Detection of Communication Signals using Stochastic Quantization

Ryan M. Corey and Andrew C. Singer
University of Illinois at Urbana-Champaign

***Abstract*—We consider the problem of receiving a communication signal using a stochastic signal converter with uncertain quantization levels. Deeply scaled comparator circuits may offer advantages in size and power, but can exhibit random variations in their parameters. We can take advantage of these devices by incorporating the uncertainty of the devices into the system model used for inference. In this paper, we develop a statistical model for such a converter and show how it can be applied to estimate sequences of symbols. We demonstrate the performance of the converter in a simulated communication receiver.**

***Keywords*—Analog-to-digital conversion, quantization, stochastic systems, communication**

I. INTRODUCTION

We consider the problem of measuring a communication signal using a stochastic signal converter: an analog-to-digital converter built using unreliable components for which the output is not a deterministic function of the input. These converters may offer advantages in size, speed, and power at the cost of reliability. For example, deeply scaled CMOS comparators consume less area and energy than larger circuits, but are more sensitive to process variations in fabrication. We can compensate for the uncertainty introduced by the stochastic converter using statistical inference techniques. These converters are particularly well suited to communication problems because the received signal is already noisy; we can simply incorporate a statistical model of the converter into the existing channel model.

In this paper, we focus on the design of a flash analog-to-digital converter built from an array of comparators with unreliable reference levels. In deeply scaled CMOS, these comparators exhibit random offsets to their nominal reference levels due to mismatch between transistor threshold voltages [1]. In the circuits literature, many authors have proposed designs that rely on some form of calibration to measure the true levels after fabrication [2], [3], [4]. More recently, some authors have proposed using the uncalibrated comparator outputs to estimate the signal based on the statistics of the offsets [5], [6]. In [7], the authors demonstrated an architecture using redundant sets of deeply scaled comparators and an efficient statistical estimator. This approach directly leverages the device uncertainty to improve system performance. In our previous work [8], we gave theoretical results on the achievable performance of such an architecture for measuring a single sample, showing that mean square error is proportional to the standard deviation of the offsets and inversely proportional to the average density of reference levels.

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six STARnet centers sponsored by MARCO and DARPA. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant Number DGE-1144245.

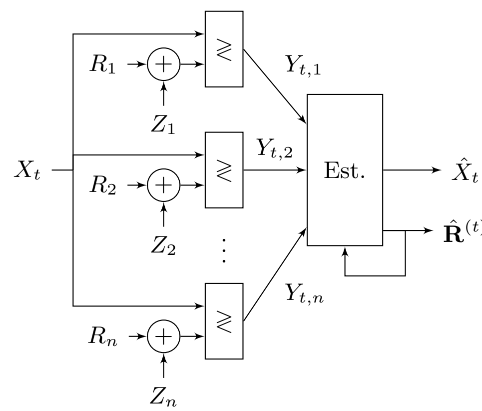

Figure 1. The stochastic signal converter measures the input signal by comparing it with a number of random threshold levels.

This paper generalizes our previous work to sequences of inputs. We model the converter uncertainty as a combination of time-varying noise and static offsets. Because the static offsets do not vary with time, they can be inferred from sequences of observations. The receiver thus performs joint estimation of the input signal and the reference levels. We first develop a statistical model of the stochastic signal converter. Next, we use that model to characterize the uncertainty introduced by the converter; we show that under certain conditions, it can be well approximated by additive noise. Finally, we demonstrate the performance of the stochastic converter for symbol detection using a simulated communication receiver.

II. STOCHASTIC SIGNAL CONVERTER

A. Converter Architecture

The converter architecture, shown in Figure 1, is similar to a flash analog-to-digital converter: each real-valued input sample X_t is measured by comparing it to each of n reference levels $R_1, \ldots, R_n$. The references are random variables that do not change over time. To account for time-varying uncertainty within the converter, we add a noise variable $Z_{t,i}$ to each reference for each sample. The outputs of the converter are binary variables $Y_{t,i} = 1$ if $X_t \geq R_i + Z_{t,i}$ and 0 otherwise for $i = 1, \ldots, n$.

