Low-Power Filtering Via Minimum Power Soft Error Cancellation

Jun Won Choi, *Student Member, IEEE*, Byonghyo Shim, Andrew C. Singer, *Senior Member, IEEE*, and Nam Ik Cho, *Member, IEEE*

***Abstract*—In this paper, an energy-efficient estimation and detection problem is formulated for low-power digital filtering. Building on the soft digital signal processing technique proposed by Hegde and Shanbhag, which combines algorithmic noise tolerance and voltage scaling to reduce power, the proposed minimum power soft error cancellation (MP-SEC) technique detects, estimates, and corrects transient errors that arise from voltage overscaling. These timing violation-induced errors, called soft errors, can be detected and corrected by exploiting the correlation structure induced by the filtering operation being protected, together with a reduced-precision replica of the protected operation. By exploiting a spacing property of soft errors in certain architectures, MP-SEC can achieve up to 30% power savings with no signal-to-noise ratio (SNR) loss and up to 55% power savings with less than 1-dB SNR loss, according to the logic-level simulations performed for an example 25-tap frequency-selective filter.**

***Index Terms*—Algorithmic noise tolerance, digital filter, low power, overscaling, soft error, supply voltage scaling.**

I. Introduction

RELIABILITY and power efficiency in digital signal processing (DSP) systems are important yet often conflicting goals in complex systems. A wide variety of techniques have been developed in the last decade to reduce power in DSP systems [1]–[14]. In general, dynamic power dissipation in a DSP architecture is a quadratic function of the supply voltage, denoted V_{dd}, i.e.,

$$P = C_L V_{dd}^2 f_s \tag{1}$$

where C_L is the effective switching capacitance and f_s is the clock frequency [1]. Due to the quadratic effect on power, a supply voltage reduction scheme, called *dynamic voltage scaling*, is often used to achieve significant power savings. Techniques to minimize V_{dd} include variable voltage scaling [5], [6], multiple supply voltages [7], and retiming techniques [8].

Manuscript received March 7, 2006; revised January 19, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Shuvra S. Bhattacharyya. This work was supported by the Defence Advanced Research Projects Agency (DARPA) and the International Research Internship Program of the Korea Science and Engineering Foundation (KOSEF).

J. W. Choi and A. C. Singer are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL 61801 USA (e-mail: jwchoi@uiuc.edu; acsinger@uiuc.edu).

B. Shim is with School of Electrical Engineering, Korea University, Seoul 136-701, Korea (e-mail: bshim@korea.ac.kr).

N. I. Cho is with the School of Electrical Engineering and Computer Sciences, Seoul National University, Seoul 151-744, Korea (e-mail: nicho@snu.ac.kr).

Digital Object Identifier 10.1109/TSP.2007.896072

In practice, due to increased execution delay at reduced voltage, the extent of supply voltage reduction is limited by the worst case path delay in a given architecture. Specifically, a system is designed such that the critical path delay at the given supply voltage be less than the clock period to void timing errors. Therefore, existing voltage scaling methods [4]–[8] have performed supply voltage reduction up to the point that the critical path delay in the architecture and the sampling period are nearly equal. We refer to this as a critically scaled system, and the supply voltage as the critical supply voltage.

However, in [9], the authors suggested that the supply voltage might be scaled further, below the critical supply voltage for additional power savings, i.e.,

$$V_{dd} = k_{\rm vos} V_{dd-{\rm crit}}, \quad 0 < k_{\rm vos} < 1 \tag{2}$$

where $V_{dd-{\rm crit}}$ is the critical supply voltage. This technique, referred to as *voltage overscaling* (VOS), is motivated by the possibility of controlling the transient errors caused by timing violations, within a tolerable margin, via *algorithmic noise tolerance* [9]. These algorithmic errors are called *soft errors* and the mitigation of them is a key factor for enabling VOS-based DSP architectures.

Previous work to mitigate soft errors include *prediction-based error-correction* methods [9], which estimate the current output sample of the system from previous samples by using a reduced-length forward linear predictor such that a corrupted sample can be replaced whenever an error is detected; *reduced precision replica* methods [10], which approximately calculate the current output to detect and correct errors, and *adaptive error correction* methods [11], which attempt to estimate soft errors directly in a minimum mean-square error (MMSE) sense.

In this paper, we propose a new soft error cancellation technique, called minimum power soft error cancellation (MP-SEC), which can detect, estimate, and correct soft errors. A statistical detection and estimation problem is formulated for the soft errors. We show that the best, in an MMSE sense, unbiased linear estimator, followed by a local maximum-likelihood (ML) detector, provides accurate estimates of soft errors, under some mild assumptions. This formulation enables power dissipation of the soft error canceller to be traded off against error resilience. For this setup, observable signals at the input and output of the main filter to be protected are collected as shown in Fig. 1. While the main filter is operating in a VOS regime to reduce power, an error cancellation unit will not suffer soft errors for much of this regime, due to the reduced complexity of the MP-SEC units used to detect and correct any soft errors induced by VOS. Necessarily, the power consumed by the soft error cancellation

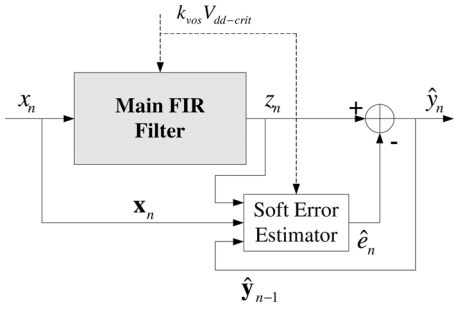

Fig. 1. Proposed MP-SEC soft error estimator applied to an FIR filter. The vector $\mathbf{x}_n$ contains past and present samples of x_n and the vector $\hat{\mathbf{y}}_{n-1}$ contains past values of the sequence of $\hat{y}_n$.

(SEC) unit must be small as compared to the savings achieved through VOS. We explore the minimum power configuration for such an SEC unit and develop an adaptive power control algorithm, which optimizes the power dissipation of the SEC unit with respect to the selection of which observations to use and their numerical precision in the SEC unit.

An important observation that makes this approach possible is that soft errors can be characterized as discrete, i.e., finite alphabet, signals. As most arithmetic units perform least significant bit (LSB)-first computation, erroneous bits due to VOS will occur largely for bits near the most significant bit (MSB). While this may seem problematic, any resulting soft errors will take on a small set of large amplitudes as their possible outcomes, and these possible amplitudes will be spaced apart, such that soft error estimation can be treated as an M-ary pulse amplitude modulation (PAM) signal detection problem [22].

The remainder of this paper is organized as follows. In Section II, we derive the soft error estimation and detection algorithm and provide performance analysis. In Section III, we present a power-optimized algorithm for the soft error canceller. In Section IV, we describe a hardware design and some simulation results, and in Section V, some conclusions are given.

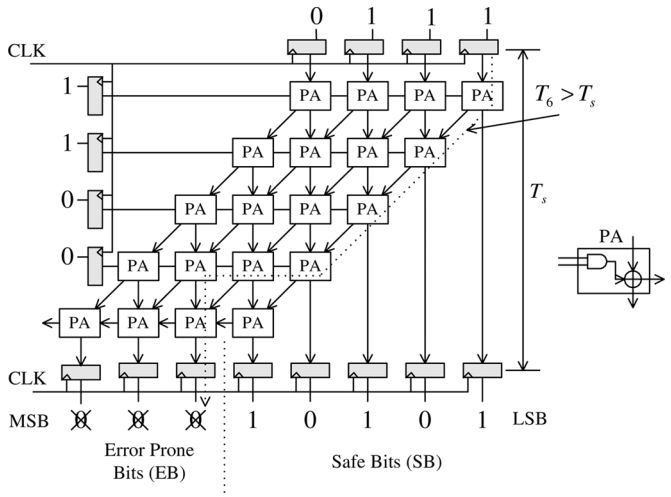

Fig. 2. A 4×4 carry–save multiplier for inputs (in two's compliment) $0111_2 \times 0011_2$. For this $B = 8$-bit example, the clock period, T_s is shown to be less than the critical path T_6. Three of the resulting bits become error prone, while five of the bits are "safe." ($B = 8$ and $M = 5$.)

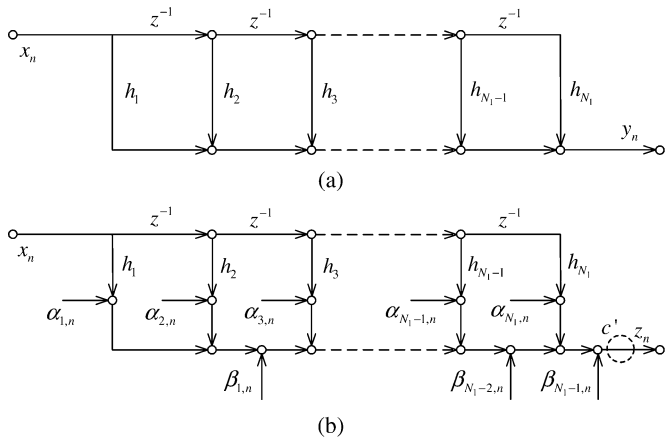

Fig. 3. Flow-graph of direct form I FIR filter : (a) ideal model and (b) soft error model.

II. Soft Error Cancellation Approach

In this section, we investigate statistical estimation and detection of soft errors. We first describe the framework where soft errors arise and their statistical description, and then derive the soft error estimator and detector.

A. Soft Error Model

By appropriate design, soft errors can be constrained to appear near the MSB in the binary representation of the signal samples for LSB-first arithmetic units used in many structures for computation. For example, we will assume a two's complement number representation such that $x = -b_0 + \sum_{i=1}^{B-1} b_i 2^{-i}$ in B-bit precision. As a simple illustration, consider the 4×4 carry–save multiplier shown in Fig. 2. Let T_s and T_i be the sampling period and the worst path delay to ith output bit, respectively. When $k_{\rm vos} = 1$ (no VOS), it is evident from the figure that $T_8 \geq \cdots \geq T_1$, due to the use of LSB-first computation. However, as we scale the supply voltage below $V_{dd-\rm crit}$, the worst-case delays T_i increase for all i, so that output bits become divided into two sets: error prone bits (EBs), where the timing conditions may be violated $(T_i > T_s)$, and safe bits (SBs), where the timing relation is guaranteed. If we let B and M be the number of output bits and safe bits in the multiplier, then the soft errors are expressed as a combination of bits from the EB region, and their magnitudes become a multiple of $2^M/2^B$. This implies that the possible magnitudes of soft errors are equally spaced by $2^M/2^B$. This property, which we refer to as a *spacing property*, plays a key role in estimating soft errors.

