A Simple and Efficient Dithering Method for Vector Quantizer Based Mismatch-Shaped $\Delta\Sigma$ DACs

Arindam Sanyal and Nan Sun
Department of Electrical and Computer Engineering
University of Texas at Austin
Austin, TX 78712, USA
Email: arindam3110@utexas.edu, nansun@mail.utexas.edu

***Abstract*—This paper presents an in-depth analysis of the generation of tones in the output spectra of vector-quantizer (VQ) based multibit mismatch-shaped $\Delta\Sigma$ digital-to-analog converters (DACs). Building upon the analysis, a simple yet elegant method of adding dither to remove tones from the output spectra is presented. It achieves a better mismatch shaping performance with a low hardware cost compared to existing dithering techniques.**

I. Introduction

Multibit $\Delta\Sigma$ DACs are more favorable than binary $\Delta\Sigma$ DACs for they allow the use of a more aggressively shaped noise transfer function without stability degradation, which leads to a higher in-band signal-to-noise-ratio (SNR). However, unlike binary $\Delta\Sigma$ DACs that are intrinsically linear, multibit $\Delta\Sigma$ DACs cannot guarantee linearity due to device mismatches.

To solve this mismatch problem in multibit $\Delta\Sigma$ DACs, researchers have developed a variety of mismatch shaping techniques, such as [1]–[14]. Their core idea is to insert an element selection logic (ESL) block between the $\Delta\Sigma$ modulator and a multibit unit-element DAC, which consists of N equally-weighted 1-b DACs (see Fig. 1). The output of the ESL block is a vector $\vec{sv}[n]$, each element of which takes a value of either '0' or '1' and controls a 1-bit DAC. The ESL block ensures that the summation of $\vec{sv}[n]$ is equal to the $\Delta\Sigma$ modulator output $d[n]$ to maintain the intended output amplitude, but it has the freedom to select which elements of $\vec{sv}[n]$ are '1' or '0'. By exploiting this freedom, the ESL block ensures that each element of $\vec{sv}[n]$ exhibits a shaped spectrum, and hence, guarantees that mismatch errors are always shaped for any mismatch distribution.

Fig. 1. Block diagram of a mismatch-shaped $\Delta\Sigma$ DAC.

The ESL can be implemented by individual level averaging (ILA) [1], data weighted averaging (DWA) [2]–[4], tree-structured switching blocks [5]–[7], and vector quantization (VQ) [8]–[14]. All these ESL algorithms give rise to tones in the output spectrum. In this work, we will concentrate on the generation of tones in a VQ-based ESL and the remedial measures. The primary motivation behind concentrating on VQ-based ESL is that it can achieve higher order (> 2) mismatch shaping with excellent stability [10], [11], whereas other ESL techniques can achieve stable mismatch shaping only up to the second order. Analyses and removal of tones for other ESL techniques are provided in [3]–[5].

This paper is organized as follows. Section II provides an in-depth analysis of tone generations in standard VQ-based ESL, and forms the basis of a new and simple dithering method proposed in Section III. Finally, the conclusion is brought up in Section IV.

II. Analysis of Spurious Tones in VQ

The structure of a basic VQ-based ESL block is shown in Fig. 2. It includes filters followed by a VQ that consists of a sorter and an array of N comparators.

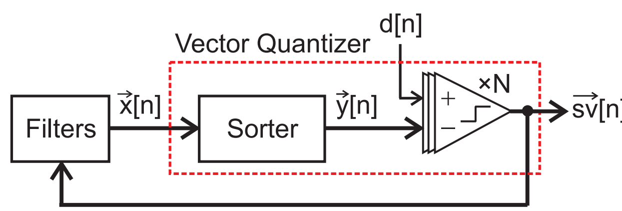

Fig. 2. Structure of a standard VQ-based ESL block.

The sorter used for the analysis in this paper is the partial sorter of [9] (see Fig. 3). We have chosen a partial sorter over the complete sorter for simplicity of presentation. The hardware for the complete sorter and its dithering logics are more complicated, for a complete sorter performs a global comparison between all inputs, while a partial sorter achieves the same by breaking the global comparison into intermediate stages in which only neighboring elements are compared. Nonetheless, the analysis presented in this paper for a partial sorter can be readily applied to a complete sorter.

