APPROXIMATE ML ESTIMATOR FOR COMPENSATION OF TIMING MISMATCH AND JITTER NOISE IN TI-ADCS

*Hesam Araghi, Mohammad Ali Akhaee, Arash Amini**

College of Engineering, University of Tehran, Tehran, Iran
* Department of Electrical Eng., Sharif University of Technology
Email: {h.araghi, akhaee}@ut.ac.ir, aamini@sharif.ir

ABSTRACT

Time-interleaved analog to digital converters (TI-ADC) offer high sampling rates by passing the input signal through C parallel low-rate ADCs. We can achieve C-times the sampling rate of a single ADC if all the shifts between the channels are identical. In practice, however, it is not possible to avoid mismatch among shifts. Besides, the samples are also subject to jitter noise. In this paper, we propose a blind method to mitigate the joint effects of sampling jitter and shift mismatch in the TI-ADC structure. We assume the input signal to be bandlimited and incorporate the jitter via a stochastic model. Next, we derive an approximate model based on a first-order Taylor series and use an iterative maximum likelihood estimator to reconstruct the uniform samples of the input signal. The simulation results show that with a slight increase in the mean square-error, we obtain a fast blind compensation algorithm.

Index Terms— Bandlimited signals, jitter noise, maximum likelihood estimation, mismatch compensation, time-interleaved ADC

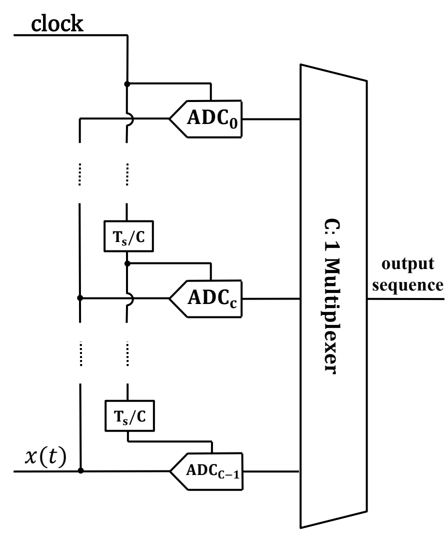

Fig. 1. Time-interleaved ADC structure (taken from [13]).

1. INTRODUCTION

The time-interleaved ADC (TI-ADC) is a low-cost solution for achieving high sampling rates [1, 2]. This structure contains C ADC channels, each with a sampling rate of f_s, yielding an overall rate of $C\,f_s$. However, this structure suffers from mismatch error between channels which can potentially degrade the performance of TI-ADC [3, 4]. Due to its nonlinear effect, the interchannel timing mismatch is known to have the most detrimental effect among the channel mismatches. As a result, there has been numerous attempts to estimate and compensate this timing mismatch; the reader is referred to [5–9] for some of the blind techniques that operate at standard conditions without requiring training signal for calibration.

Jitter noise is an additional non-ideality in ADCs in general. Simply, it causes unpredictable but small deviation in the sampling instances from their nominal periodic values [3, 10]. To compensate this effect, the deviations need to be estimated first. A Bayesian approach based on the Gibbs sampling technique has been devised in [11] to estimate the timing deviations in a single channel ADC. However, the method demands high computational complexity, which makes it unsuitable for on-line applications. An alternative with low computational cost is proposed in [12] that estimates the uniform samples of the signal using an iterative linear estimator. The method is able to mitigate small amount of jitter noise. Recently, we extended the method of [11] to a multichannel framework in [13] for compensating both the interchannel timing mismatch and the jitter noise. Similar to [11], the algorithm is suitable for post-processing applications where the computational cost is not of major concern. In this paper, we propose a computationally efficient alternative to estimate the uniform samples of a bandlimited signal captured by a TI-ADC suffering from both the jitter and the timing mismatch with a slight performance loss.

2. ACQUISITION MODEL

In this section, we first explain the signal acquisition structure in a TI-ADC setup and the input signal model in separate subsections. Then, we present a first-order Taylor series approximation to incorporate the timing mismatches into our model.

