TIMING MISMATCH COMPENSATION IN TI-ADCS USING BAYESIAN APPROACH

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

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

ABSTRACT

A TI-ADC is a circuitry to achieve high sampling rates by passing the signal and its shifted versions through a number of parallel ADCs with lower sampling rates. When the time shifts between the C channels of a TI-ADC are properly tuned, the aggregate of the obtained samples is equivalent to that of a single ADC with C-times the sampling rate. However, the performance of a TI-ADC can be seriously degraded under interchannel timing mismatch. As this non-ideality cannot be avoided in practice, we need to first estimate the mismatch value, and then, compensate it. In this paper, by adopting a stochastic bandlimited signal model we study the signal recovery problem from the samples of a TI-ADC affected by timing mismatch and jitter. In particular, we derive the Bayesian model and implement the minimum mean square error (MMSE) estimator. The latter is achieved by means of Gibbs sampling technique.

Index Terms— Bandlimited signals, Bayesian modeling, Gibbs sampling, Multichannel sampling, Timing mismatch.

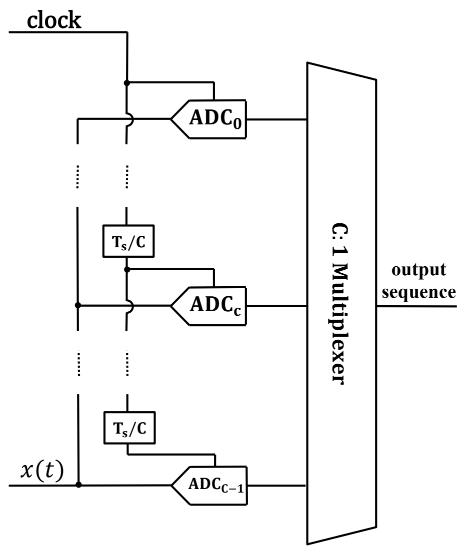

Fig. 1. Time-interleaved ADC with C channels

1. INTRODUCTION

The expansive developments in micro-electronic technologies have facilitated processing of the signals in discrete domain. However, prior to such processes, the analog signal needs to be converted to a discrete form. The electronic modules responsible for such a task are commonly referred to as analog to digital converters (ADC). The samples of a signal provide a unique representation only when the sampling rate is beyond a given threshold, e.g., Nyquist rate. In certain applications such as software-defined radios, the minimum sampling rate for unique representation could be potentially very high [1]. The realization of reliable ADCs with high sampling rates is quite challenging and usually involve considerable dissipation of power, which might not be feasible in many applications [2].

The time-interleaved ADC (TI-ADC) architecture, is a low-cost alternative for achieving high sampling rates. In this method, multiple sampling channels with equal interchannel delays are combined to form a high-rate ADC (Figure 1) [3, 4]. Besides the increase in the sampling rate, the TI-ADCs have the advantage of better power efficiency. For a given efficient number of bits (ENOB), the dissipated power as a function of the sampling rate is superlinear in high-speed ADCs. However, the power consumption scales linearly in a TI-ADC when the number of channels increases. The main drawback of TI-ADCs compared to high-speed ADCs is their increased sensitivity to sampling non-idealities such as jitter and interchannel timing mismatch. The errors introduced by the interchannel synchronizing circuitry can lead to drastic performance degradations in TI-ADCs [4–6].

The structure of TI-ADCs is a special case of filter-bank samplers introduced in the seminal work of Papoulis [7]. The filter-bank sampling theorem demonstrates that a bandlimited signal with the maximum frequency $f_c/2$ can be reconstructed using multichannel samples of certain C-channel filter-banks, in which each channel is sampled with the rate f_c/C. The TI-ADC corresponds to the simple case where the filters are time shifters. In this case, the ideal TI-ADC setup is when the c-th channel has a time shift of $\frac{c}{f_c}$ with respect to the first channel (the first channel corresponds to $c = 0$). Nevertheless, other shift values can still be acceptable. For a TI-ADC setup whose time shifts deviate from the ideal case, [8] investigates the performance of the least squares signal recovery method under noisy samples. The performance is found to be similar to the ideal case when the shift values are known. In practice, the exact value of the time shifts is unknown and should be estimated from the samples. An estimation method based on minimizing the rank of a matrix generated from the samples is proposed in [9]. As the minimization requires an exhaustive search over the set of all possible time shifts, the method becomes impractical for moderate to large number of channels. A Bayesian jitter mitigation technique is presented in [10] for a single channel ADC. The method tries to estimate the uniform samples of the bandlimited signal given the nonuniform ADC samples (due to unknown jitter) using the minimum mean-squared error criterion. For the method to work well, the average sampling rate should be strictly above the Nyquisr rate. Our goal in this paper is to extend this method to a multichannel TI-ADC setup, where it is crucial to compensate for the timing mismatches.