For $N = 8$, the partial sorter has three stages. The first stage comparators perform the comparisons $x_{2j-1} < x_{2j}$, $j \in [1, 4]$; the second stage comparators perform the comparisons $\sum_{i=4j-3}^{4j-2} x_i < \sum_{i=4j-1}^{4j} x_i$, $j \in [1, 2]$; and the third stage comparator performs the comparison $\sum_{i=1}^{4} x_i < \sum_{i=5}^{8} x_i$.

It follows from the comparison operation that when presented with two inputs a and b, the comparator performs the operation $a < b$. When $a < b$, the output of the comparator

978-1-4673-0219-7/12/$31.00 ©2012 IEEE

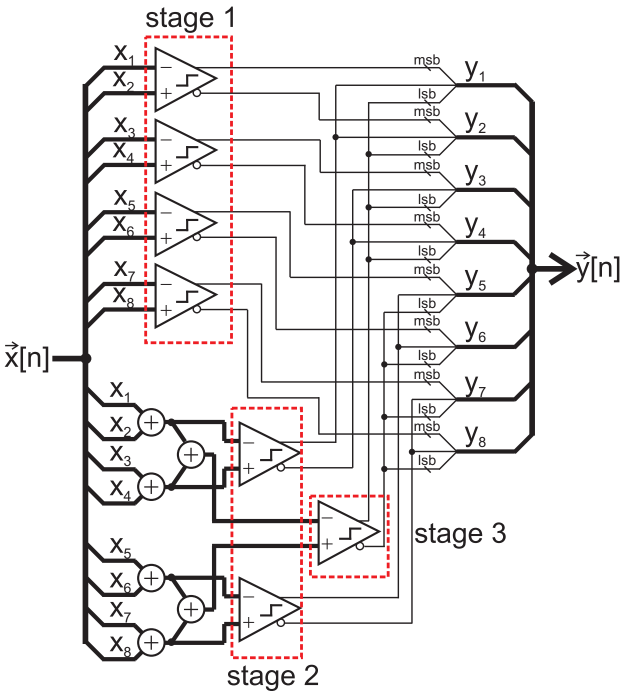

Fig. 3. Block diagram of the partial sorter.

is '1', and when $a > b$, the output of the comparator is '0'. When $a = b$, the ideal comparator output would be '0.5' (the average of '1' and '0'), but the real output is '0', which is the same as the result for $a > b$. This nonideality is usually referred to as the **comparator quantization error**. It causes spurious tones in the ESL output $\vec{sv}[n]$. The periodicity of these spurious tones depends on the ESL input, $d[n]$, as well as the initial conditions of the filters (see Fig. 2).

A. Constant $d[n]$

To get an understanding of how the periodicity changes with the input, let us first examine a simple case of a dc input. First-order mismatch shaping is chosen for this analysis for simplicity of presentation. Analysis of the first-order shaping, nonetheless, provides key insights that can be readily extended to higher-order mismatch shaping.

Let us assume the initial condition on $\vec{x}[n]$ is such that

$$\forall\ x_i[n], x_j[n] \in \vec{x}[n],\ \ |x_i[n] - x_j[n]| \leq 1 \qquad (1)$$

It can be easily seen that even if initially $|x_i[n] - x_j[n]| > 1$, $\vec{x}[n]$ will eventually reach the condition of (1). This is because in each cycle, the partial sorter will always assign a lower rank to the higher value from the pair $(x_{2i-1}, x_{2i}) \in \vec{x}[n]$, $i \in [1, N/2]$ whenever possible. This holds true across all the stages of the partial sorter, i.e, in addition to the first stage, the partial sorter will also assign a lower rank to the higher value from the pairs $(\sum_{i=1}^{2} x_i, \sum_{i=3}^{4} x_i)$, $(\sum_{i=5}^{6} x_i, \sum_{i=7}^{8} x_i)$ and $(\sum_{i=1}^{4} x_i, \sum_{i=5}^{8} x_i)$. Since the $d[n]$ elements of $\vec{x}[n]$ with the lowest ranks are decremented by '1' in each cycle, even if initially any $|x_i[n] - x_j[n]| > 1$, after some m cycles, it will converge to $|x_i[n+m] - x_j[n+m]| \ \leq 1$. Once the condition of (1) is achieved, it will continue to hold as each of the elements of $\vec{x}[n]$ can only be decremented by '1' or retain its value every cycle, and no element of the 2-tuple (x_{2i-1}, x_{2i}) can be decremented by '1' in two successive cycles unless the other element is decremented at least once. Thus, it is easy to see that