To illustrate the impact of soft errors on finite-impulse-response (FIR) filters, we consider an N_1-tap causal FIR filter whose direct-form I implementation is shown in Fig. 3(a). Under VOS, the processing units in the main filter violating the timing requirement may suffer from soft errors. As shown in Fig. 3(b), the system can be described by an equivalent linear additive model. In this model, a soft error, denoted $\alpha_{i,n}$ for ith multiplier and $\beta_{j,n}$ for jth adder, are injected at the output of each arithmetic unit. If soft errors do not appear, $\alpha_{i,n} = 0$ and $\beta_{j,n} = 0$.

These sources of soft error can be collected together and merged into one signal source, e_n at the node c', where e_n is given by

$$e_n = \sum_{i=1}^{N_1} \alpha_{i,n} + \sum_{j=0}^{N_1-1} \beta_{j,n} + \gamma_n \tag{3}$$

where γ_n represents the errors that might occur due to overflow from the adders. Due to the spacing property, e_n takes on values in $\Omega = \{k2^{M-B} \,|\, k \in \mathcal{Z}, k \in [-2^B/2^M, 2^B/2^M)\}$, where B and M are the numerical precision of output and the smallest number of SBs of all arithmetic units, respectively. The noisy output z_n in the presence of a soft error is given by

$$z_n = y_n + e_n = \sum_{k=1}^{N_1} h_k x_{n-k+1} + e_n \tag{4}$$

where x_n and y_n are the nth sample of the input and the error-free output, respectively, and h_k is the kth FIR filter coefficient. In the sequel, we will use the vector notation $\mathbf{x}_n = [x_n, \ldots, x_{n-N_1+1}]^T, \mathbf{y}_n = [y_{n-1}, \ldots, y_{n-N_2}]^T$, and $\bar{h} = [h_1, \ldots, h_{N_1}]^T$, for convenience, where a boldface vector denotes a vector of random variables, and an overbar denotes a deterministic vector.

B. Soft Error Cancellation

The objective of soft error cancellation is to subtract an estimate of the soft error from the erroneous output if necessary, i.e.,

$$\hat{y}_n = z_n - \hat{e}_n = y_n + (e_n - \hat{e}_n) \tag{5}$$

where $\hat{y}_n$ is the error-corrected output, and $\hat{e}_n$ is an estimate of the soft error. Hence, ideal soft error cancellation provides that $(e_n - \hat{e}_n)$ is zero such that $\hat{y}_n = y_n$.

1) Soft Error Estimation: Assume that the input and therefore the output signals x_n and y_n are zero-mean stationary random processes. As mentioned in the previous section, the soft error estimator makes decisions based on the observations of a subset of the sets $\{z_n, \{x_n, \ldots, x_{n-N_1+1}\}, \{\hat{y}_{n-1}, \ldots, \hat{y}_{n-N_2}\}\}$, where the elements are selected to trade off performance of the estimation for the added power of the soft error cancellation. To limit the added complexity or power drawn by the estimator, we restrict to the precision used to describe the sets $\{x_n, \ldots, x_{n-N_1+1}\}$ and $\{\hat{y}_{n-1}, \ldots, \hat{y}_{n-N_2}\}$ to only p bits, producing the vectors $\mathbf{x}_{q,n} = [x_{q,n}, \ldots, x_{q,n-N_1+1}]^T$ and $\hat{\mathbf{y}}_{q,n} = [\hat{y}_{q,n-1}, \ldots, \hat{y}_{q,n-N_2}]^T$, where x_n and $\hat{y}_n$ are quantized to $x_{q,n}$ and $\hat{y}_{q,n}$. Then, we mask the vectors, $\mathbf{x}_{q,n}$ and $\hat{\mathbf{y}}_{q,n}$ using the switching vectors $\bar{c} = [c_1, \ldots, c_{N_1}]^T$ and $\bar{d} = [d_1, \ldots, d_{N_2}]^T$, producing $\mathbf{x}_{c,n} = [c_1 x_{q,n}, \ldots, c_{N_1} x_{q,n-N_1+1}]^T$ and $\hat{\mathbf{y}}_{c,n} = [d_1 \hat{y}_{q,n-1}, \ldots, d_{N_2} \hat{y}_{q,n-N_2}]^T$, where c_i and d_i take on values 0 or 1. As a result, the reduced observation takes the form

$$\begin{bmatrix} z_n \\ \mathbf{x}_{c,n} \\ \hat{\mathbf{y}}_{c,n} \end{bmatrix} = \begin{bmatrix} y_n \\ \mathbf{x}_{c,n} \\ \hat{\mathbf{y}}_{c,n} \end{bmatrix} + e_n \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{6}$$

Note that the resolution and selection of observations are controlled by $p, \bar{c}$, and $\bar{d}$. For the time being, we assume that $p, \bar{c}$, and $\bar{d}$ are given. Based on these reduced observations, a linear (affine) unbiased estimate of e_n can be expressed

$$\tilde{e}_n = a_0 z_n + [\,\bar{w}^T \quad \bar{v}^T\,] \begin{bmatrix} \mathbf{x}_{c,n} \\ \hat{\mathbf{y}}_{c,n} \end{bmatrix} + A \tag{7}$$

where $\bar{w}^T = [w_1, \ldots, w_{N_1}]^T$ and $\bar{v}^T = [v_1, \ldots, v_{N_2}]^T$. To satisfy the unbiased constraint, $E[\hat{e}_n] = e_n$ for all $e_n \in \Omega, a_0 = 1$, and

$$A = -[\,\bar{w}^T \quad \bar{v}^T\,] E \begin{bmatrix} \mathbf{x}_{c,n} \\ \hat{\mathbf{y}}_{c,n} \end{bmatrix} \tag{8}$$

where $E[\,\cdot\,]$ denotes a statistical expectation. The vectors $\bar{w}$ and $\bar{v}$ are determined to minimize the variance of the estimate

$$E[(\tilde{e}_n - e_n)^2] = E\left[\left(y_n + [\,\bar{w}^T \quad \bar{v}^T\,] \begin{bmatrix} \mathbf{x}_{c,n} \\ \hat{\mathbf{y}}_{c,n} \end{bmatrix} + A\right)^2\right]. \tag{9}$$

The estimator coefficients, $\bar{w}$ and $\bar{v}$ are obtained by finding a linear minimum mean-square error (LMMSE) estimate of y_n based on $\mathbf{x}_{c,n}$ and $\hat{\mathbf{y}}_{c,n}$. Substituting (8) into (9), we write this variance

$$\begin{aligned} &E[(\tilde{e}_n - e_n)^2 \,|\, \bar{c}, \bar{d}] \\ &\quad = E\left[\left(y_n + [\,\bar{w}_c^T \quad \bar{v}_c^T\,] \left(\begin{bmatrix} \mathbf{x}_{q,n} \\ \hat{\mathbf{y}}_{q,n} \end{bmatrix} - E \begin{bmatrix} \mathbf{x}_{q,n} \\ \hat{\mathbf{y}}_{q,n} \end{bmatrix} \right)\right)^2\right] \end{aligned} \tag{10}$$

where $\bar{w}_c = [c_1 w_1, \ldots, c_{N_1} w_{N_1}]^T$, and $\bar{v}_c = [d_1 v_1, \ldots, d_{N_2} v_{N_2}]^T$. Note that the entries $w_{c,i}$ and $v_{c,j}$ of $\bar{w}_c$ and $\bar{v}_c$ are constrained to be zero when c_i and d_j are zero. The coefficients of $\bar{w}_c$ and $\bar{v}_c$, which minimize (10) subject to the coefficient constraint, are given by

$$\begin{aligned} \begin{bmatrix} w_{c,i_1} \\ \vdots \\ w_{c,i_{N_1'}} \\ v_{c,j_1} \\ \vdots \\ v_{c,j_{N_2'}} \end{bmatrix} &= -\mathrm{Cov} \begin{pmatrix} x_{q,n-i_1+1} \\ \vdots \\ x_{q,n-i_{N_1'}+1} \\ \hat{y}_{q,n-j_1} \\ \vdots \\ \hat{y}_{q,n-j_{N_2'}} \end{pmatrix}^{-1} \\ &\quad \cdot \mathrm{Cov} \left(\begin{bmatrix} x_{q,n-i_1+1} \\ \vdots \\ x_{q,n-i_{N_1'}+1} \\ \hat{y}_{q,n-j_1} \\ \vdots \\ \hat{y}_{q,n-j_{N_2'}} \end{bmatrix}, \mathbf{x}_n \right) \bar{h} \end{aligned} \tag{11}$$

where $i_1 \cdots i_{N_1'} \in \Lambda_1 = \{i : c_i = 1\}, j_1 \cdots j_{N_2'} \in \Lambda_2 = \{j : d_j = 1\}$, and $\mathrm{Cov}(a,b) = E[ab^T] - E[a]E[b]^T$ and $\mathrm{Cov}(a) = \mathrm{Cov}(a,a)$. Note that $w_{c,i} = 0$ for $i \in \Lambda_1^c$ and $v_{c,j} = 0$ for $j \in \Lambda_2^c$, where the superscript c denotes a set complement. Let the quantization error be $\Delta x_n = x_n - x_{q,n}$ and $\Delta \hat{y}_n = \hat{y}_n - \hat{y}_{q,n}$. When a random variable x has a two's complement representation with independent and identically distributed (i.i.d.) bits and is truncated from B bits to p bits, the

mean and variance of the resulting quantization errors can be shown to be given by

$$\mu_{\Delta x} = 2^{-p} - 2^{-B} \tag{12}$$

$$\mathrm{Var}_{\Delta x} = \frac{1}{3}(2^{-2p} - 2^{-2B}). \tag{13}$$

When the observations are quantized from B bits to p bits, and using (12), the parameter A is given by

$$A = (2^{-p} - 2^{-B}) \left(\sum_{i=1}^{N_1} w_{c,i} + \sum_{i=1}^{N_2} v_{c,i} \right) \tag{14}$$

and using (13), the coefficients, $w_{c,i_1}, \ldots, w_{c,i_{N_1'}}$ and $v_{c,j_1}, \ldots, v_{c,j_{N_2'}}$ are given by