The rest of this paper is organized as follows. In Section 2, we introduce the Bayesian model for multichannel TI-ADC setup. The extended method for the estimation of the unknown variables and the reconstruction of the bandlimited signal is presented in Section 3. We investigate the performance of the proposed method in Section 4 by means of numerical simulations. Finally, Section 5 concludes the paper.

2. PROBLEM MODELING

We model the input $x(t)$ to be a random bandlimited signal with the maximum frequency of $f_c/2$. From the sampling theorem it follows that

$$x(t) = \sum_{k\in\mathbb{Z}} x[k] \operatorname{sinc}\big(\frac{t}{T_c} - k\big), \tag{1}$$

where $T_c = 1/f_c$, $x[k] = x(kT_c)$ are the discrete-time uniform samples of $x(t)$ at Nyquist rate, and $\operatorname{sinc}(t) = \sin(\pi t)/\pi t$. The samples $x[k]$ are assumed to be i.i.d. random variables following a zero-mean Gaussian distribution with variance σ_x^2 (the independent assumption amounts to a worst-case scenario). We consider a C-channel TI-ADC setup as depicted in Figure 1. Further, we assume that the sampling period in each of the channels is given by $T_s = T_c/M$, where M stands for the oversampling ratio. The multichannel samples are then, multiplexed to form a single sequence of signal samples. Without loss of generality, we consider $T_c = 1$ throughout the paper. As explained previously, the ideal time shift for the c-th channel (with respect to the first channel) is $\frac{c}{M.C}$ for $0 \le c \le C-1$; in practice, however, the shift values deviate from these nominal values. We denote the shift mismatch of the c-th channel by t_c.

To include the effect of sampling jitter, we model the timing error for each of the samples using additive i.i.d. random variables. In summary, for jointly modeling the jitter and interchannel timing errors, we include a timing mismatch $z_c[m]$ for the m-th sample of the c-th channel. We model $z_c[m]$s as i.i.d. Gaussian random variables with mean t_c and variance σ_z^2. The interchannel offsets t_c are also represented by i.i.d. zero-mean Gaussian random variables with variance σ_t^2.

According to the explained model, the set of sampling times for channel c is given by

$$\mathcal{T}_c = \left\{ \frac{m}{M} + \frac{c}{M.C} + z_c[m] \,\middle|\, m \in \mathbb{Z} \right\}, \tag{2}$$

where $0 \le c \le C-1$. As explained earlier, the time shifts are referenced relative to the first channel, thus we let $t_0 = 0$.

To include the noise effect on the samples caused by the thermal noise of the system and quantization of samples, we add i.i.d. zero-mean Gaussian random variables $w_c[m]$ with variance σ_w^2 to the samples. Hence, the received samples of c-th channel, $y_c[m]$, can be expressed as:

$$y_c[m] = \sum_{k\in\mathbb{Z}} x[k] \operatorname{sinc}\Big(\frac{mC+c}{MC} + z_c[m] - k\Big) + w_c[m], \tag{3}$$

where $m, k, c \in \mathbb{Z}$, and $0 \le c \le C-1$.