$$|x_i[n+1] - x_j[n+1]| \leq 1 \qquad (2)$$

Since the above selection process will take place in every subsequent cycle, (2) can be extended by induction to hold for any n. Therefore,

$$|x_i[n] - x_j[n]| \leq 1,\ \forall\ i, j \in [1, N],\ \ i \neq j \qquad (3)$$

The periodicity of the **comparator quantization noise** is equal to the periodicity with which the inputs to the comparators in the sorter become equal. The number of comparators having both inputs equal is strongly correlated with the initial conditions and the value of $d[n]$, but once a comparator has both inputs equal, the periodicity with which it will have the inputs to be equal again is the same as the periodicity of the rank vector $\vec{y}[n]$. By using (3), we can estimate the periodicity of $\vec{y}[n]$. As already stated, the sorter selects the $d[n]$ highest ranked elements of $\vec{x}[n]$ which are decremented by 1 in the next cycle. Therefore, in order to satisfy (3), the sorter will always try to select the elements that have not been selected in the previous cycles. If $d[n]$ perfectly divides N, then the rank vector $\vec{y}[n]$ will start repeating after $N/d[n]$ cycles. However, if $d[n]$ and N are mutually prime, the ranks will repeat only after N cycles. It is easy to see that the periodicity of $\vec{y}[n]$ is given by $N/gcd(N, d[n])$. The same can be seen from the auto-correlation coefficients of the rank vector for $N = 8$ and $d[n] = 2$ and 3 in Fig. 4; for $d[n] = 3$, the periodicity of the ranks is 8, and for $d[n] = 2$, the periodicity is 4.

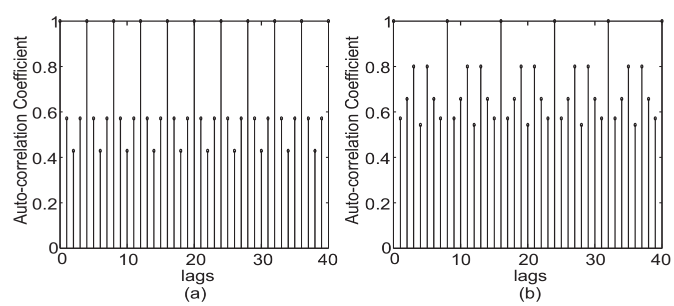

Fig. 4. Auto-correlation coefficients of the rank vector $\vec{y}[n]$ for $N = 8$ and (a) $d[n] = 2$; (b) $d[n] = 3$.

The periodicity can also be seen by looking at the power spectrum of the output of ESL, $\vec{sv}[n]$, and is shown in Fig. 5. For $d[n] = 3$, there are tones at multiples of $f_s/8$ and for $d[n] = 2$, there are tones at multiples of $f_s/4$.

The filters are designed using D flip-flops, which are typically initialized to all zeros. This will result in all the comparators introducing quantization error, which depending on $d[n]$, can have at most a periodicity of N, resulting in significant spurious tone levels in the output. A better initial

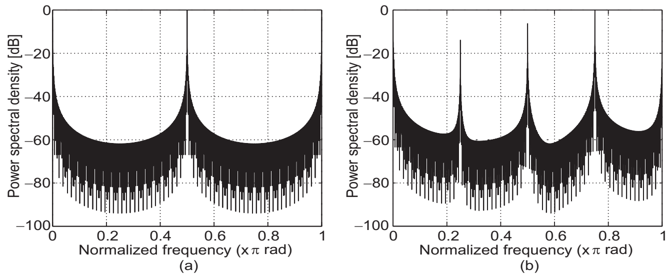

Fig. 5. Power spectral density for first order mismatch shaping with $N = 8$ and (a) $d[n] = 2$; (b) $d[n] = 3$.

condition can be to assign a ‘1’ to the output of some of the D flip-flops and ‘0’ to the rest. However, even then the quantization error can be severe depending on the input magnitude $d[n]$. Also, if $d[n]$ is mutually prime with N, there will be tones at f_s/N, where f_s is the sampling frequency, and its multiples, and can appear in the signal band depending on the value of N. Any analog non-linearities will also give rise to intermodulation products, appearing as distortion terms, in the signal-band.