$$\begin{bmatrix} w_{c,i_1} \\ \vdots \\ w_{c,i_{N_1'}} \\ v_{c,j_1} \\ \vdots \\ v_{c,j_{N_2'}} \end{bmatrix} d = - \left(\mathrm{Cov} \begin{pmatrix} x_{n-i_1+1} \\ \vdots \\ x_{n-i_{N_1'}+1} \\ \hat{y}_{n-j_1} \\ \vdots \\ \hat{y}_{n-j_{N_2'}} \end{pmatrix} + \frac{1}{3}(2^{-2p} - 2^{-2B}) I \right)^{-1} \cdot \mathrm{Cov} \left(\begin{bmatrix} x_{n-i_1+1} \\ \vdots \\ x_{n-i_{N_1'}+1} \\ \hat{y}_{n-j_1} \\ \vdots \\ \hat{y}_{n-j_{N_2'}} \end{bmatrix}, \mathbf{x}_n \right) \bar{h} \tag{15}$$

where I is an $(N_1'+N_2')$-by-$(N_1'+N_2')$ identity matrix, and it has been assumed that Δx_n and $\Delta \hat{y}_n$ are mutually uncorrelated and uncorrelated with x_n and $\hat{y_n}$, which is reasonable for moderate values of p. The coefficients, $\bar{w}_c$ and $\bar{v}_c$ can be implemented with linear filters, and hence we refer to $\bar{w}_c$ and $\bar{v}_c$ as the main estimation filter (MEF), and specifically to $\bar{w}_c$ as the feedforward MEF (FF-MEF) and to $\bar{v}_c$ as the feedback MEF (FB-MEF), respectively. The resulting soft error estimate is given by

$$\tilde{e}_n = z_n + [\, \bar{w}_c^T \quad \bar{v}_c^T \,] \begin{bmatrix} \mathbf{x}_{q,n} \\ \hat{\mathbf{y}}_{q,n} \end{bmatrix} + A. \tag{16}$$

Although we have obtained an unbiased estimate of e_n, it will not generally satisfy the constraint that the estimate of e_n lie in Ω, while we know that the true e_n must lie in Ω. We consider

$$\tilde{e}_n = e_n + \left(y_n + [\, \bar{w}_c^T \quad \bar{v}_c^T \,] \begin{bmatrix} \mathbf{x}_{q,n} \\ \hat{\mathbf{y}}_{q,n} \end{bmatrix} + A \right) \tag{17}$$

where we refer to the term in (17) in braces as the estimation error of the MMSE estimator, which will be called the residual error. Given the unbiased estimate, $\tilde{e}_n$, we may be able to refine the estimate by maximizing the log-likelihood function of $\tilde{e}_n$ with respect to e_n, i.e.,

$$\hat{e}_n = \arg \max_{e_n \in \Omega} \ln p(\tilde{e}_n; e_n). \tag{18}$$

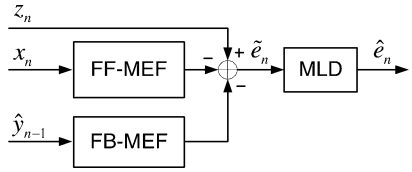

Fig. 4. Soft error estimator.

When the PDF of the residual error is symmetric about zero and unimodal, then a local maximum-likelihood estimate (MLE) of soft error is given by

$$\hat{e}_n = T(\tilde{e}_n) \tag{19}$$

where

$$T(x) = \frac{2^M}{2^B} i, \quad \text{where } \left| x - \frac{2^M}{2^B} i \right| \le \left| x - \frac{2^M}{2^B} j \right|, \text{ for all } j \tag{20}$$

for i and j are integers between $-2^B/2^M$ and $2^B/2^M - 1$. This MLE is based on the statistic $\tilde{e}$, not on $[z_n, \mathbf{x}_n, \hat{\mathbf{y}}_n]^T$. We see that $\hat{e}_n$ would become the MLE obtained based on $[z_n, \mathbf{x}_n, \hat{\mathbf{y}}_n]^T$, if $z_n, \mathbf{x}_n$, and $\hat{\mathbf{y}}_n$ were jointly Gaussian and not truncated. In the general case, the estimate, $\hat{e}_n$, would be suboptimal but can be practically implemented with low complexity.

Based on (17) and (19), the total estimation error $\hat{e}_n - e_n$ is given by

$$\hat{e}_n - e_n = T \left(y_n + [\, \bar{w}_c^T \quad \bar{v}_c^T \,] \begin{bmatrix} \mathbf{x}_{q,n} \\ \hat{\mathbf{y}}_{q,n} \end{bmatrix} + A \right). \tag{21}$$

Equation (21) implies that when the residual error is smaller than $2^M/2^B$, the estimate would be accurate i.e., $\hat{e}_n = e_n$. This means that wider soft error spacing leads to better estimation. We may be able to reduce the variance of residual error by increasing p i.e., the resolution of the quantizer or the number of ones in $\bar{c}, \bar{d}$. Accordingly, we can assume that we set the values of $\bar{c}, \bar{d}$ and p such that

$$E[(\hat{e}_n - e_n)^2] \ll \sigma_y^2 \tag{22}$$

where σ_y^2 is the variance of y_n. As long as this assumption holds, the impact of soft error correction errors on the statistics of y_n can be assumed to be negligible. This low-error regime allows the use of y_{n-k}, instead of $\hat{y}_{n-k}$, in the estimate formulation, (15).

The soft error estimator derived herein consists of two parts: 1) the feedforward and feedback linear filters, which enhance the quality of estimation, and 2) the sequent maximum-likelihood detector (MLD), which maps the input to the nearest soft error candidate, as shown in Fig. 4. This estimation mechanism shows an interesting analogy to that of a PAM receiver which consists of a matched filter correlator that improves the estimation signal-to-noise ratio (SNR) and the maximum-likelihood symbol detector.

2) Soft Error Detection: To determine when an error has occurred and enable error correction, the following hypothesis test for soft error detection is used:

$$H_1 : z_n = \overline{h}^T \mathbf{x}_n + e_n$$
$$H_0 : z_n = \overline{h}^T \mathbf{x}_n \tag{23}$$

where e_n takes a value from Ω. As the parameter e_n is unknown, this problem can be interpreted as a composite hypothesis test for which a generalized-likelihood ratio test (GLRT) is often used [21]. Subject to complexity constraints, we may also base the detection on the test statistic $\hat{e}_n$, not on $[z_n, \mathbf{x}_n, \hat{\mathbf{y}}_n]^T$. We compare the log-likelihood ratio, maximized over e_n with a threshold, τ, i.e.,

$$\Lambda = \max_{e_n \in \Omega} \ln \frac{p(\tilde{e}_n \,|\, H_1; e_n)}{p(\tilde{e}_n \,|\, H_0)} \underset{H_1}{\overset{H_0}{\gtrless}} \tau \tag{24}$$

where τ may be chosen using a constant false alarm rate (CFAR) criterion. To simplify the development, we assume the condition (22) and that the residual error is well approximated by a zero-mean Gaussian; then, the maximizer of the log-likelihood ratio becomes the MLE given in (19), and we can substitute e_n by $\hat{e}_n$ in (24). The resulting approximated detection rule is given by

$$\tilde{e}_n^2 - (\tilde{e}_n - \hat{e}_n)^2 \underset{H_1}{\overset{H_0}{\gtrless}} 2\tau\sigma_r^2 \tag{25}$$

where σ_r^2 is the variance of the residual error and is given by

$$\sigma_r^2 = \sigma_y^2 - \begin{bmatrix} \overline{w}_c \\ \overline{v}_c \end{bmatrix}^T \left(\mathrm{Cov} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{y}_n \end{bmatrix} + \frac{1}{3}\left(2^{-2p} - 2^{-2B}\right) I \right) \times \begin{bmatrix} \overline{w}_c \\ \overline{v}_c \end{bmatrix}. \tag{26}$$

In practice, this soft error detector may not be necessary when the error spacing is large, because the quantization function, $T(\cdot)$ in (19) performs the task of detecting the errors if the residual error exceeds $2^M/2^B$.

3) Algorithmic Performance Measure: In order to analyze the performance of the soft error canceller, we use the signal power to soft error power ratio (SSR), defined as

$$\mathrm{SSR} = \frac{\text{power of desired signal}}{\text{power of residual soft error}}. \tag{27}$$

From (5), the SSR at the output of the main filter is given by

$$\mathrm{SSR} = 10\log_{10} \frac{\sigma_y^2}{E[(\hat{e}_n - e_n)^2]}. \tag{28}$$

We may also use other measures such as the SNR. As an example, consider an application in which the output signal y_n consists of both a desired signal d_n and an undesired noise signal η_n, i.e.,

$$y_n = d_n + \eta_n. \tag{29}$$

After applying the soft error canceller, the SNR is given by

$$\mathrm{SNR} = 10\log_{10} \frac{\sigma_d^2}{\sigma_\eta^2 + E[(\hat{e}_n - e_n)^2]}. \tag{30}$$

Note that both measures in (28) and (30) depend on the power of the estimation error $\hat{e}_n - e_n$, or residual mean-square error (RMSE). Hence, the algorithmic performance of MP-SEC can be thoroughly analyzed by deriving the RMSE.

4) RMSE Analysis: In this subsection, we provide an analysis of RMSE when employing the soft error estimator. In deriving the estimate, we neglected the effect of previous decisions assuming the condition (22) to hold. However, in practice, inaccurate estimates of soft errors may cause subsequent errors in estimating soft errors, permitting error propagation. Hence, we need to analyze the RMSE considering the consequences of using tentative decisions.

First, we assume that the soft error detector is not employed. We assume that the residual error has a Gaussian distribution $\mathcal{N}(0, \sigma_r^2)$ for analysis. If the previous errors are essentially correct, the probability that the soft error estimate is not correct is

$$P(\hat{e}_n \neq e_n) = 1 - 2Q\left(\sqrt{\frac{\lambda^2}{4\sigma_r^2}}\right) \tag{31}$$

where $Q(x) = \int_x^\infty (1/\sqrt{2\pi}) \exp(-t^2/2)$, and $\lambda = 2^M/2^B$. Note that the probability of error will increase as λ increases or σ_r^2 decreases. Hence, the ratio λ^2/σ_r^2 is a crucial factor that affects the quality of soft error estimation, even when error propagation happens. The RMSE is expressed as

$$E[(e_n - \hat{e}_n)^2] = \sum_{k=-\infty}^{\infty} (\lambda k)^2 P(\hat{e}_n - e_n = \lambda k). \tag{32}$$

When ignoring the effect of previous decision errors, the RMSE is given by

$$\mathrm{SNR} = 2\sum_{k=1}^{\infty} \lambda^2 k^2 \left\{ Q\left(\sqrt{\frac{\lambda^2}{\sigma_r^2}}\left(k - \frac{1}{2}\right)\right) \tag{33}$$

$$- Q\left(\sqrt{\frac{\lambda^2}{\sigma_r^2}}\left(k + \frac{1}{2}\right)\right)\right\}. \tag{34}$$

Note that RMSE also increases when σ_r^2 increases.