2.1. TI-ADC structure

A time-interleaved ADC consists of C parallel channels that each operate at an oversampling ratio M with $M \in \mathbb{N}$ (Figure 1). Thus, the total oversampling ratio equals $M \times C$. The ideal scenario for a TI-ADC corresponds to a setup in which the sampling instances of the cth channel occur at

$$\mathcal{T}_c = \big\{m\, T_s + c\tfrac{T_s}{C}\big\}_{m\in\mathbb{Z}}\,, \tag{1}$$

where $T_s = 1/f_s$ is the sampling period in each channel, and $0 \le c \le C-1$. In this ideal setting, the obtained samples are the same as a single channel ADC with C times the sampling rate. In practice, however, the sampling instances of the cth channel differ from the ideal values by a constant timing mismatch t_c, and the random jitter $z_c[m]$. Hence, the more realistic model for the sampling instances is given by

$$\mathcal{T}_c{}' = \big\{m\, T_s + c\tfrac{T_s}{C} + z_c[m] + t_c\big\}_{m\in\mathbb{Z}}\,, \tag{2}$$

where we consider $t_0 = 0$ as a reference, and assume $\{t_c\}_{c=1}^{C-1}$ and $\{z_c[m]\}_{c,m}$ are independent zero-mean Gaussian random variables with variance σ_t^2 and σ_z^2, respectively. The output of channel c, i.e. $y_c[m]$, is then formulated as

$$y_c[m] = x\big(m\, T_s + c\tfrac{T_s}{C} + z_c[m] + t_c\big) + w_c[m]. \tag{3}$$

Here, w_c stands for the joint effects of thermal and quantization noises, which is modeled by a white Gaussian noise with variance σ_w^2, and is independent of timing mismatches. In the rest of this article, we assume that the values of σ_t^2, σ_z^2, and σ_w^2 are known in advance.

2.2. Input signal model

Suppose the input signal $x(t)$ is a deterministic bandlimited signal with the cut-off frequency $f_c/2$. Without loss of generality, we focus on $f_c = 1$. Thus, $x(t)$ can be written as

$$x(t) = \sum_{k\in\mathbb{Z}} x(k)\ \mathrm{sinc}\big(t-k\big), \tag{4}$$

where $\{x(k)\}_k$ are the uniform samples of $x(t)$ at the Nyquist rate, and $\mathrm{sinc}(t) = \sin(\pi t)/\pi t$. For $T_s = \frac{1}{M}$, i.e., oversampling ratio of M at each channel, (3) can be rewritten as

$$y_c[m] = \sum_{k\in\mathbb{Z}} x(k)\ \mathrm{sinc}\big(\tfrac{m}{M} + \tfrac{c}{M\,C} + z_c[m] + t_c - k\big) + w_c[m]. \tag{5}$$

Our goal is to estimate the uniform samples $x(k)$ of x based on the set of samples $\{y_c[m]\}_{c=0}^{C-1}$ provided by the C channels of TI-ADC. To this end, we truncate the summation in (5) to $0 \le k \le (K-1)$, and use the samples $\{y_c[m]\}_{c=0}^{C-1}$ with $0 \le m \le (M\,K-1)$.More precisely, the output signal $y_c[m]$ is approximated as

$$y_c[m] \cong \sum_{k=0}^{K-1} x(k)\,\mathrm{sinc}\big(\tfrac{m}{M} + \tfrac{c}{M\,C} + z_c[m] + t_c - k\big) + w_c[m], \tag{6}$$

for $0 \le m \le (M\,K-1)$.

2.3. Taylor approximation of the system model

Oftentimes, the timing mismatches $z_c[m] + t_c$ are small compared to the sampling period $T_s = \frac{1}{M}$. This allows us to apply the first-order Taylor series approximation in (6), which describes the effect of timing mismatches in a linear way:

$$\begin{aligned} y_c[m] \approx &\sum_{k=0}^{K-1} x(k)\mathrm{sinc}'\big(\tfrac{m}{M} + \tfrac{c}{M\,C} - k\big)\big(z_c[m] + t_c\big) \\ &+ \sum_{k=0}^{K-1} x(k)\,\mathrm{sinc}\big(\tfrac{m}{M} + \tfrac{c}{M\,C} - k\big) + w_c[m], \end{aligned} \tag{7}$$

where $\mathrm{sinc}'(t) = \frac{d}{dt}(\sin(\pi t)/\pi t)$. For the sake of simplicity, we express (7) in a matrix form as