To reconstruct the $x[k]$ samples, we apply a rectangular window of length N_C to the channels output sequences and estimate a frame of length K samples of $x[k]$, where $N_C = K\,M$. Here, we ignore the windowing effects and the influence of the other frames. Therefore, (3) is approximated as:

$$y_c[m] \cong \sum_{k=0}^{K-1} x[k] \operatorname{sinc}\Big(\frac{mC+c}{MC} + z_c[m] - k\Big) + w_c[m], \tag{4}$$

where $0 \le m \le N_C - 1$. After receiving all the C frames of length N_C, we multiplex them into a single one of length $N = C\,N_C$. Let $y[n]$, $z[n]$, and $w[n]$ be unified sequences by multiplexing $\{y_c[m]\}_c$, $\{z_c[m]\}_c$, and $\{w_c[m]\}_c$, respectively. Thus, $y[n] = y_c[m]$ for $n = mC + c$, $0 \le c \le C-1$, and $0 \le m \le N_C - 1$ (similar relations hold for z and w). Now, we rewrite (4) without using the index c:

$$y[n] \cong \sum_{k=0}^{K-1} x[k] \operatorname{sinc}(n + z[n] - k) + w[n], \tag{5}$$

where $0 \le n \le N-1$. If we group $z[n]$, $w[n]$, and $x[n]$, respectively, into vectors $\mathbf{z}$, $\mathbf{w}$, and $\mathbf{x}$ we obtain that

$$\mathbf{z} \sim \mathcal{N}(\mathbf{t}, \sigma_z^2 \mathbf{I}_N), \tag{6}$$
$$\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma_w^2 \mathbf{I}_N), \tag{7}$$
$$\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \sigma_x^2 \mathbf{I}_K), \tag{8}$$

where $\mathbf{z} = [z[0], \ldots, z[N-1]]^T$, $\mathbf{w} = [w[0], \ldots, w[N-1]]^T$, $\mathbf{x} = [x[0], \ldots, x[K-1]]^T$, $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, $\mathbf{I}_N$ is the $N \times N$ identity matrix, and

$$\mathbf{t} = \mathbf{1}_{N_C \times 1} \otimes \mathbf{t}_C, \tag{9}$$

with $\mathbf{t}_C = [t_0, t_1, \ldots, t_{C-1}]^T$, $\mathbf{1}_{N_C \times 1}$ is an $N_C \times 1$ all-one vector, and $\otimes$ represents the Kronecker product. Moreover, we assume that $\mathbf{z}$, $\mathbf{w}$, and $\mathbf{x}$ are mutually independent.

For the sake of simplicity, we restate (5) in the matrix form as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{10}$$

where $\mathbf{y} = [y[0], \ldots, y[N-1]]^T$, and $\mathbf{H}[n, k] = \mathrm{sinc}(n + z[n] - k)$ for $0 \leq n \leq N-1$ and $0 \leq k \leq K-1$. We reserve the notation $\mathbf{H}[n, \cdot]$ for the n-th row of $\mathbf{H}$.

3. BANDLIMITED SIGNAL ESTIMATION

In this section, we aim to reconstruct the input signal based on the samples. Specifically, we shall estimate the input sequence/vector $\mathbf{x}$ by observing the samples obtained through TI-ADC channels, $\mathbf{y}$, in presence of both interchannel timing mismatch and sampling jitter modeled by a random vector $\mathbf{z}$. With respect to Bayesian modeling introduced in the previous section, MMSE criterion in (11) is applied for measuring the reconstruction error:

$$\mathrm{MSE}_{\mathbf{x}} = \mathbb{E}[\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2], \tag{11}$$

where $\hat{\mathbf{x}}$ denotes the estimate of $\mathbf{x}$.

In subection 3.1, we present the MMSE estimator of $\mathbf{x}$ when $\mathbf{z}$ is known. For the general case of unknown $\mathbf{z}$, we extend the Gibbs/slice method proposed in [10] to the multichannel problem in subection 3.2. In both subsections, the values of σ_x^2, σ_z^2, σ_t^2, and σ_w^2 are assumed to be known in advance.