Now, assume that the previous decision errors are no longer negligible. We use a Markov chain model to describe the sequence of previous errors and evaluate the RMSE in the steady state. This approach has been employed in the analysis of decision feedback equalization [12], [13]. Let the estimation error at time $n-k$ be s_{n-k}, i.e., $s_{n-k} = e_{n-k} - \hat{e}_{n-k}$. Without loss of generality, we consider the feedback errors to come from the length N_2 sequence $\hat{y}_{n-1}, \ldots, \hat{y}_{n-N_2}$. Then, we define the state $(i_1, \ldots, i_{N_2})$ at time n to be

$$\mathrm{state}_n(i_1, \ldots, i_{N_2}) = \{s_{n-1} = i_1\lambda, \ldots, s_{n-N_1} = i_{N_2}\lambda\}. \tag{35}$$

where $i_1, \ldots, i_{N_2}$ are integers in $[-2\cdot 2^{B-M}+1, 2\cdot 2^{B-M}-1]$. The number of possible decision errors is $(4\cdot 2^{B-M}-1)$, and the total number of states should be $(4\cdot 2^{B-M}-1)^{N_2}$. However, for moderate configurations of the estimator, the possibility that a decision error occurs with large magnitude is small as shown

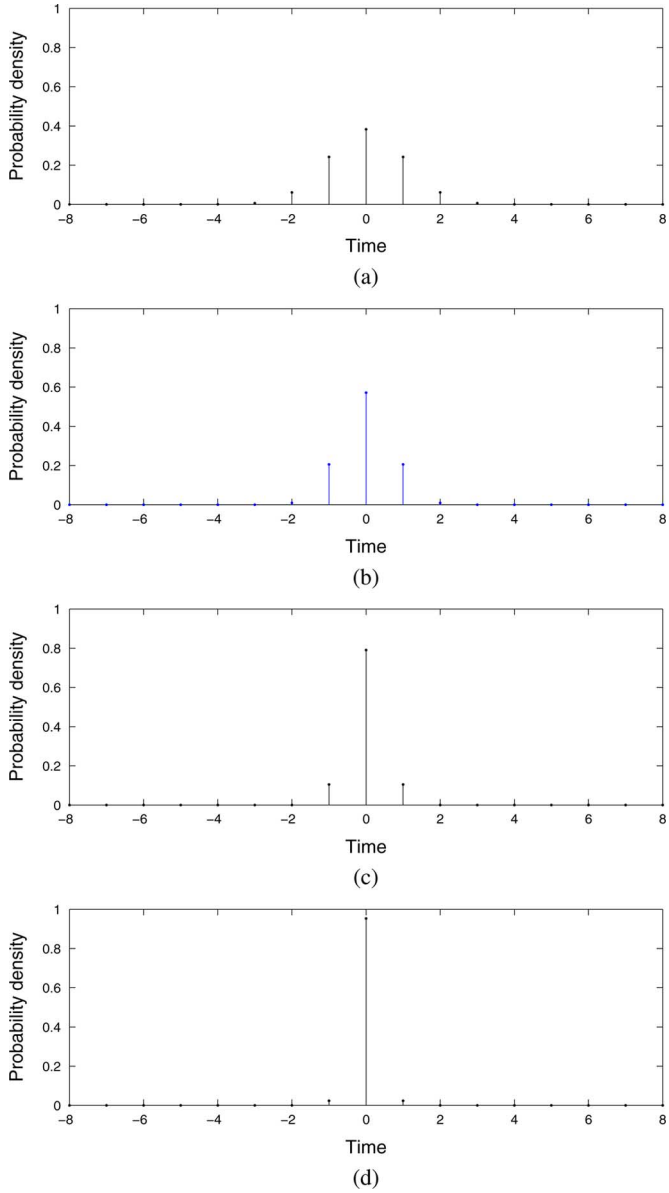

Fig. 5. Analytically modeled pdf of the estimation error, or $\hat{e}_n - e_n$ when λ^2/σ_r^2 is (a) 0 dB, (b) 4 dB, (c) 8 dB, and (d) 12 dB. Note that time axis is scaled by $1/\lambda$.

in Fig. 5, which is derived under modest assumptions in the following analysis. This reduces the number of states by counting only $2m+1 (< 2^{B-M})$ error magnitudes near zero, where m is a small integer and, hence, the total number of states becomes $(2m+1)^{N_2}$. Then, we can present the following conditional probabilities:

$$P(s_n = i\lambda \,|\, \text{state}_n(i_1, \ldots, i_{N_2})) = Q\left(\frac{(2i-1)\lambda - 2\sum_{k=1}^{N_2} v_{c,k} i_k \lambda}{2\sigma_r}\right) - Q\left(\frac{(2i+1)\lambda - 2\sum_{k=1}^{N_2} v_{c,k} i_k \lambda}{2\sigma_r}\right), \quad (i = -m+1, \ldots, m-1). \quad (36)$$

and

$$P(s_n = \pm m\lambda \,|\, \text{state}_n(i_1, \ldots, i_{N_2})) = Q\left(\frac{(\pm 2m-1)\lambda \mp 2\sum_{k=1}^{N_2} v_{c,k} i_k \lambda}{2\sigma_r}\right). \quad (37)$$

We list from $\text{state}_n(-m, \ldots, -m)$ to $\text{state}_n(m, \ldots, m)$ in an appropriate but arbitrary order and number them from state 1 to state $(2m+1)^{N_2}$. The $(2m+1)^{N_2}$-by-$(2m+1)^{N_2}$ state-transition matrix T from time n to time $n+1$ is given by

$$T\left(\text{state}_n\left(i_1, \ldots, i_{N_2}\right), \text{state}_{n+1}\left(j_1, \ldots, j_{N_2}\right)\right) \begin{cases} = P\left(s_n = j_1\lambda \,|\, \text{state}_n\left(i_1, \ldots, i_{N_2}\right)\right), \\ \quad \text{if } i_1 = j_2, i_2 = j_3, \ldots, i_{N_2-1} = j_{N_2} \\ = 0, \quad \text{otherwise.} \end{cases}$$

Let the probability of staying in the ith state in the steady state to be π_i. The vector $\left[\pi_1, \ldots, \pi_{(2m+1)^{N_2}}\right]$ can be found by solving the following simultaneous equations:

$$\left[\pi_1 \quad \cdots \quad \pi_{(2m+1)^{N_2}}\right] T = \left[\pi_1 \quad \cdots \quad \pi_{(2m+1)^{N_2}}\right] \quad (38)$$

$$\sum_{k=1}^{(2m+1)^{N_2}} \pi_k = 1. \quad (39)$$

The solution to these equations is obtained by finding the null vector of the matrix $(T^T - I)$ and normalizing the null vector to satisfy (39). The probability that the decision error equals λk in the steady state is given by

$$P(\hat{e}_n - e_n = \lambda k) = \sum_{\{j : i_1 \text{ of } j\text{th state} = k\}} \pi_j. \quad (40)$$

for $-m \le k \le m$.

Fig. 5 shows the probability density functions (pdf's) of $\hat{e}_n - e_n$ using this result, in the case that two feedback observations are employed, i.e., $N_2 = 2, B = 16$ and $M = 12$. Note that the pdf is defined on the discrete event space due to the soft error spacing property and more concentrated around zero with increasing λ^2/σ_r^2. Finally, the RMSE can be obtained by plugging (40) into (32). In Fig. 6, we plot the RMSE versus λ^2/σ_r^2 when N_2 is 2 and 3. The RMSE decreases as λ^2/σ_r^2 increases and as N_2 decreases. The experimental values of RMSE are close to the analytically derived curves. It should be noted that the performance of the soft error canceller depends on σ_r^2, or equivalently how well the MMSE estimator estimates the desired output y_n using the given information inferred in the observations.

III. Energy Minimum Soft Error Cancellation

In this section, we introduce two approaches to optimize the power dissipation in the soft error canceller. First, an energy-minimum design is provided, which assumes stationary statistics of the signals of interest. The key feature of this strategy is enabling "one shot" design that can be fixed, e.g., in a VLSI design. Second, we introduce a dynamic power optimization technique, which controls the configuration of the SEC unit in real time to cope with time-varying environments.

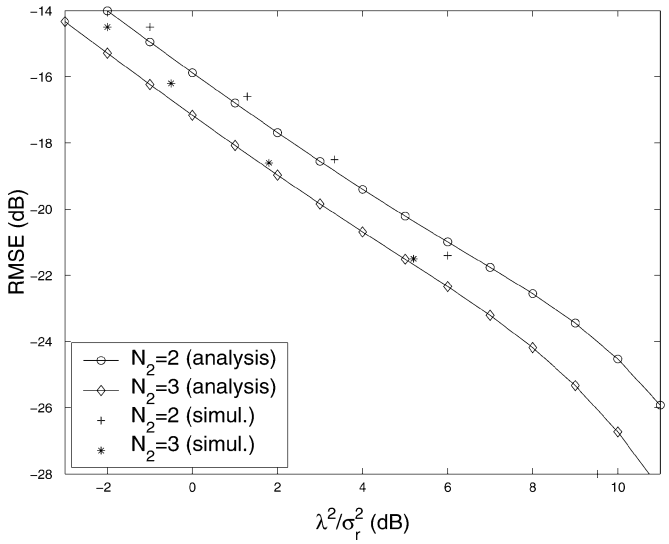

Fig. 6. Analytically derived and simulated RMSEs are shown versus λ^2/σ_r^2.

A. Power Optimization Criterion

In general, the power dissipation P_v of a system may follow

$$P_v \propto (C_0 + C_s)(k_{\text{vos}} V_{dd-\text{crit}})^2 \tag{41}$$

where C_0 and C_s are the switched capacitances in the main filter block and SEC block, respectively. This relationship implies that, for maximum power savings, we have to reduce the complexity of the SEC block as much as possible. A rather general approach to optimizing power consumption has been addressed in [14] and [15] in the form of a constrained optimization that minimizes an estimate of the power dissipation subject to particular performance constraints. We express this formulation in a form relevant to our framework

$$\begin{aligned} &\text{Minimize: } P_{\text{SEC}}(\bar{c}, \bar{d}, p) \\ &\text{Subject to: } E[(\hat{e}_n - e_n)^2 \,|\, \bar{c}, \bar{d}, p] \le D \end{aligned} \tag{42}$$

where D is the desired RMSE given as a system requirement, and $P_{\text{SEC}}(\bar{c}, \bar{d}, p)$ is the power consumed in the SEC unit. In the following, we search for a feasible solution to this problem with respect to $\bar{c}, \bar{d}$, and p.