$$\mathbf{y}_c \approx \mathbf{H}_c\,\mathbf{x} + t_c\,\mathbf{H}_c'\,\mathbf{x} + \mathbf{Z}_c\,\mathbf{H}_c'\,\mathbf{x} + \mathbf{w}_c, \tag{8}$$

where

$$\begin{aligned} \mathbf{y}_c &= [y_c[0], \ldots, y_c[N_c-1]]^T, \\ \mathbf{x} &= [x[0], \ldots, x[K-1]]^T, \\ \mathbf{H}_c[m,k] &= \mathrm{sinc}(\tfrac{m}{M} + \tfrac{c}{M\,C} - k), \\ \mathbf{H}_c'[m,k] &= \mathrm{sinc}'(\tfrac{m}{M} + \tfrac{c}{M\,C} - k), \end{aligned}$$

and $\mathbf{Z}_c$ is an $N_c \times N_c$ diagonal matrix with $\mathbf{Z}_c[m,m] = z_c[m]$ for $0 \le m \le N_c - 1$, $0 \le k \le K-1$, and $N_c = M\,K$.

Since $w_c[m]$, $z_c[m]$, and t_c are independent Gaussian random variables, vector $\mathbf{y_c}$ in (8) is also a Gaussian vector [14]:

$$\mathbf{y}_c \sim \mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_{cc}), \tag{9}$$

with the mean being

$$\boldsymbol{\mu}_c = \mathbf{H}_c\,\mathbf{x}, \tag{10}$$

and the covariance matrix being

$$\begin{aligned} \boldsymbol{\Sigma}_{cc} &= \mathbb{E}[(\mathbf{y}_c - \mathbf{H}_c\mathbf{x})(\mathbf{y}_c - \mathbf{H}_c\mathbf{x})^T] \\ &= \mathbb{E}[\mathbf{Z}_c(\mathbf{H}_c'\mathbf{x})(\mathbf{H}_c'\mathbf{x})^T\mathbf{Z}_c] + \sigma_t^2\,(\mathbf{H}_c'\mathbf{x})(\mathbf{H}_c'\mathbf{x})^T + \sigma_w^2\,\mathbf{I}_{N_c} \\ &= \sigma_z^2\,\mathbf{I}_{N_c}\odot(\mathbf{H}_c'\mathbf{x})(\mathbf{H}_c'\mathbf{x})^T + \sigma_t^2\,(\mathbf{H}_c'\mathbf{x})(\mathbf{H}_c'\mathbf{x})^T + \sigma_w^2\,\mathbf{I}_{N_c} \end{aligned} \tag{11}$$

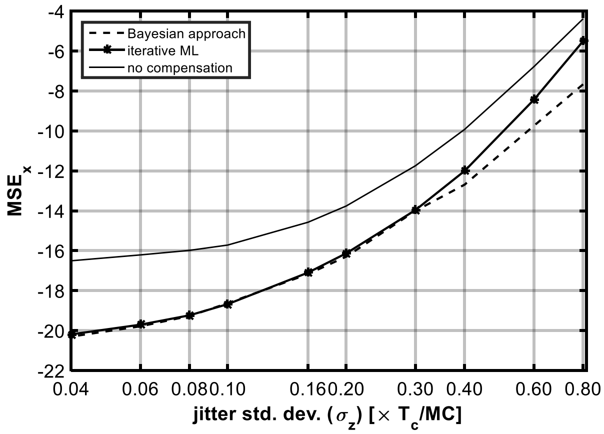

Fig. 2. MSE performance of the iterative ML estimator for different values of σ_z, with $\sigma_w^2 = 0.05^2$, $\sigma_t = 0.2\frac{T_c}{M\,C}$, $M = 1$, and $C = 4$.