B. Varying $d[n]$

If $d[n]$ comes from an ADC, $d[n]$ will vary with time. If $d[n]$ changes, the periodicity of the quantization noise sequence will be modulated by $d[n]$, and might be significantly greater than $N/gcd(N, d[n])$. If $d[n]$ has an equal probability of assuming all the values in $[0, N-1]$, the quantization noise will be sufficiently random to approximate white noise. However, if $d[n]$ is confined to the neighborhood of a certain value k, there will be tones at multiples of $gcd(N, k) \cdot f_s/N$ and some of them will potentially fall in the signal band. As an example, it was seen in [4], that for very low amplitude input to a $\Delta\Sigma$ ADC, $d[n]$ has a very high probability to be $N/2$, and hence, the output spectrum will show tones at $f_s/2$.

The present study on the generation of tones in a VQ based ESL gives rise to the question how these tones can be removed. It has been shown that the **comparator quantization noise** gives rise to these tones. Hence, we need look no further than the comparison operation to find a corrective measure. A way to remove the tones is presented in Section III.

III. Our Dithering Technique

Typically, for breaking the periodicity of the **comparator quantization noise**, a random sequence $r[n]$ is added to $d[n]$ or to the output of the filters as shown in Fig. 6. Adding a random sequence to $d[n]$ however needs to be done with caution as this can easily destabilize the modulator in case of a $\Delta\Sigma$ ADC [3]. This is because the output of the quantizer of the $\Delta\Sigma$ ADC is $d[n]$, while the feedback from the DAC is based on the DAC seeing an input of $d[n]$ added to a random sequence. Successive introduction of erroneous feedback can lead to a buildup of the error leading to clipping at the integrators in the loop filter of the ADC. The dither sequence, therefore, has to be of a low amplitude and have a frequency which is a sub-multiple of the sampling frequency. Another problem with this approach is that the dither sequence is not shaped by the filter in the DAC as the dither is added directly to the input, and can result in a significant loss of SNR in the signal band. An approach proposed in [13] is to remove the dither at the output of the DAC [see Fig. 6(a)]. However, the dither sequence has to be passed through another DAC before cancelation at the output, and for perfect cancelation, the two DACs have to matched, which can potentially limit the benefits of this approach as shown by simulation results later. Adding dither to the output of the filters, as in [14] [see Fig. 6(b)], provides better mismatch shaping performance than adding dither to $d[n]$, as the dither is shaped by the loop filter. However, this approach dithers all the comparators in the sorter. We will show that by selectively applying dither to the comparators, only when they need it, can provide a much better mismatch shaping without sacrificing the ability to break up the periodicity of the tones appearing at the DAC output.

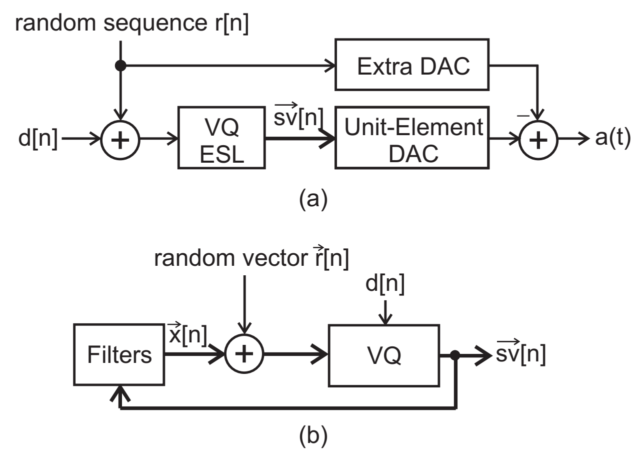

Fig. 6. Block diagrams for (a) the dithering technique in [13]; (b) the dithering technique in [14].

Since the quantization noise occurs only when both the inputs to a comparator in the partial sorter are equal, a random sequence can be used to replace the comparator output every time its inputs are equal, and is shown in Fig. 7. The block diagram shown in Fig. 7 can be used to replace the comparators in Fig. 3. This method does not lead to stability problems when used in a $\Delta\Sigma$ DAC, and the dither sequence is also shaped by the loop filter. It can be easily seen that the dither sequence randomly assigns ranks from the set of 2-tuples $\{(y_{2i-1}, y_{2i})\},\ |y_{2i-1} - y_{2i}| = 4,\ i \in [1, N/2]$ to elements of $\vec{x}[n]$ when the inputs to a comparator are equal.