3.1. MMSE estimator with known $\mathbf{z}$

First, we assume that the interchannel timing errors, $\mathbf{z}$, and thus, the $\mathbf{H}$ matrix are known. Then, the random vectors $\mathbf{x}$ and $\mathbf{y}$ are jointly Gaussian, which implies that the MMSE estimator of $\mathbf{x}$ given $\mathbf{y}$ coincides with the LMMSE estimator for the same problem [11]. The LMMSE estimator and its MSE are given below:

$$\hat{\mathbf{x}}_{\text{MMSE}|\mathbf{z}}(\mathbf{y}) = \hat{\mathbf{x}}_{\text{LMMSE}|\mathbf{z}}(\mathbf{y}) = \mathbf{C}_{\mathbf{xy}} \mathbf{C}_{\mathbf{yy}}^{-1} \mathbf{y}, \tag{12}$$

$$\mathrm{MSE}_{\mathbf{x}} = \mathrm{trace}[\mathbf{C}_{\mathbf{xx}} - \mathbf{C}_{\mathbf{xy}} \mathbf{C}_{\mathbf{yy}}^{-1} \mathbf{C}_{\mathbf{yx}}], \tag{13}$$

where $\mathbf{C}_{\mathbf{xy}}$ is the covariance matrix and $\mathbf{C}_{\mathbf{yy}}$ is the autocorrelation matrix,

$$\mathbf{C}_{\mathbf{xy}} = \mathbf{C}_{\mathbf{xx}} \mathbf{H}^T = \sigma_x^2 \mathbf{H}^T = \mathbf{C}_{\mathbf{yx}}^T, \tag{14}$$

$$\mathbf{C}_{\mathbf{yy}} = \mathbf{H} \mathbf{C}_{\mathbf{xx}} \mathbf{H}^T + \mathbf{C}_{\mathbf{ww}} = \sigma_x^2 \mathbf{H}\mathbf{H}^T + \sigma_w^2 \mathbf{I}. \tag{15}$$

Incorporating (14) and (15) into (12) and (13), we have that

$$\hat{\mathbf{x}}_{\text{MMSE}|\mathbf{z}}(\mathbf{y}) = \mathbf{H}^T (\mathbf{H}\mathbf{H}^T + \frac{\sigma_w^2}{\sigma_x^2} \mathbf{I})^{-1} \mathbf{y}, \tag{16}$$

$$\mathrm{MSE}_{\mathbf{x}} = \sigma_x^2 \ \mathrm{trace}[\mathbf{I} - \mathbf{H}^T (\mathbf{H}\mathbf{H}^T + \frac{\sigma_w^2}{\sigma_x^2} \mathbf{I})^{-1} \mathbf{H}]. \tag{17}$$

3.2. General method for MMSE estimation of $\mathbf{x}$

For the case of unknown $\mathbf{z}$, the MMSE estimator no longer simplifies to the LMMSE estimator. Therefore, we need to consider the posterior-mean estimator in its general form:

$$\begin{aligned} \hat{\mathbf{x}}(\mathbf{y}) &= \mathbb{E}\left[\mathbf{x} \mid \mathbf{y}\right] \\ &= \int \mathbf{x}\, p(\mathbf{x} \mid \mathbf{y})\, d\mathbf{x} \\ &= \frac{1}{p(\mathbf{y})} \int \mathbf{x}\, p(\mathbf{y} \mid \mathbf{x})\, p(\mathbf{x}) d\mathbf{x}. \end{aligned} \tag{18}$$

Due to the non-standard form of $p(\mathbf{y}|\mathbf{x})$, calculating the latter multivariate integral is not straightforward [12]. To overcome this difficulty, we extend the Gibbs/slice sampling method proposed in [10] to the multichannel framework.

Gibbs sampling is an instance of the more general Markov chain Monte Carlo (MCMC) techniques. In this iterative method, the expectation in (18) is approximated by the empirical mean using randomly generated samples of the posterior distribution $p(\mathbf{x}|\mathbf{y})$ [13]. In each iteration, we generate random samples following the conditional distribution of unknown variables (individually) given known variables. By including enough inner iterations, the distribution of these samples converges to the posterior probability density of $p(\mathbf{x}, \mathbf{z}, \mathbf{t}_C \mid \mathbf{y})$, which is the conditional distribution of all unknown random variables given the known variables. By assuming the validity of the generated samples, this procedure is repeated for a number of (outer) iterations in order to collect sufficient number of valid samples to obtain the empirical mean. In the multichannel problem, the unknown variables are $z[n]$, $x[k]$, and t_c, and the known variables are $y[n]$ where $0 \leq n \leq N-1$, $0 \leq k \leq K-1$, and $0 \leq c \leq C-1$.