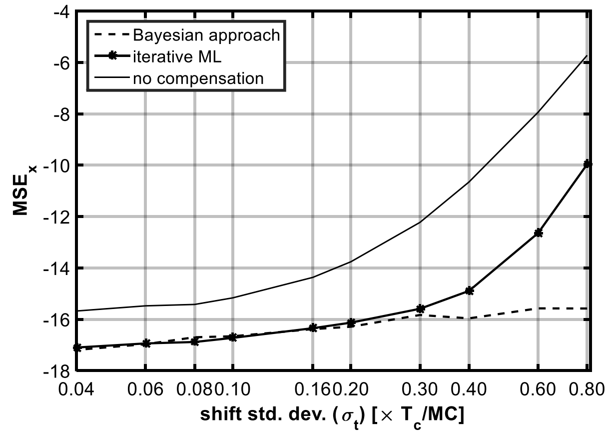

Fig. 3. MSE performance of the iterative ML estimator for different values of σ_t, with $\sigma_w^2 = 0.05^2$, $\sigma_z = 0.2\frac{T_c}{M\,C}$, $M = 1$, and $C = 4$.

where $\odot$ denotes the Hadamard matrix product, $1 \leq c \leq C-1$, and $\mathbf{I}_{N_c}$ is the $N_c \times N_c$ identity matrix. As the vectors $\{\mathbf{y}_c\}_c$ are independent of each other, we can write that

$$p(\mathbf{y};\mathbf{x}) = \prod_{c=0}^{C-1} p(\mathbf{y}_c;\mathbf{x}), \tag{12}$$

where $\mathbf{y} \triangleq [\mathbf{y}_0^T, \ldots, \mathbf{y}_{C-1}^T]^T$. Note that the covariance matrix of the channel 0, i.e. $\mathbf{\Sigma}_{00}$, does not contain the term $\sigma_t^2(\mathbf{H}_c'\mathbf{x})(\mathbf{H}_c'\mathbf{x})^T$, because t_0 is assumed to be zero.

3. MAXIMUM LIKELIHOOD ESTIMATION

We derive an iterative method to approximate the maximum likelihood (ML) estimation of vector $\mathbf{x}$. The ML method tries to find the value of $\mathbf{x}$ that maximizes the likelihood function or equally log-likelihood (LL) function [14]. In our approximated observation model, it is convenient to concentrate the LL function:

$$\begin{aligned}\mathcal{L}(\mathbf{x};\mathbf{y}) \triangleq \ln p(\mathbf{y};\mathbf{x}) &= \sum_{c=0}^{C-1} \ln p(\mathbf{y}_c;\mathbf{x}) \\ &= -\frac{N_c\,C}{2}\ln(2\pi) - \frac{1}{2}\sum_{c=0}^{C-1}\ln(\det(\mathbf{\Sigma}_{cc})) \\ &\quad - \frac{1}{2}\sum_{c=0}^{C-1}(\mathbf{y}_c - \mathbf{H}_c\,\mathbf{x})^T\mathbf{\Sigma}_{cc}^{-1}(\mathbf{y}_c - \mathbf{H}_c\,\mathbf{x}).\end{aligned} \tag{13}$$

Since covariance matrices $\mathbf{\Sigma}_{cc}$s are dependent on vector $\mathbf{x}$, maximizing (13) is not straightforward. We use the similar iterative solution as in [12] to approximate the ML estimator. In this iterative approach, first the algorithm is initialized with

$$\hat{\mathbf{x}} = \Big(\sum_{c=0}^{C-1}\mathbf{H}_c^T\mathbf{H}_c\Big)^{-1}\Big(\sum_{c=0}^{C-1}\mathbf{H}_c^T\mathbf{y}_c\Big), \tag{14}$$

assuming that there is no timing mismatches. Then, we calculate the covariance matrices $\hat{\mathbf{\Sigma}}_{cc}$s by replacing (14) into (11) (the term $\sigma_t^2(\mathbf{H}_c'\mathbf{x})(\mathbf{H}_c'\mathbf{x})^T$ is excluded in computing $\hat{\mathbf{\Sigma}}_{00}$). Now, we fix the $\hat{\mathbf{\Sigma}}_{cc}$s and solve the following linear equation

$$\frac{\partial\mathcal{L}(\mathbf{x};\mathbf{y})}{\partial\mathbf{x}} = 0 \tag{15}$$

to obtain the estimate of $\mathbf{x}$:

$$\hat{\mathbf{x}} = \Big(\sum_{c=0}^{C-1}\mathbf{H}_c^T\hat{\mathbf{\Sigma}}_{cc}^{-1}\mathbf{H}_c\Big)^{-1}\Big(\sum_{c=0}^{C-1}\mathbf{H}_c^T\hat{\mathbf{\Sigma}}_{cc}^{-1}\mathbf{y}_c\Big). \tag{16}$$

By replacing $\hat{\mathbf{x}}$ in (11) an computing the new $\hat{\mathbf{\Sigma}}_{cc}$s, the repetition cycle is completed.