To check the effectiveness of applying dither to break up the quantization noise periodicity, let us first apply dither while keeping $d[n]$ fixed. The auto-correlation coefficients for the rank vector for $d[n] = 2$ and 3, are plotted in Fig. 8, and the power spectral density of $\vec{sv}[n]$ for $d[n] = 2$ and 3, are plotted in Fig. 9.

It can be clearly seen from Fig. 8 that the added dither has completely randomized the ranks. As a result, the tones should now be absent from the power spectrum of $\vec{sv}[n]$, and this is

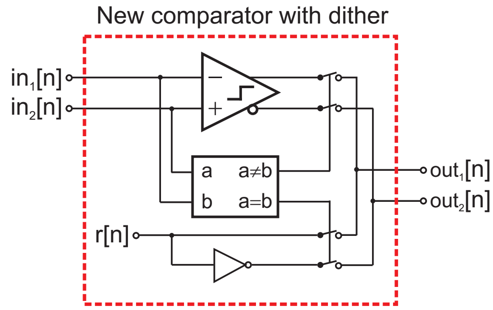

Fig. 7. Proposed dithering technique to remove the **comparator quantization noise**.

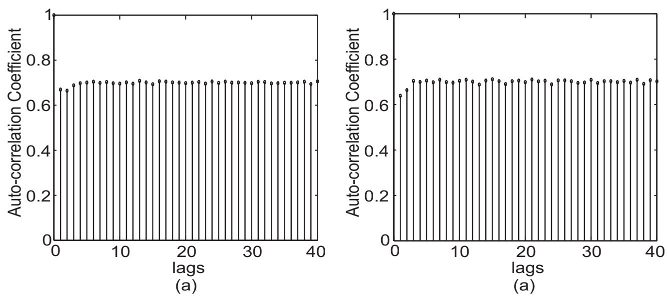

Fig. 8. Auto-correlation coefficients of the rank vector, with dithering, for $N = 8$ and (a) $d[n] = 2$; (b) $d[n] = 3$.

confirmed by the power spectra plots in Fig. 9.

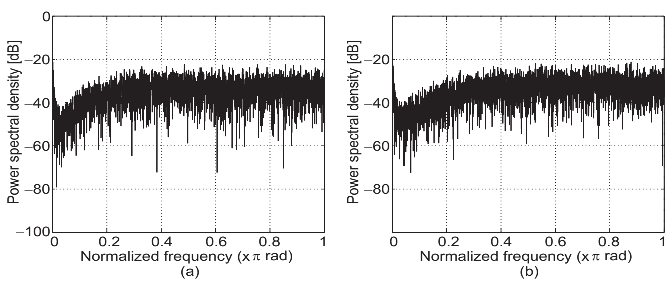

Fig. 9. Power spectral density for first order mismatch shaping with dithering for $N = 8$ and (a) $d[n] = 2$; (b) $d[n] = 3$.

Let us now see the effect of dithering when $d[n]$ is not constant. Fig. 10 shows the simulated output spectra for a 3-bit $\Delta\Sigma$ DAC with 1% device mismatch and an over-sampling ratio of 64. Without dithering, there are appreciable tones due to comparator quantization noise. With our dithering technique applied, the tones are completely removed.

Fig. 10 also compares the simulation result of our dithering technique to those of existing dithering techniques for the VQ-based ESL [13], [14]. It can be seen that our dithering technique has the lowest in-band noise. The in-band noise for the technique of [13] is much larger, for its added dither cannot be perfectly canceled at the output due to device mismatches. The in-band noise for the technique of [14] is a significant improvement over [13], but is still 5-dB larger than ours.