B. Static Power-Optimum Design

In this subsection, we find the selection of observations and their precisions to optimize the constrained objective function in (42).

1) Observation Selection: An important result from the previous section is in that the algorithmic performance of the system depends on the MSE of the MMSE estimator, or σ_r^2. Taking more observations leads to smaller σ_r^2, or a better quality soft error estimate; however, it will require higher complexity, and thus higher power. Therefore, we need to limit the number of observations, and more systematically, we should select the best combination of observations on the basis of (42). In the following, we will describe an efficient observation selection method that uses a tree search procedure based on the branch and bound principle [17], [18]. Let ξ_m be the set of m samples of observations that are masked by the vector $\bar{c}$ and $\bar{d}$. The variance σ_r^2, based on the observation set ξ_m, is given by

$$\begin{aligned} \sigma_r^2(\xi_m) = \sigma_y^2 - \text{Cov}(y_n, [\xi_m])^T \\ \times \text{Cov}([\xi_m])^{-1} \text{Cov}(y_n, [\xi_m]) \end{aligned} \tag{43}$$

where $[\xi_m]$ is a column vector associated with ξ_m. If the signal subsets, $\xi_1, \ldots, \xi_m$ are nested such that

$$\xi_1 \subset \xi_2 \subset \cdots \subset \xi_m \tag{44}$$

then

$$\sigma_r^2(\xi_1) \ge \sigma_r^2(\xi_2) \ge \cdots \ge \sigma_r^2(\xi_m) \tag{45}$$

since $(y_n - E[y_n \,|\, \xi_{m+1}])$ and $(E[y_n \,|\, \xi_{m+1}] - E[y_n \,|\, \xi_m])$ are orthogonal so that

$$\sigma_r^2(\xi_m) = \sigma_r^2(\xi_{m+1}) + E[(E[y_n \,|\, \xi_m] - E[y_n \,|\, \xi_{m+1}])^2]. \tag{46}$$

On the other hand, the power dissipation P_{SEC}, given $\bar{c}$ and $\bar{d}$, may be estimated via the multiplier energy model [14] or any available power modeling technique [19]. For the nested sets in (44), it is reasonable to assume that

$$P_{\text{SEC}}(\xi_1) \le P_{\text{SEC}}(\xi_2) \le \cdots \le P_{\text{SEC}}(\xi_m) \tag{47}$$

where $P_{\text{SEC}}(\xi_i)$ is the power dissipation associated with ξ_i. The two monotonicity properties in (45) and (47) enable the use of the branch and bound technique to solve for the optimal observation selection.

As a simple example, assume that we select a combination from five candidates, denoted $\{a_1 = x_{q,n}, a_2 = x_{q,n-1}, a_3 = x_{q,n-2}, a_4 = \hat{y}_{q,n-1}, a_5 = \hat{y}_{q,n-2}\}$, which constitutes the full set of observations. We can construct a tree as shown in Fig. 7. The search begins from the root with "no observation," or $[\bar{c}^T, \bar{d}^T] = [0, \ldots, 0]^T$ and traverses down when adding each "new observation." Beginning from the rightmost branch, we calculate P_{SEC} and the resulting RMSE and save them at each branch (see "A" in Fig. 7). We continue traversing the tree while decreasing RMSE, and stop when the RMSE begins to be less than D. We call the first node at which the traversal stops the *initial best subset* and the corresponding P_{SEC} a *bound*. Next, we search the left-side branches and cut off any branches and subbranches whose P_{SEC}'s are larger than the current bound. However, if a leaf of the tree is reached with no pruning, we replace the bound by the current P_{SEC} and update performance of the branch and bound with the current subset. In this example, the final best subset is $\{x_{q,n}, x_{q,n-2}, \hat{y}_{q,n-2}\}$.

In order to improve the algorithm, it is common to position the "good" signals, i.e., those that cause significant RMSE decrease, to the right-hand side of the tree (note the signals at the first level ordered "a_4, a_3, a_2, a_5, a_1"). This ordering reduces the average search path. The initial best subset under this ordering provides a good, but not optimal, solution. Since the initial subset improves successively as the update proceeds, we can terminate the procedure early to obtain a good solution if the search time is excessive.

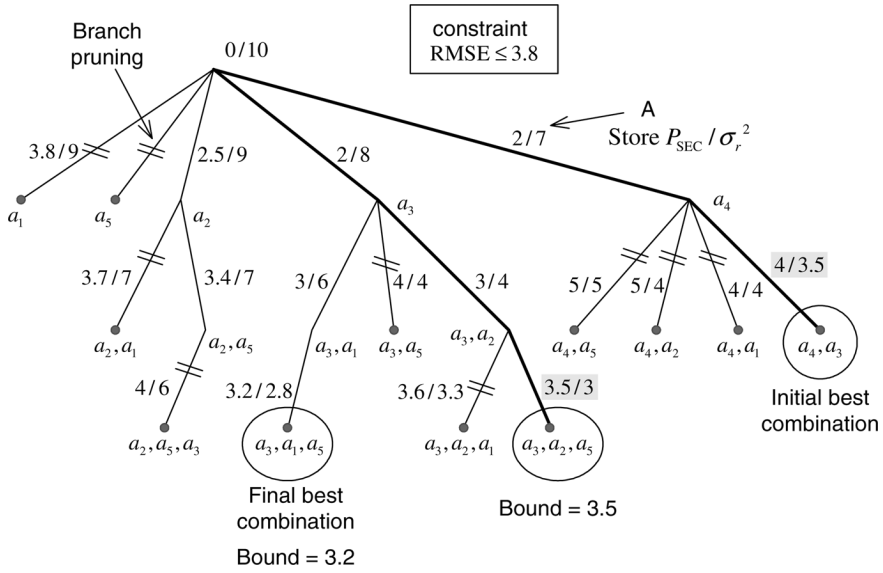

Fig. 7. Observation selection based on the branch and bound technique.

2) Precision Selection: The precision parameter p has a significant impact on the power dissipation of the SEC block. To obtain the jointly optimal values of $\bar{c}$, $\bar{d}$, and p, we can construct the search tree for each value of p over a nominal range. After finding the optimal $\bar{c}$ and $\bar{d}$ for each p, we select the p, which results in the minimum $P_{\rm SEC}$, as the optimal p, and the corresponding $\bar{c}$ and $\bar{d}$ as optimal switching vectors.

C. Dynamic Power-Optimum Configuration

In this subsection, we introduce an automatic power control algorithm that adapts the control vectors $\bar{c}$ and $\bar{d}$ to the variation of input statistics or a given target performance.

1) Adaptive Soft Error Cancellation: To develop a procedure to control the vectors $\bar{c}$ and $\bar{d}$, we need to update the estimator weights $\bar{w}_c$, $\bar{v}_c$ and A automatically, for each $\bar{c}$ and $\bar{d}$. We can employ the least-mean-square (LMS) algorithm, which adapts to minimize the MSE (10) over the weights:

$$w_{c,i}^{(n+1)} = c_i\left(w_{c,i}^{(n)} + \mu\epsilon_n x_{q,n-i+1}\right), \quad \text{for } i = 1,\ldots,N_1 \tag{48}$$

$$v_{c,i}^{(n+1)} = d_i\left(v_{c,i}^{(n)} + \mu\epsilon_n \hat{y}_{q,n-i}\right), \quad \text{for } i = 1,\ldots,N_2 \tag{49}$$

$$A^{(n+1)} = A^{(n)} + \mu\epsilon_n \tag{50}$$

where $w_{c,i}^{(n)}, v_{c,i}^{(n)}$, and $A^{(n)}$ are the values of $w_{c,i}, v_{c,i}$ and A at time n, respectively, and μ is the step size. We denote ϵ_n as the result of subtracting the MEF output from y_n. Since the value of y_n is not available, we use the current restored output instead of y_n as a training symbol, and hence ϵ_n is given by

$$\epsilon_n = \hat{y}_n + \sum_{i=1}^{N_1} w_{c,i}^{(n)} x_{q,n-i+1} + \sum_{i=1}^{N_2} v_{c,i}^{(n)} \hat{y}_{q,n-i} + A^{(n)}. \tag{51}$$

Whenever there is a change in $\bar{c}$ or $\bar{d}$, the adaptive algorithm begins operation until it converges to the correspondingly optimal weight vector. Since $\bar{c}$ and $\bar{d}$ can power down the update algorithm for each weight, the power consumption of weight update block (WUB) can be reduced depending on $\bar{c}$ and $\bar{d}$. We can now address the control of $\bar{c}$ and $\bar{d}$.

2) Automatic Power Control Algorithm: We assume that the computation of each tap , $w_{c,i}$ or $v_{c,i}$ and its weight update algorithm consumes the same power E_s. Then, the power dissipation in the SEC block is given by

$$P_{\rm SEC} = \left(\sum_{i=1}^{N_1} c_i + \sum_{i=1}^{N_2} d_i\right) E_s + \mathcal{H} \tag{52}$$

where $\mathcal{H}$ includes the power dissipation in the MLD block. This means that as we power down more taps, the power dissipation in SEC will proportionally decrease. Hence, we can rewrite the energy optimization problem (42) by

$$\begin{aligned} &\text{Minimize: } \sum_{i=1}^{N_1} c_i + \sum_{i=1}^{N_2} d_i \\ &\text{Subject to: } \sigma_r^2 \le J \end{aligned} \tag{53}$$

where the RMSE constraint can be replaced by a new constraint $\sigma_r^2 \le J$, since RMSE is an increasing function in σ_r^2. In [14], the authors presented a solution to a similar problem for an adaptive equalizer, using a Lagrange multiplier method under the assumption that the input signals to the adaptive equalizer are white. The solution suggested that the best strategy involves powering down the taps of the equalizer with less contribution to the performance metric if each tap consumes the same energy. Unfortunately, it is considerably more difficult to find the associated control vectors for the case with correlated input. Hence, we adopt the strategy of switching off the taps with the smallest coefficients.