The approximate ML estimator of the vector $\mathbf{x}$ can be represented as the following iterative procedure:

Step 1) compute $\hat{\mathbf{x}}^{(0)}$ from (14).

Step 2) $i = 0$.

Step 3) compute $\hat{\mathbf{\Sigma}}_{cc}^{(i)}$ from (11) ($\hat{\mathbf{\Sigma}}_{00}^{(i)}$ does not have the $\sigma_t^2(\mathbf{H}_c'\hat{\mathbf{x}}^{(i)})(\mathbf{H}_c'\hat{\mathbf{x}}^{(i)})^T$ term).

Step 4) compute $\hat{\mathbf{x}}^{(i+1)}$ from (16).

Step 5) if $\|\hat{\mathbf{x}}^{(i+1)} - \hat{\mathbf{x}}^{(i)}\|_2 < \epsilon$ or $i > \max\_\mathrm{iter}$ return $\hat{\mathbf{x}}^{(i+1)}$ as output of the algorithm, otherwise $i = i + 1$ and go to *Step 3*.

4. SIMULATION RESULTS

In this section, the performance of the iterative method approximating the ML estimator is presented. We model the input vector $\mathbf{x}$ as a zero-mean Gaussian vector with covariance matrix $\sigma_x^2\mathbf{I}_K$ where $\sigma_x = 1$ and $K = 10$. In all the simulations, the variance of the additive noise is assumed to

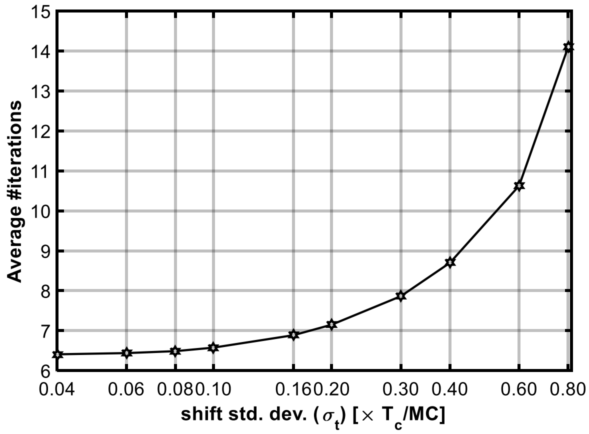

Fig. 4. Average number of iterations that iterative ML estimator needs to $\|\hat{\mathbf{x}}^{(i+1)} - \hat{\mathbf{x}}^{(i)}\|_2 < 10^{-8}$ for different values of σ_t ($\sigma_w^2 = 0.05^2$, $\sigma_z = 0.2\frac{T_c}{MC}$, $M = 1$, and $C = 4$).

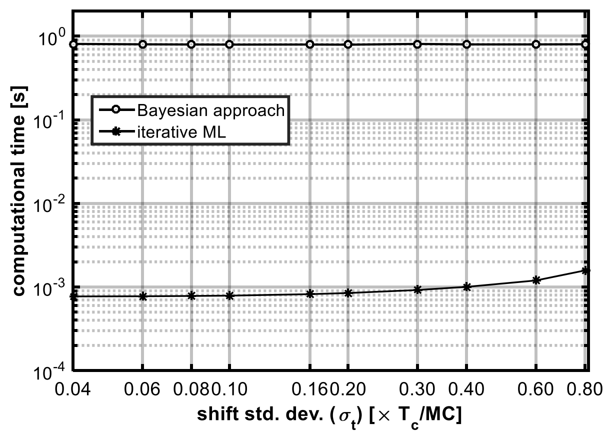

Fig. 5. Comparing average computational time of iterative ML estimator and Bayesian approach method for different values of σ_t ($\sigma_w^2 = 0.05^2$, $\sigma_z = 0.2\frac{T_s}{MC}$, $M = 1$, and $C = 4$).