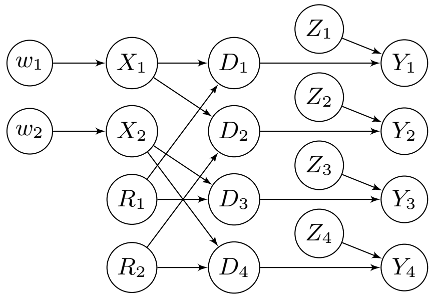

Figure 2. Relationships between variables defined in this section. $\mathbf{w}$ is the sequence of symbols, $\mathbf{X}$ is the noisy channel output, $\mathbf{R}$ is the set of random references, $\mathbf{D}$ is a vector of pairwise differences between X's and R's, $\mathbf{Z}$ is internal noise, and $\mathbf{Y}$ is the observed data.

Because the random references are fixed, they can be learned over time using the statistics of the input data and noise. Instead of a single sample X_t, consider a block of inputs $X_1, \ldots, X_m$. It will be useful to express the observations in terms of a linear mapping. Let $\mathbf{\Theta} = \left(\mathbf{X}^T, \mathbf{R}^T\right)^T$ be the length-$(m+n)$ random parameter vector containing both the samples and the references. Define the intermediate random variable $\mathbf{D} = J\mathbf{\Theta}$, a length-$mn$ vector containing the pairwise differences between the components of $\mathbf{X}$ and $\mathbf{R}$, where

$$J = \left[I_m \otimes \mathbf{1}_n \quad -\mathbf{1}_m \otimes I_n\right]. \tag{1}$$

Here $\otimes$ denotes the Kronecker product, I_a denotes the $a \times a$ identity matrix, and $\mathbf{1}_a$ is a length-a vector of 1's. Using this representation, the outputs can be written as a single vector $\mathbf{Y} = u\left(\mathbf{D} - \mathbf{Z}\right) = u\left(J\mathbf{\Theta} - \mathbf{Z}\right)$, where $u\left(\cdot\right)$ is the unit step function applied element-wise.

B. Statistical Model

To recover the input signal from the observations, we must develop a statistical model relating $\mathbf{\Theta}$ and $\mathbf{Y}$. Assume that the elements of $\mathbf{Z}$ are independent and identically distributed according to a continuous cumulative distribution function F_Z. To ensure that the parameters are identifiable, we require that F_Z is differentiable with density $f_Z\left(z\right) > 0$ for all $z \in \mathbb{R}$. For a given observation vector $\mathbf{y}$, the likelihood of a parameter vector $\boldsymbol{\theta}$ can be expressed in terms of $\mathbf{d} = J\boldsymbol{\theta}$ as

$$\begin{aligned}
p_{\mathbf{Y}|\mathbf{\Theta}}\left(\mathbf{y} \mid \boldsymbol{\theta}\right) &= p_{\mathbf{Y}|\mathbf{D}}\left(\mathbf{y} \mid \mathbf{d}\right) & (2)\\
&= \prod_{k=1}^{mn} \mathrm{P}\left(\{d_k \geq Z_k, y_k\} \cup \{d_k < Z_k, \bar{y}_k\}\right) & (3)\\
&= \prod_{k=1}^{mn} F_Z\left(d_k\right)^{y_k} \bar{F}_Z\left(d_k\right)^{\bar{y}_k}, & (4)
\end{aligned}$$

where $\bar{y}_k = 1 - y_k$ and $\bar{F}_Z = 1 - F_Z$. By the assumption that f_Z is strictly positive, F_Z and $\bar{F}_Z$ are also strictly positive and the log-likelihood can be expressed as

$$\ln p_{\mathbf{Y}|\mathbf{\Theta}}\left(\mathbf{y} \mid \boldsymbol{\theta}\right) = \mathbf{y}^T \ln F_Z\left(J\boldsymbol{\theta}\right) + \bar{\mathbf{y}}^T \ln \bar{F}_Z\left(J\boldsymbol{\theta}\right). \tag{5}$$