In order to find the conditional probability distribution of $z[n]$ we have that

$$\begin{aligned} p(\, z[n] \mid \mathbf{z}[\cdot \backslash n], \mathbf{y}, \mathbf{x}, \mathbf{t}) &\propto p(\mathbf{z}, \mathbf{y}, \mathbf{x}, \mathbf{t}) \\ &\propto \mathcal{N}(\mathbf{y}; \mathbf{H}\mathbf{x}, \sigma_w^2 \mathbf{I})\, \mathcal{N}(\mathbf{t}, \sigma_z^2 \mathbf{I}) \\ &\propto \mathcal{N}(y[n]; \mathbf{H}[n, \cdot]\mathbf{x}, \sigma_w^2) \\ &\quad \times \mathcal{N}(z[n]; \mathbf{t}[n], \sigma_z^2), \end{aligned} \tag{19}$$

where $\mathbf{z}[\cdot \backslash n]$ is the $\mathbf{z}$ vector after omitting the n-th entry. The last equality is validated by the fact that the n-th row

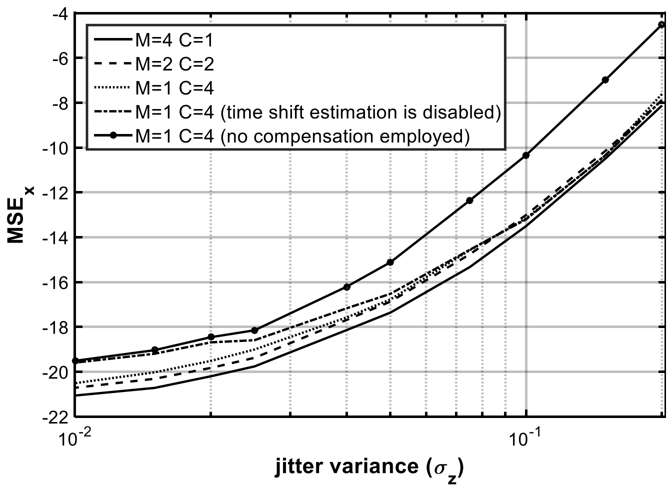

Fig. 2. Performance of the $MC = 4$ samplers for different values of σ_z^2, with $\sigma_w^2 = 0.05^2$ and $\sigma_t^2 = 0.025^2$. M denotes the ratio of the sampling rate to the Nyquist rate and C is the number of channels. We only have the jitter estimation in case that the time shift estimation is disabled.

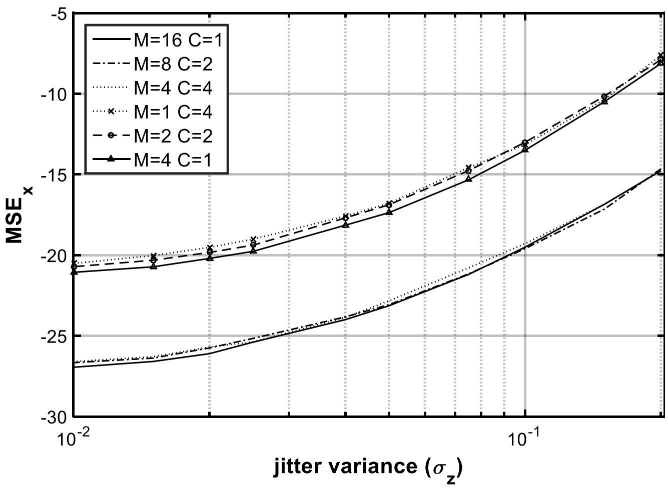

Fig. 3. Comparison of general method between the samplers with $MC = 4$ and $MC = 16$, for $\sigma_w^2 = 0.05^2$ and $\sigma_t^2 = 0.025^2$.

of $\mathbf{H}$ (i.e. $\mathbf{H}[n, \cdot]$) depends only on $z[n]$. Similar to [10], we utilize the slice sampling to produce $z[n]$ samples following the probability distribution of (19). The iterative slice sampling achieves the samples of a given distribution by repetitive generation of the samples from a specific uniform distributions [14].

