and M. R. Stan, "Error-Latency Trade-Off for Asynchronous Stochastic Computing with $\Sigma\Delta$ Streams for the IoT," in *2019 32nd IEEE International System-on-Chip Conference (SOCC)*. IEEE, 2019, pp. 97–102.
- [26] P. Gonzalez-Guerrero, X. Guo, and M. Stan, "SC-SD: Towards Low Power Stochastic Computing using Sigma Delta Streams," in *2018 IEEE International Conference on Rebooting Computing (ICRC)*. IEEE, 2018, pp. 1–8.
- [27] M. José, "Sigma-Delta Modulators: Tutorial Overview, Design Guide, and State-Of-The-Art Survey," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 1, pp. 1–21, 2010.
- [28] A. Frappé, A. Flament, B. Stefanelli, A. Kaiser, and A. Cathelin, "An All-Digital RF Signal Generator Using High-Speed $\Delta\Sigma$ Modulators," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 10, pp. 2722–2732, 2009.
- [29] S. Hein and A. Zakhori, "On the Stability of Sigma Delta Modulators," *IEEE Transactions on Signal Processing*, vol. 41, no. 7, pp. 2322–2348, 1993.
- [30] P. Madoglio, A. Ravi, L. Cuellar, S. Pellerano, P. Seddighrad, I. Lomeli, and Y. Palaskas, "A 2.5-GHz, 6.9-mW, 45-nm-LP CMOS, $\Delta\Sigma$ Modulator Based on Standard Cell Design With Time-Interleaving," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 7, pp. 1410–1420, 2010.
- [31] S. Pavan, R. Schreier, and G. C. Temes, *Understanding Delta-Sigma Data Converters*. John Wiley & Sons, 2017.

- [32] K. Papachatzopoulos, C. Andriakopoulos, and V. Palioras, "Novel Noise-Shaping Stochastic-Computing Converters for Digital Filtering," in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2020.
- [33] A. V. Oppenheim, *Discrete-Time Signal Processing*. Pearson Education India, 1999.
- [34] Y. Zhang, R. Wang, X. Zhang, Y. Wang, and R. Huang, "Parallel Hybrid Stochastic-Binary-Based Neural Network Accelerators," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 12, pp. 3387–3391, 2020.
- [35] K. K. Parhi and Y. Liu, "Computing Arithmetic Functions Using Stochastic Logic by Series Expansion," *IEEE Transactions on Emerging Topics in Computing*, vol. 7, no. 1, pp. 44–59, 2016.
- [36] Y. Liu and K. K. Parhi, "Computing Complex Functions using Factorization in Unipolar Stochastic Logic," in *2016 International Great Lakes Symposium on VLSI (GLSVLSI)*. IEEE, 2016, pp. 109–112.
- [37] A. Alaghi, W. Qian, and J. P. Hayes, "The Promise and Challenge of Stochastic Computing," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 8, pp. 1515–1531, 2017.
- [38] W. Qian, X. Li, M. D. Riedel, K. Bazargan, and D. J. Lilja, "An Architecture for Fault-Tolerant Computation with Stochastic Logic," *IEEE Transactions on Computers*, vol. 60, no. 1, pp. 93–105, 2010.
- [39] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Englewood Cliffs: Prentice-Hall, 1975.
- [40] X. Chen, B. Li, Y. Wang, Y. Liu, and H. Yang, "A Unified Methodology for Designing Hardware Random Number Generators Based on Any Probability Distribution," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 8, pp. 783–787, 2016.
- [41] R. Govindaraj, S. Ghosh, and S. Katkoori, "CSRO-Based Reconfigurable True Random Number Generator Using RRAM," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 12, pp. 2661–2670, 2018.
- [42] S. H. Ardalan and J. J. Paulos, "Stability Analysis of High-Order Sigma-Delta Modulators," North Carolina State University. Center for Communications and Signal Processing, Tech. Rep., 1986.
- [43] P. Steiner and W. Yang, "A Framework for Analysis of High-Order Sigma-Delta Modulators," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44, no. 1, pp. 1–10, 1997.
- [44] B. Yuan and Y. Wang, "High-Accuracy FIR Filter Design using Stochastic Computing," in *2016 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*. IEEE, 2016, pp. 128–133.
- [45] *Window-based FIR filter design*, accessed: Mar, 2022. [Online]. Available: <https://www.mathworks.com/help/signal/ref/fir1.html>



Kleanthis Papachatzopoulos (S'16) received the Diploma degree in Electrical and Computer Engineering, and the M.Sc. degree in Integrated Hardware-Software Systems from the University of Patras, Greece, in 2016 and 2018, respectively.

Currently, he is pursuing a PhD degree and he is working as a Research Assistant with the VLSI Design Laboratory, ECE Dept., University of Patras, Patras, Greece. His current research interests include VLSI architectures for signal processing and computer arithmetic.



Vassilis Palioras (Member, IEEE) is a Full Professor with the Electrical and Computer Engineering Department, University of Patras, Greece. His research interests are in the areas of VLSI architectures for machine learning, signal processing and communications, low-power systems and computer arithmetic. He is advisor to five Ph.D. students, and has supervised four Ph.D., 36 masters', and 40 diploma theses. Prof. Palioras has received the IEEE CASS Guillemin—Cauer Best-Paper Award for the year 2000. He has served as the General Co-Chair for International Workshop on Power and Timing Modeling, Optimization and Simulation (PATMOS) 2004. He has also served as a Technical Program Chair of PATMOS 2005, the IEEE Workshop on Signal Processing Systems Implementation (SIPS) 2005, and Technical Program Co-Chair of the IEEE International Conference on Electronics Circuits and Systems (ICECS) 2010 and a European liaison for the IEEE ISCAS 2012, South Korea.