be $\sigma_w^2 = 0.05^2$ (SNR = 26 dB). The mean square errors of $\mathbf{x}$ ($\mathrm{MSE}_{\mathbf{x}} = \|\hat{\mathbf{x}} - \mathbf{x}\|^2$) are displayed in Figures 2 and 3. To examine the performance of our algorithm, we use the numerical Bayesian approach presented in [13]. This technique utilizes Gibbs sampling which is a Markov chain Monte Carlo algorithm for implementing optimal MMSE estimator. However, the algorithm suffers from high computational time due to the need for great number of iterations. Since the Bayesian approach supposes that the distribution of $\mathbf{x}$ is known a-priori, it is expected to offer a lower MSE bound. Additionally, the performance of the algorithm is compared against the case in which timing mismatches (jitter noise and interchannel timing mismatch) are not compensated, i.e., $\hat{\mathbf{x}}$ is only estimated according to (14). Each point in the figures is the average of 10000 independent runs except for more computationally expensive Bayesian approach, where it is the average of 1000 trials. Also, the stopping criterions are set to $\epsilon = 10^{-8}$ and $\mathrm{max\_iter} = 1000$.

In Figure 2, the performance of the iterative ML estimator together with the Bayesian approach and no compensation case is depicted for different values of σ_z, where $M = 1$, $C = 4$, $\sigma_t = 0.2\frac{T_c}{MC}$, and $T_c = 1/f_c$. We observe that for $\sigma_z < 0.3\frac{T_c}{MC}$ the proposed iterative method has almost the same MSE compared to the Bayesian approach. It is important to note that for $\sigma_z > 0.3\frac{T_c}{MC}$ there is a significant probability that the ordering of the timing instances gets violated (Figure 7), which is unlikely in practice. In Figures 3 to 6, we fix $\sigma_z = 0.2\frac{T_c}{MC}$ and sweep σ_t. Again we can see in Figure 3 that the method performs near the Bayesian approach for $\sigma_t < 0.3\frac{T_c}{MC}$, while keeping a promising distance with the uncompensated curve.

The average number of iterations required for $\|\hat{\mathbf{x}}^{(i+1)} - \hat{\mathbf{x}}^{(i)}\|_2 < \epsilon$ is displayed in Figure 4. For the reasonable range $\sigma_t < 0.3\frac{T_c}{MC}$, the algorithm converges in less than 10 iterations on average. To compare the computational complexities, we measure the running times of the iterative ML algorithm and the numerical Bayesian approach with MATLAB simulations (`tic` and `toc` commands). The employed machine had 8.0 GB of RAM and a Quad-Core processor with the clock speed of 3.40 GHz. Figure 5 shows the average time needed for each algorithm to estimate the vector $\mathbf{x}$. Our method is found to be roughly 1000 times faster than the Bayesian approach.

Figure 6 shows the sensitivity of our method to the precision of the parameters σ_t, σ_z, and σ_w. We considered three cases defined by setting **(i)** $\sigma_t^{(est.)} = 5\sigma_t$, **(ii)** $\sigma_z^{(est.)} = 2\sigma_z$, and **(iii)** $\sigma_w^{(est.)} = 0.5\sigma_w$. Although these cases cause slight performance loss compared to the exact estimation case, there is still a reasonable gap between them and the uncompensated curve, and the method is relatively resistant to the parameter imprecision.

5. CONCLUSION

The TI-ADC is a low-cost analog to digital structure that can achieve high sampling rates. However, its performance is greatly affected by interchannel timing mismatch and sampling jitter. Therefore, it is essential to estimate and compensate the timing nonidealities. In this paper, we proposed a blind yet computationally efficient technique to compensate the jitter noise and the interchannel timing mismatch. The method relies on a first-order Taylor series approximation of the timing mismatches, which simplifies their nonlinear nature into an almost linear form. Then, we employ an iterative ML estimator to obtain the interchannel timing mismatch, and eventually achieve the uniform samples of the signal. Simulation results reveal a significant performance improvement compared to the case where no compensation is applied. Fur-