For the estimation of $\mathbf{x}$ we have the following multivariate distribution:

$$\begin{aligned} p(\mathbf{x}|\,\mathbf{z}, \mathbf{t}, \mathbf{y}) &\propto p(\mathbf{y}|\,\mathbf{z}, \mathbf{t}, \mathbf{x})\, p(\mathbf{z}, \mathbf{t}, \mathbf{x}) \\ &\propto \mathcal{N}(\mathbf{y}; \mathbf{H}\mathbf{x}, \sigma_w^2 \mathbf{I})\, \mathcal{N}(\mathbf{x}; \mathbf{0}, \sigma_x^2 \mathbf{I}) \\ &\propto \mathcal{N}(\mathbf{x}; \frac{\sigma_x^2}{\sigma_w^2}\mathbf{H}^T\mathbf{y}, \sigma_w^2[\mathbf{H}^T\mathbf{H} + \frac{\sigma_w^2}{\sigma_x^2}\mathbf{I}]^{-1}). \quad (20) \end{aligned}$$

Finally, the conditional distribution of time shift, t_c, is obtained as follows:

$$\begin{aligned} p(t_c|\,\mathbf{z}, \mathbf{x}, \mathbf{y}, \mathbf{t}_C[\cdot\backslash c]) &\propto p(\mathbf{z}|\,\mathbf{t}_C)\, p(\mathbf{t}_C) \\ &\propto \mathcal{N}(\mathbf{z}_c; t_c \mathbf{1}_{N_C \times 1}, \sigma_z^2 \mathbf{I}_{N_C})\, \mathcal{N}(t_c; 0, \sigma_t^2) \\ &\propto \mathcal{N}(t_c; \frac{\sum_{i=0}^{N_C-1} z_c[i]}{N_C + \frac{\sigma_z^2}{\sigma_t^2}}, \frac{\sigma_z^2}{N_C + \frac{\sigma_z^2}{\sigma_t^2}}), (21) \end{aligned}$$

where $\mathbf{t}_C[\cdot\backslash c]$ is the $\mathbf{t}_C$ vector with its c-th entry removed, and $\mathbf{z}_c = [z_c[0], z_c[1], \ldots, z_c[N_C - 1]]^T$. Again, we observe that the posterior distribution of t_c is normal.

We, sequentially generate samples according to the probability density functions (19) to (21). After including a necessary number of inner iterations that guarantee the convergence in the distribution of the generated samples to the posterior distribution $p(\,\mathbf{x}, \mathbf{z}, \mathbf{t}_C|\,\mathbf{y})$, we repeat the whole procedure for a number of outer iterations to collect samples. Finally, by averaging the samples, we estimate the unknown variables, $\mathbf{x}$, $\mathbf{z}$, and $\mathbf{t}_C$.

4. SIMULATIONS

In this section, we examine the performance of the multichannel sampling system for various jitter variances σ_z^2 and interchannel time shift variances σ_t^2. For all scenarios, we assume $K = 10$ and $\sigma_x^2 = 1$. We represent the number of channels and the oversampling rate of each channel by C and M, respectively. We evaluate the performances according to the MSE of the estimated $\mathbf{x}$. The numbers of both inner and outer iterations of the extended Gibbs/slice method are set to 500. Figures 2 to 4 are the results of averaging 1000 independent realizations.

In Figure 2, the performance of the method (Section 3.2) for one, two and four channels are depicted for various values of σ_z^2. Here, we have $MC = 4$, $\sigma_w^2 = 0.05^2$, and $\sigma_t^2 = (\frac{0.1}{MC})^2 = 0.025^2$. For the case where neither the jitter nor the interchannel timing mismatch are compensated, (16) is used with $\mathbf{z} = \mathbf{0}$. When only the jitter is compensated, the general method is run by $\mathbf{t} = \mathbf{0}$ (t_cs are not updated within the iterations). It can be seen that the compensation of the interchannel timing mismatch causes the time-interleaved converter to exhibit a closer performance to the single channel sampling in the higher rate ($M = 4$ and $C = 1$). For better clarification, the performance of the method for various cases resulting in $MC = 4$ and $MC = 16$ are shown in Figure 3. According to Figures 2 and 3, we confirm the improvement in the TI-ADCs system equipped with jitter and time shift estimation.