IV. Conclusion

In this paper, we have presented an in-depth analysis of the generation of tones in the output spectra of a vector quantizer

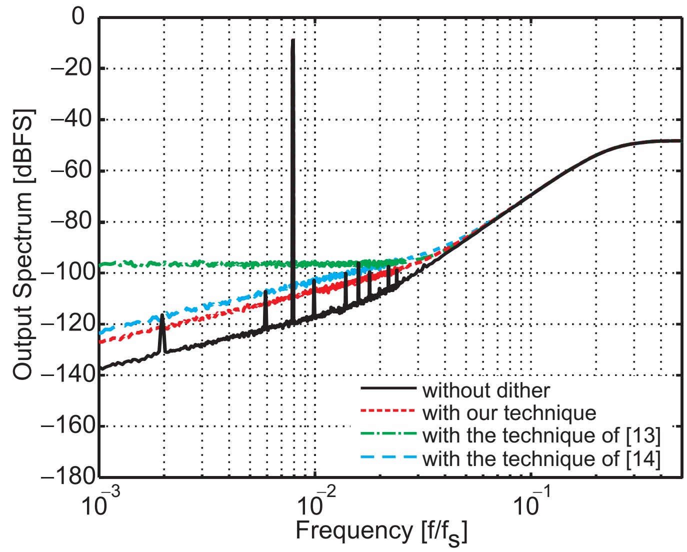

Fig. 10. Power spectra comparing the different dithering techniques for 1% mismatch.

based ESL for a first order mismatch shaping. The key mathematical results were verified by simulations. A simple way of dithering the ESL, in order to remove the tones, has also been presented, and has been shown to provide superior mismatch shaping performance than the existing techniques that rely on adding the dither at the input to the ESL or to the output of the filters preceding the ESL.

References

[1] F. Chen et al., "A high resolution multibit sigma-delta modulator with individual level averaging," *IEEE JSSC*, vol. 30, no. 4, pp. 453–460, April 1995.

[2] I. Fujimori et al., "A 1.5 V, 4.1 mW dual-channel audio delta-sigma D/A converter," *IEEE JSSC*, vol. 33, no. 12, pp. 1863–1870, Dec. 1998.

[3] R. T. Baird et al., "Linearity enhancement of multi-bit $\Sigma\Delta$ A/D and D/A converters using data weighted averaging," *IEEE Trans. Circ. Sys. – II*, vol. 42, no. 12, pp. 753–762, Dec. 1995.

[4] K. D. Chen et al., "An improved technique for reducing baseband tones in sigma-delta modulators employing data weighted averaging algorithm without adding dither," *IEEE Trans. Circ. Sys. – II*, vol. 46, no. 1, pp. 63–68, Jan. 1999.

[5] I. Galton, "Spectral shaping of circuit errors in digital-to-analog converters", *IEEE Trans. Circ. Sys. – II*, vol. 44, pp. 808–817, Nov. 1997.

[6] J. Welz, I. Galton, and E. Fogleman, "Simplified logic for first-order and second-order mismatch-shaping digital-to-analog converters," *IEEE Trans. Circ. Sys. – II*, vol. 48, no. 11, pp. 1014–1027, Nov. 2001.

[7] J. Grilo et al., "A 12-mW ADC delta-sigma modulator with 80 dB of dynamic range integrated in a single-chip bluetooth transceiver" *IEEE JSSC*, vol. 37, no. 3, pp. 271–278, Mar. 2002.

[8] R. Schreier et al., "Noise-shaped multi-bit D/A convertor employing unit elements," *Elec. Lett.* , vol. 31, no. 20, pp. 1712–1713, Sep. 1995.

[9] A. Yasuda et al., "A third-order Δ-Σ modulator using second-order noise shaping dynamic element matching," *IEEE JSSC*, vol. 33, pp. 1879–1886, Dec. 1998.

[10] N. Sun, "High-order mismatch shaping in multi-bit $\Delta\Sigma$ DAC," *IEEE Trans. Circ. Sys. – II*, vol. 58, no. 6, pp. 346–350, Jun. 2011.

[11] N. Sun, "Mismatch-shaped segmented multibit $\Delta\Sigma$ DACs," *to appear in IEEE Trans. Circ. Sys. – I.*

[12] N. Sun and P. Cao, "Low-complexity high-order vector-based mismatch shaping in multi-bit $\Delta\Sigma$ ADCs," *to appear in IEEE Trans. Circ. Sys. –II.*

[13] A. Yasuda, "United States Patent", Patent no. 5872532, Feb. 16, 1999.

[14] L. Risbo et al., "A 108dB-DR 120dB-THD and 0.5V_{rms} output audio DAC with inter-symbol-interference-shaping algorithm in 45nm CMOS," *IEEE ISSCC*, pp. 484–485, 2011.