The performance estimate (P-estimate) P_n is monitored in real time and compared with two preset thresholds τ_1 and τ_2, where $\tau_1 > \tau_2$. The P-estimate is computed by averaging the square of the LMS update error ϵ_n^2, i.e.,

$$P_n = (1-\rho)P_{n-1} + \rho\varepsilon_n^2 \tag{54}$$

TABLE I
POWER OPTIMIZATION ALGORITHM

STEP 1	Start with $\bar{c}_0$, and $\bar{d}_0$ preset to yield P-estimate smaller than τ_2.
STEP 2	Wait until the estimator coefficients converge.
STEP 3	Monitor P-estimate : if $P_n - \tau_2 \leq 0$ go to STEP 4, and if $P_n - \tau_1 \geq 0$ go to STEP 5.
STEP 4	For $\lvert w_{c,i}\rvert < \lvert w_{c,j}\rvert\ \forall i \neq j$ and $\lvert v_{c,k}\rvert < \lvert v_{c,l}\rvert$ $\forall k \neq l$, set $c_i = 0$ if $\lvert w_{c,i}\rvert < \lvert v_{c,k}\rvert$, else set $d_k = 0$. Go to STEP 2.
STEP 5	Set $c_i = 1$ or $d_i = 1$ for the mask coefficient subject to the last change. Go to STEP 2.

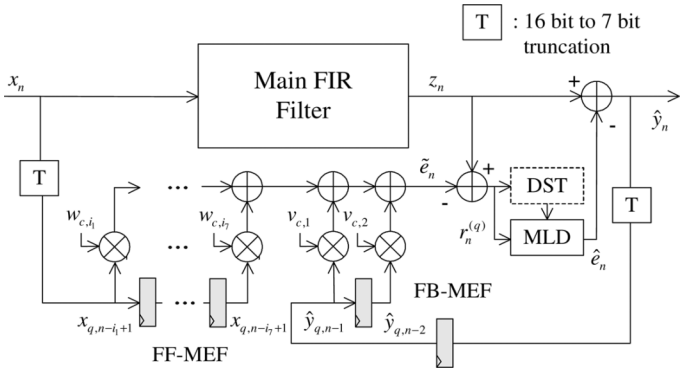

Fig. 8. Hardware flow-graph for an MP-SEC system.

where $0 < \rho < 1$ is a constant for experimental averaging. The automatic power control algorithm (PCA) changes $\bar{c}$ and $\bar{d}$ only when the P-estimate is larger than τ_1 or smaller than τ_2. Starting from the initial setup, $\bar{c}_0$ and $\bar{d}_0$, which are preset to provide a small P-estimate, we set $c_i = 0$ or $d_i = 0$ if the P-estimate is smaller than τ_2, where i corresponds to smallest coefficients. On the contrary, we set $c_i = 1$ or $d_i = 1$, if P_n exceeds τ_1 in the reverse order. Once the control vector $\bar{c}$ or $\bar{d}$ has been changed, the PCA waits until the adaptive algorithm converges. When the P-estimate lies within $[\tau_2, \tau_1]$, we power down the PCA and keep monitoring the P-estimate to detect any changes. This procedure is summarized in Table I.

IV. DISCUSSIONS

In this section, we discuss a hardware design of an MP-SEC system and present a simulation framework and some results.

A. Hardware Design

In Fig. 8, we depict an implementation of the MP-SEC system, which protects a 26-tap FIR frequency selective filter. The system is designed via the static design methodology described in Section III-B-1). It should be noted that the critical path delay of the MEF is shorter than the main filter due to its reduced complexity, and therefore soft errors cannot occur in the MEF for values of up to $k_{\text{VOS}} = 0.6$.

Since the detection rule (25) requires the use of a multiplier, it is desirable to further simplify the detector structure complexity. Under H_1, it follows that $\tilde{e}_n^2 \gg (\tilde{e} - \hat{e}_n)^2$, and under $H_0, (\tilde{e} - \hat{e}_n)^2$ is close to zero. As such, the detection rule can be simplified to

$$|\tilde{e}_n| < \sqrt{2\gamma\sigma_r^2}. \tag{55}$$

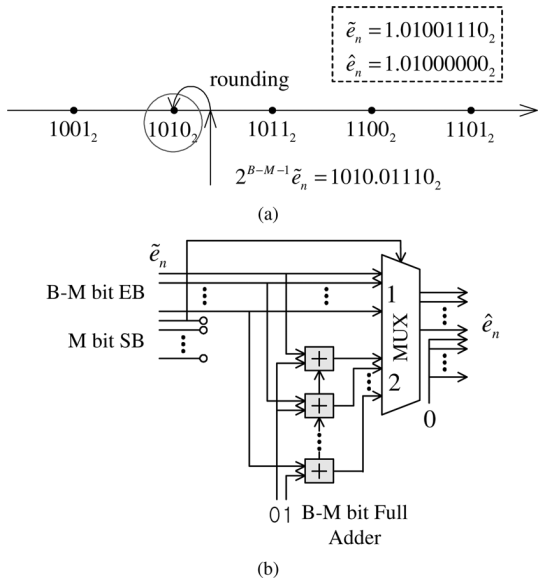

Fig. 9. Implementation of MLD block.

This rule is called a double-sided test (DST), which compares $\tilde{e}_n$ to both positive and negative thresholds. This rule can be expressed as

$$I_b(\tilde{e}_n)\tilde{e}_n - \sqrt{2\gamma\sigma_r^2} \underset{H_1}{\overset{H_2}{\gtrless}} 0 \tag{56}$$

where $I_b(\cdot)$ outputs 1 for a positive input and -1, otherwise. As a result, the hardware design of the detector in (56) requires only one inverter for computing $I_b(\cdot)$ and one subtractor.

Next, consider the hardware design of the MLD block. We can write (18) as

$$\hat{e}_n = \frac{\text{round}\left(2^{B-M-1} \cdot \tilde{e}_n\right)}{2^{B-M-1}}. \tag{57}$$

Fig. 9(a) illustrates the computation of $\hat{e}_n$ based on $\tilde{e}_n$, when $B = 9$ and $M = 5$. Fig. 9(b) describes its implementation. The MLD block can be implemented with only one MUX and one full adder.

Next, we depict an SEC system that employs dynamic PCA in Fig. 10. To implement the power control block in hardware, we devise the stack-based power control which memorizes the tap-positions of the inactive MEF coefficients in the stack on a first-input last-output (FILO) basis. When the P-estimate drops below τ_2, the tap position of the smallest MEF weight enters the stack deactivating the coefficient. When the P-estimate rises above τ_1, the tap at the top of stack is released. Though the hardware implementation including WUB and PCA block appear complicated, the algorithm powering down the SEC subblocks dramatically reduces power consumption.

Table II summarizes the number of basic arithmetic units required for the SEC block depicted in Fig. 8, compared with that of main filter when the PCA is not employed. Note that the hardware complexity of the SEC block is simple, compared to that of main filter.

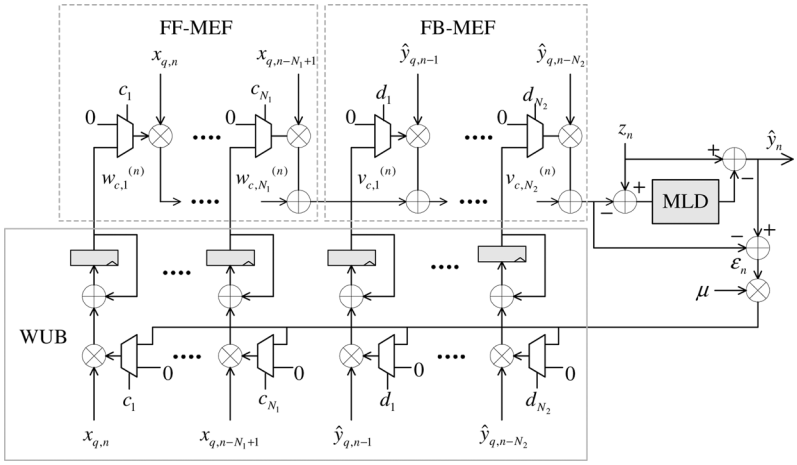

Fig. 10. Adaptive SEC block.

TABLE II
HARDWARE UNITS FOR MP-SEC SYSTEM

Block	Sub-blocks	Necessary units	Number of full adders
Main filter		26 multipliers (16 bit), 27 adders (21 bit)	7223
SEC Block	MEF	9 multipliers (7 bit), 8 adders (12 bit)	537
	DST	1 adder (5 bit), 1 inverter	5
	MLD	1 MUX, 1 adder (4 bit)	4
	ect	2 adders (21 bit)	42

TABLE III
SOFT ERROR RATE, SPACING AND SNR DEGRADATION VERSUS VOS FACTOR

$k_{\rm VOS}$	1.0	0.9	0.8	0.7	0.6
Soft error rate, $P(H_1)$ (%)	0%	2.69%	8.13%	29.88%	54.55%
Number of SBs, M	None	13	12	8	7
Output SNR (dB)	22.91	−0.16	−4.34	−8.24	−10.92

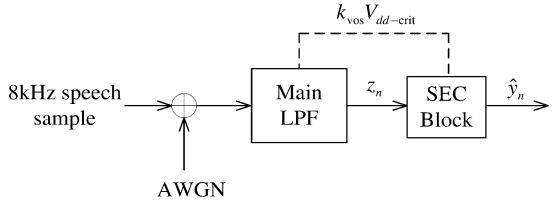

Fig. 11. Simulation setup.

B. Energy Saving Measure

The energy savings $E_{\rm sav}(\%)$ of the ML-EC system is defined by

$$E_{\rm sav} = \frac{P_{\rm ORG} - P_{\rm VOS}}{P_{\rm ORG}} \times 100 \tag{58}$$

where $P_{\rm ORG}$ and $P_{\rm VOS}$ are the power dissipation before and after applying MP-SEC technique, respectively. Note that the power required by SEC block, $P_{\rm SEC}$ is included in calculating $P_{\rm VOS}$.

C. Simulation Setup

The simulation setup used in our simulations is illustrated in Fig. 11. The context that we chose for experiments is a low-pass FIR filter (LPF) that removes the out-of-band noise in front of a speech recognizer. A sequence of 10 000 speech samples of bandwidth 8 kHz are filtered to remove the out-of-band corruption from additive white Gaussian noise (AWGN). A 25-tap linear-phase FIR LPF with cut-off frequency $\pi/4$ is used as the main filter to be protected.