Figure 4 illustrates the performance of the multichannel sampler for various values of σ_t^2, with $MC = 4$, $\sigma_w^2 = 0.05^2$, and $\sigma_z^2 = 0.05^2$. As it can be seen, when σ_t^2 exceeds a certain limit, the difference between MSE curves of the general method and the one in which we only compensate the jitter is increased. Therefore, for high values of σ_t^2, the effect of the interchannel timing mismatch can be sever and jitter-only-

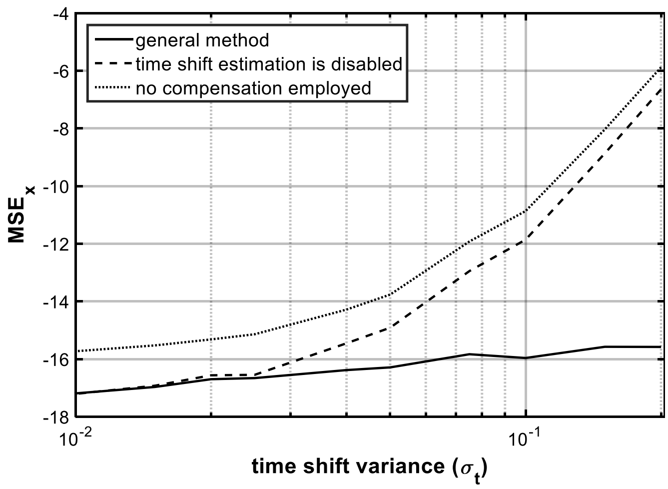

Fig. 4. Performance comparison of the general method for various σ_t^2 values against the cases which time shifts estimation is disabled, and neither jitter nor the interchannel timing mismatch compensation is applied. $M = 1, C = 4, \sigma_z^2 = 0.05^2$, and $\sigma_w^2 = 0.05^2$. The curve for general method is almost constant with σ_t.

compensation is not sufficient. An interesting observation is that compensating the interchannel timing mismatch is robust to σ_t^2. These results confirm the necessity of estimating time shifts in TI-ADCs.

5. CONCLUSION

In this paper, we studied the problem of compensating jitter and interchannel timing mismatch in TI-ADC setups. We first derived a Bayesian model for the problem by assuming the input to be a stochastic process. Then, to estimate the mismatch values, we extended the technique introduced in [10] to a multichannel scenario. In particular, we implemented the optimal MMSE estimator using the Gibbs sampling technique. The simulation results reveal a considerable gap between the performance of the converters with and without estimation-compensation of jitter and timing mismatch. Indeed, we observe that the performance of a TI-ADC under timing mismatch, where the mismatch is estimated and compensated, is very close the performance of a high-rate ADC.

Finally, we remark that the computational complexity for this iterative-based method is relatively high due to the need for generating sufficient valid samples, and therefore the method may encounter difficulties in real-time applications. However, we can achieve power-efficient signal acquisition systems by processing the gathered signals off-chip on the computing systems that power consumption is not matter of concern.

REFERENCES

[1] A. A. Abidi, "The path to the software-defined radio receiver," *IEEE J. Solid-State Circuits*, vol. 42, no. 5, pp. 954–966, May 2007.

[2] B. Murmann, "A/D converter trends: Power dissipation, scaling and digitally assisted architectures," in *IEEE Custom Integrated Circuits Conf.*, Sept. 2008, pp. 105–112.

[3] 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.

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

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

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

[7] A. Papoulis, "Generalized sampling expansion," *IEEE Trans. Circuits Syst.*, vol. 24, no. 11, pp. 652–654, Nov. 1977.

[8] V. Divi and G. Wornell, "Bandlimited signal reconstruction from noisy periodic nonuniform samples in time-interleaved ADCS," in *IEEE Int. Conf. Acoust., Speech and Signal Process.*, March 2008, pp. 3721–3724.

[9] P. Vandewalle et al., "Super-resolution from unregistered and totally aliased signals using subspace methods," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3687–3703, July 2007.

[10] 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.

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

[12] D. S. Weller, "Mitigating timing noise in adcs through digital post-processing," M.S. thesis, Electr. Eng. Comput. Sci. Dept., Massachusetts Inst. Technol., Cambridge, MA, 2008.

[13] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[14] Radford M. Neal, "Slice sampling," *Ann. Statist.*, vol. 31, no. 3, pp. 705–767, June 2003.