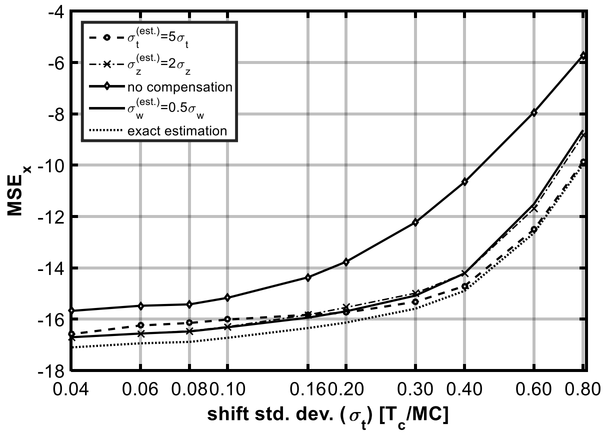

Fig. 6. Performance sensitivity of the iterative ML method to estimation errors for cases $\sigma_t^{(est.)} = 5\sigma_t$, $\sigma_z^{(est.)} = 2\sigma_z$, and $\sigma_w^{(est.)} = 0.5\sigma_w$. ($\sigma_w^2 = 0.05^2$, $\sigma_z = 0.2\frac{T_c}{M\,C}$, $M = 1$, and $C = 4$.

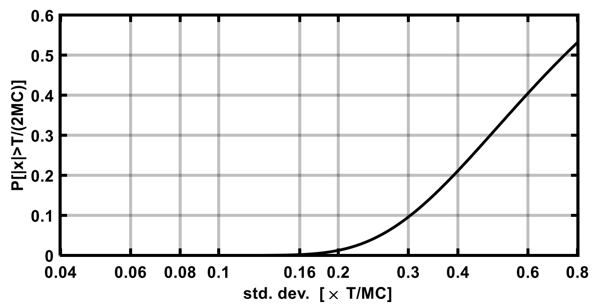

Fig. 7. The probability that $x \sim \mathcal{N}(0, \sigma_x^2)$ exceeds its sampling region ($p[|x| > \frac{T_c}{2M\,C}]$) for different values of σ_x.

thermore, the method tightly follows the performance of the optimal Bayesian approach for small to moderate timing mismatch values. Nevertheless, as the first-order Taylor approximation becomes less sharp for large mismatch values, we expect the performance to deviate from the Bayesian technique at large mismatch variances. Fortunately, in most of the practical cases, such large mismatch values are uncommon.

6. REFERENCES

[1] W. C. Black and D. A. Hodges, "Time interleaved converter arrays," *IEEE J. Solid-State Circuits*, vol. 15, no. 6, pp. 1022–1029, Dec. 1980.

[2] B. Razavi, "Design considerations for interleaved ADCs," *IEEE J. Solid-State Circuits*, vol. 48, no. 8, pp. 1806–1817, Aug. 2013.

[3] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, April 1999.

[4] C. Vogel, "The impact of combined channel mismatch effects in time-interleaved ADCs," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 1, pp. 415–427, Feb. 2005.

[5] J. Elbornsson, F. Gustafsson, and J. E. Eklund, "Blind adaptive equalization of mismatch errors in time-interleaved A/D converter system," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 1, pp. 151–158, Jan. 2004.

[6] V. Divi and G. Wornell, "Blind calibration of timing skew in time-interleaved analog-to-digital converters," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 3, pp. 509–522, Jun. 2009.

[7] T. Strohmer and J. Xu, "Fast algorithms for blind calibration in time interleaved analog-to-digital converters," in *IEEE Int. Conf. Acoust., Speech and Signal Process.*, Apr. 2007, pp. 1225–1228.

[8] C. Luo, L. Zhu, and J. McClellan, "Coordinated blind calibration for time interleaved ADCS," in *IEEE Int. Conf. Acoust., Speech and Signal Process.*, May 2013, pp. 3890–3894.

[9] S. Saleem and C. Vogel, "Adaptive blind background calibration of polynomial-represented frequency response mismatches in a two-channel time-interleaved ADC," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 6, pp. 1300–1310, Jun. 2011.

[10] B. Brannon, "Sampled systems and the effects of clock phase noise and jitter," Analog Devices, Inc., Norwood, MA, Tech. Rep. AN-756, 2004.

[11] D. S. Weller and V. K. Goyal, "Bayesian post-processing methods for jitter mitigation in sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2112–2123, May 2011.

[12] S. Kumar, V. K. Goyal, and S. E. Sarma, "Efficient parametric signal estimation from samples with location errors," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5285–5297, Nov. 2013.

[13] H. Araghi, M. A. Akhaee, and A. Amini, "Timing mismatch compensation in TI-ADCs using Bayesian approach," in *23rd European Signal Process. Conf. (EUSIPCO)*, Aug. 2015, pp. 1391–1395.

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing : Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.