We assume a 0.25-μm, 2.5-V CMOS process technology and that 16×16 bit Baugh–Wooley multipliers [16] are used in the main filter. We also assume that the supply voltage, set to 2.5 V, is scaled down by $k_{\rm VOS} = 0.9, 0.8, 0.7$, and 0.6. First, we compute the logic gate delay for each $k_{\rm VOS}$ via a circuit-level simulator, HSPICE, and obtain the worst path delays reaching the intermediate bits of each processing unit via a logic-level simulator. The intermediate bits with larger path delay than the sampling period exhibit a timing violation, thereby causing a soft error. The power dissipation of the system is obtained via a gate-level power simulation tool, MED [20].

D. Simulation Results

Table III tabulates the soft error rate, number of SBs, M, and output SNR versus $k_{\rm VOS}$, when the speech samples are used as the input. As $k_{\rm VOS}$ decreases, the error rate increases, and the number of SBs decreases or equivalently λ decreases. The original output SNR before applying VOS was 22.91 dB, but the system experiences a catastrophic SNR drop with $k_{\rm VOS}$.

Fig. 12(a) shows the original 400 samples of the desired speech signal y_n. The output signal, which is corrupted by soft errors when $k_{\rm VOS} = 0.7$, is shown in Fig. 12(b). We employ the MP-SEC unit to restore the degraded signal. The predesigned 13-tap FF-MEF and 1-tap FB-MEF are employed to meet the target SNR, or 22 dB. The signal $\tilde{e}_n$ and restored $\hat{y}_n$ are shown in Fig. 12(c) and (d). Note that the MEF modifies the noisy signal to $\tilde{e}_n$ to readily estimate e_n. The resulting estimation error $\hat{y}_n - y_n$ after error correction is shown in Fig. 12(e).

Table IV summarizes the design specifications of the MP-SEC system for each $k_{\rm VOS}$ when the power-optimum design strategy described in Section III-B-1) is employed. The table also includes the resulting SNR and achieved power savings.

TABLE IV
DESIGN SPECIFICATION AND ENERGY SAVINGS OF ENERGY-MINIMUM MP-SEC SYSTEM

$k_{\rm vos}$	FF-MEF Length (N_1')	FB-MEF Length (N_2')	p	SNR Achievement	$E_{\rm sav}$ (%)	$P_{\rm SEC}/P_{\rm VOS}$
0.9	2 tap	2 tap	4 bit	22.91 dB	17.76%	2.10 %
0.8	4 tap	2 tap	6 bit	22.91 dB	33.40%	4.69 %
0.7	13 tap	1 tap	7 bit	22.46 dB	42.23%	14.03 %
0.6	15 tap	1 tap	9 bit	22.08 dB	57.57%	23.30 %

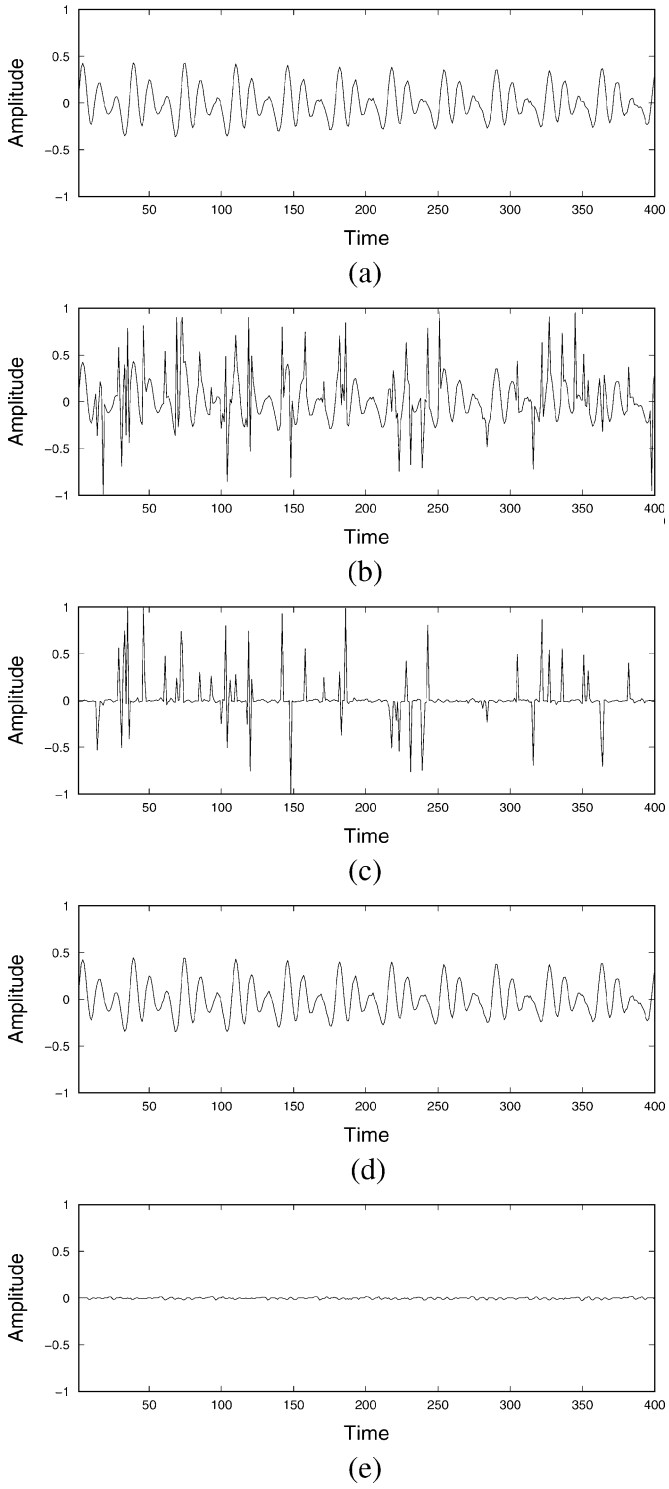

Fig. 12. Shown are 400 samples of (a) the clean output y_n, (b) the noisy output z_n, (c) $\bar{e}_n$, (d) the corrected output $\hat{y}_n$, and (e) the estimation error $\hat{y}_n - y_n$, when $k_{\rm vos}$ is set to 0.7.

When $k_{\rm vos}$ is 0.9 and 0.8, the degraded SNR is completely restored with at most six coefficients, and 6-bit precision, since wide soft error spacing allows for relatively loose MSE requirements for error correction. As $k_{\rm vos}$ is scaled to 0.7 and then to 0.6, we can achieve power savings of 57%, with at most a 0.83-dB SNR loss. This is why the quadratic effect of $k_{\rm vos}$ on power dissipation becomes dominant even when the complexity overhead in the SEC block is increased.

Table V tabulates the design specifications and the resulting energy savings depending on various input signal and filter bandwidths. To generate the input with a particular bandwidth, random white noise is shaped by a linear filter with the given cut-off frequency. The SEC block is designed to allow at most 1-dB SNR loss. For brevity, we present the result only for $k_{\rm vos} = 0.7$, since similar behavior trends were observed when $k_{\rm vos}$ was set to 0.6, 0.8, and 0.9. The higher the input bandwidth, the less power savings. However, the decrease is slight, and we gain relatively consistent power savings for all input bandwidths. Furthermore, the filter bandwidths hardly appear to influence the power savings, which may be a desirable feature compared with the prediction-based error-correction technique [9].

Fig. 13 shows the nature of the adaptation performed by the automatic PCA described in Section III-C-2) to changes in the input signal characteristics. Fig. 13(a) contains a plot of the input signal x_n. The input x_n exhibits a dramatic change in statistics at time 10 000 and 20 000. Specifically, over the intervals, [0, 10 000], [10 000, 20 000], and [20 000, 30 000], the signal has bandwidth of $0.2\pi, 0.5\pi$ and 0.9π and variance of $0.04, 0.0025$, and 0.06, respectively. Fig. 13(b) shows the P-estimate in (54) when ρ is set to 0.999. The evolution of the number of powered-up coefficients is plotted in Fig. 13(c). The largest power savings is achieved during the interval, between the samples of 10 000 and 20 000, where only seven coefficients are used for soft error cancellation. Note that the coefficient adaptation and power control operate only when the P-estimate remains outside of the range $[\tau_2, \tau_1]$. Hence, when the P-estimate lies within $[\tau_2, \tau_1]$, the power overhead of automatic PCA comes from the task of computing the P-estimate and comparing it with the thresholds.

Fig. 14 compares the performance and power trade-off of MP-SEC with those of the prediction-based error-correction [9] and reduced precision replica [10] techniques, when the speech samples are used as the system input. The MP-SEC technique yields better SNR performance over the range of 0% to 50% power savings, and is 9 dB better than the prediction-based error-correction method and 1 dB better from the reduced precision replica method at 40% power savings. When employing

TABLE V
DESIGN SPECIFICATION AND ENERGY SAVINGS DEPENDING ON INPUT AND FILTER BANDWIDTH ($k_{\rm vos} = 0.7$)

Filter bandwidth	Input bandwidth	FF-MEF length	FB-MEF length	p	SNR loss (dB)	Energy savings (%)
0.2π	0.2π	14	2	5 bit	0.99	47.13
	0.4π	18	1	5 bit	0.79	46.18
	0.6π	20	1	5 bit	1.00	45.94
	0.8π	21	1	5 bit	0.86	45.71
0.4π	0.2π	6	3	6 bit	0.76	48.00
	0.4π	12	1	5 bit	0.97	47.84
	0.6π	10	3	6 bit	0.82	46.70
	0.8π	21	1	5 bit	0.94	45.71
0.6π	0.2π	6	1	6 bit	0.86	48.66
	0.4π	7	1	5 bit	0.97	49.04
	0.6π	7	3	6 bit	0.98	47.68
	0.8π	12	3	5 bit	0.95	45.37
0.8π	0.2π	5	1	6 bit	0.94	48.99
	0.4π	5	3	6 bit	0.68	48.33
	0.6π	12	2	5 bit	0.90	47.61
	0.8π	13	1	5 bit	0.93	47.61

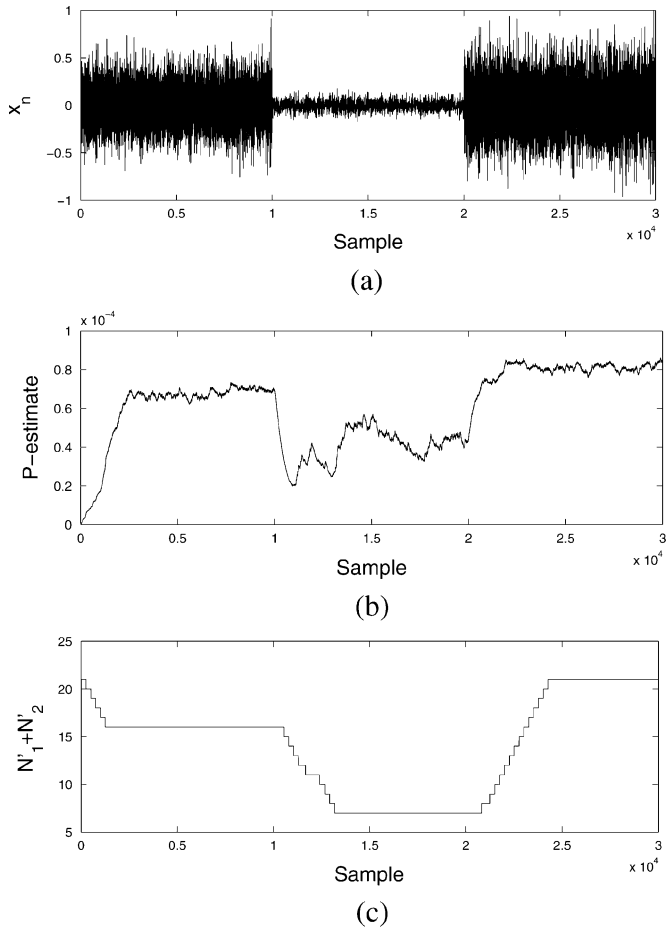

Fig. 13. Shown are (a) the signal x_n, (b) the P-estimate, and (c) $N_1' + N_2'$.

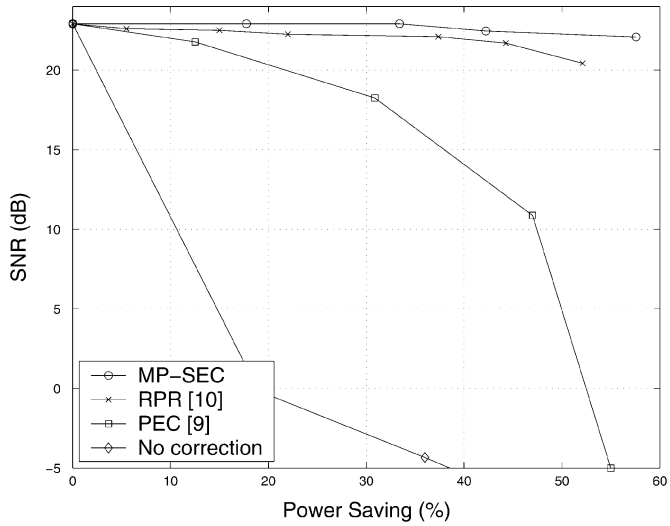

Fig. 14. Performance and power tradeoff of MP-SEC, precision-based error-correction [9] and reduced precision replica [10].

the PCA, the MP-SEC achieves 8% power savings with no SNR loss and 15 % power savings within 1 dB SNR loss.

V. CONCLUSION

In this paper, we have addressed two problems: 1) estimation and detection of soft errors induced by VOS in low-power digital filtering and 2) an approach to energy minimum design and adaptive power control for reconfiguration. Our derivation of the soft error estimator and detection algorithm is based on the observation that soft errors are created by higher order bits in the representation of the signals being processed. As such, through tracking the signal correlation over time, and using a reduced-precision replica of the filtering operation to be protected, such high-order bit errors can be readily detected and ultimately corrected. Through a low-power implementation of the error detection and correction unit, the MP-SEC approach shows promise for achieving significant power savings for digital filtering as well as a variety of DSP applications.

REFERENCES

[1] A. Chandrakasan and R. W. Brodersen, "Minimizing power consumption in digital CMOS circuits," *Proc. IEEE*, vol. 83, no. 4, pp. 498–523, Apr. 1995.
[2] S. Iman and M. Pedram, "An approach for multi-level logic optimization targeting low-power," *IEEE Trans. Comput.-Aided Design*, vol. 15, pp. 889–901, Aug. 1996.
[3] J. T. Ludwig, S. H. Nawab, and A. P. Chandrakasan, "Low-power digital filtering using approximate processing," *IEEE J. Solid-State Circuits*, vol. 31, pp. 395–400, Mar. 1996.
[4] R. Gonzalez, B. Gordon, and M. Horowitz, "Supply and threshold voltage scaling for low-power CMOS," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1210–1216, Aug. 1997.

[5] G. Y. Wei and M. Horowitz, "A low power switching power supply for self-clocked systems," in *IEEE Int. Symp. Low Power Electronics Design*, Aug. 1996, pp. 313–317.
[6] V. Gutnik and A. Chandrakasan, "Embedded power supply for low-power DSP," *IEEE Trans. VLSI Syst.*, vol. 5, pp. 425–435, Dec. 1997.
[7] J. M. Chang and M. Pedram, "Energy minimization using multiple supply voltages," *IEEE Trans. VLSI Syst.*, vol. 5, pp. 1–8, Dec. 1997.
[8] N. Chabini and W. Wolf, "Reducing dynamic power consumption in synchronous sequential digital designs using retiming and supply voltage scaling," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 12, pp. 573–589, Jun. 2004.
[9] R. Hegde and N. Shanbhag, "Soft digital signal processing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 9, no. 6, pp. 813–823, Dec. 2001.
[10] B. Shim, S. Sridhara, and N. Shanbhag, "Reliable low-power digital signal processing via reduced precision redundancy," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 12, pp. 497–510, May 2004.
[11] L. Wang and N. Shanbhag, "Low-power filtering via adaptive error-cancellation," *IEEE Trans. Signal Process.*, vol. 51, pp. 575–583, Feb. 2003.
[12] M. E. Austin, "Decision-feedback equalization for digital communication over dispersive channels," Research Laboratory of Electronics, Mass. Inst. Technol., Cambridge, MA, Tech. Rep. 461, Aug. 1967.
[13] P. Monsen, "Adaptive equalization of the slow fading channel," *IEEE Trans. Commun.*, vol. COM-22, no. 8, pp. 1064–1075, Aug. 1974.
[14] M. Goel and N. Shanbhag, "Dynamic algorithm transformations for low-power reconfigurable adaptive equalizer," *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 2821–2832, Oct. 1999.
[15] M. Goel and N. Shanbhag, "Dynamic algorithm transformations (DAT)—A systematic approach to low-power reconfigurable signal processing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 7, pp. 463–476, Dec. 1999.
[16] C. R. Baugh and B. A. Wooley, "A two's complement parallel array multiplication algorithm," *IEEE Trans. Comput.*, vol. C-22, pp. 1045–1047, Dec. 1973.
[17] P. M. Narendra and K. Fukunaga, "A branch and bound algorithm for feature subset selection," *IEEE Trans. Comput.*, vol. 26, pp. 917–922, Sep. 1977.
[18] P. Somol, P. Pudil, and J. Kittler, "Fast branch & bound algorithms for optimal feature selection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, pp. 900–912, Jul. 2004.
[19] G. Gupta and F. N. Najm, "Power modeling for high-level power estimation," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 8, pp. 18–29, Feb. 2000.
[20] F. N. Najm, "A survey of power estimation techniques in VLSI circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 2, pp. 446–455, Dec. 1994.
[21] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer, 1994.
[22] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.

Jun Won Choi (S'04) was born in Seoul, Korea, in 1976. He received the B.S. and M.S. degrees in electrical and computer engineering from Seoul National University (SNU), Seoul, Korea, in 2000 and 2002, respectively. He is currently working towards the Ph.D. degree at the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign.

He was a Research Assistant at SNU from 2002 until 2003. His research interests include communication signal processing, adaptive filtering, low-power DSP design, and wireless and space–time communications.

Byonghyo Shim received the B.S. and M.S. degrees in control and instrumentation engineering from Seoul National University, Seoul, Korea, in 1995 and 1997, respectively, and the M.A. degree in mathematics and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign, in 2004 and 2005, respectively.

From 1997 to 2000, he was with Department of Electronics Engineering at the Korean Air Force Academy as an officer and an academic full-time instructor. He also had a short-time research position in the DSP Group of LG Electronics and the DSP R&D Center of Texas Instruments in 1997 and 2004, respectively. In January 2005, he joined Qualcomm, Inc., San Diego, CA, where he is currently working on wireless CDMA systems. His research interests encompass a wide spectrum of digital signal processing including signal processing for communication, statistical signal processing, low-power and reliable signal processing, and also estimation and detection.

Dr. Shim received the 2005 M. E. Van Valkenburg Research Award from the Electrical and Computer Engineering Department of the University of Illinois and Tenth Samsung Humantech Paper Award in 2004. He is a member of Sigma Xi and Tau Beta Pi.

Andrew C. Singer (S'92–M'95–SM'05) was born in Akron, OH, in 1967. He received the S.B., S.M., and Ph.D. degrees, all in electrical engineering and computer science, from the Massachusetts Institute of Technology (MIT), Cambridge, in 1990, 1992, and 1996, respectively.

Since 1998, he has been with the faculty of the Department of Electrical and Computer Engineering (ECE), University of Illinois, Urbana-Champaign, where he is a Willet Faculty Scholar and an Associate Professor with the ECE and a Research Associated Professor with the Coordinated Science Laboratory. In 1996, he was a Postdoctoral Research Affiliate with the Research Laboratory of Electronics, MIT. From 1996 to 1998, he was a Research Scientist with Sanders, A Lockheed Martin Company, Manchester, NH. His research interests include statistical signal processing and communication, universal prediction and data compression, and machine learning.

Dr. Singer was a Hughes Aircraft Masters Fellow and was the recipient of the Harold L. Hazen Memorial Award for excellence in teaching in 1991. In 2000, he received the National Science Foundation CAREER Award, and in 2001, he received the Xerox Faculty Research Award. He is currently a member of the MIT Educational Council, Eta Kappa Nu, and Tau Beta Pi.

Nam Ik Cho (S'86–M'92) received the B.S., M.S., and Ph.D. degrees in control and instrumentation engineering from Seoul National University, Seoul, Korea, in 1986, 1988, and 1992, respectively.

From 1991 to 1993, he was a Research Associate of the Engineering Research Center for Advanced Control and Instrumentation, Seoul National University. From 1994 to 1998, he was with the University of Seoul, Seoul, Korea, as an Assistant Professor of Electrical Engineering. He joined the School of Electrical Engineering, Seoul National University, in 1999, where he is currently a Professor. His research interests include speech, image, video signal processing, and adaptive filtering.