In this paper, we are concerned with the detection of communication signals. Thus, $\mathbf{X}$ is a noisy version of the channel output $\mathbf{w}$, which is drawn from a finite set $\mathcal{W}^m$ of sequences. Furthermore, $\mathbf{R}$ is a noisy version of the set of nominal reference levels. For communication, the goal is to recover $\mathbf{w}$ from $\mathbf{Y}$. If the statistics of the channel noise and offsets are known and if the offsets are independent of the channel noise, then the prior distribution for $\mathbf{\Theta}$ is $f_{\mathbf{\Theta}|\mathbf{W}}\left(\boldsymbol{\theta} \mid \mathbf{w}\right) = f_{\mathbf{X}}\left(\mathbf{x} \mid \mathbf{w}\right) f_{\mathbf{R}}\left(\mathbf{r}\right)$ and we can express the likelihood of the message as

$$p_{\mathbf{Y}|\mathbf{W}}\left(\mathbf{y} \mid \mathbf{w}\right) = \int p_{\mathbf{Y}|\mathbf{\Theta}}\left(\mathbf{y} \mid \boldsymbol{\theta}\right) f_{\mathbf{\Theta}|\mathbf{W}}\left(\boldsymbol{\theta} \mid \mathbf{w}\right) \mathrm{d}\boldsymbol{\theta}, \tag{6}$$

where the integral is over $\mathbb{R}^{m+n}$. The dependencies between the variables defined in this section are shown in Figure 2.

III. Estimation Performance

A. Fisher Information

The stochastic signal converter introduces additional uncertainty into the measurement process on top of the channel noise and conventional quantization error. We can characterize this uncertainty using the Fisher information [9] provided by the observations about the channel output.

Because the observations $\mathbf{Y}$ are conditionally independent given $\mathbf{D}$, the $mn \times mn$ Fisher information matrix $\mathcal{I}_{\mathbf{D}}\left(\mathbf{d}\right)$ is diagonal with entries

$$\begin{aligned}
\mathcal{I}_{D_k}\left(\mathbf{d}\right) &= \mathbb{E}\left[\left(\frac{\partial}{\partial d_k} \ln f_{\mathbf{Y}|\mathbf{D}}\left(\mathbf{Y} \mid \mathbf{d}\right)\right)^2\right] & (7)\\
&= \mathbb{E}\left[Y_k^2 \frac{f_Z\left(d_k\right)^2}{F_Z\left(d_k\right)^2} + \bar{Y}_k^2 \frac{f_Z\left(d_k\right)^2}{\bar{F}_Z\left(d_k\right)^2}\right] & (8)\\
&= \frac{f_Z\left(d_k\right)^2}{F_Z\left(d_k\right)\bar{F}_Z\left(d_k\right)}, & (9)
\end{aligned}$$

where $\mathbb{E}\left[\cdot\right]$ is the expectation. We denote this function by $\psi_Z\left(z\right) = f_Z\left(z\right)^2 / F_Z\left(z\right)\bar{F}_Z\left(z\right)$. Let $\psi_{t,i} = \psi_Z\left(x_t - r_i\right)$. The Fisher information matrix for $\mathbf{\Theta} = \boldsymbol{\theta}$ is

$$\begin{aligned}
\mathcal{I}_{\mathbf{\Theta}}\left(\boldsymbol{\theta}\right) &= J^T \mathcal{I}_{\mathbf{D}}\left(J\boldsymbol{\theta}\right) J & (10)\\
&= \begin{bmatrix}
\sum_i \psi_{1,i} & & 0 & -\psi_{1,1} & \cdots & -\psi_{1,n}\\
& \ddots & & \vdots & & \vdots\\
0 & & \sum_i \psi_{m,i} & -\psi_{m,1} & \cdots & -\psi_{m,n}\\
-\psi_{1,1} & \cdots & -\psi_{m,1} & \sum_t \psi_{t,1} & & 0\\
\vdots & & \vdots & & \ddots & \\
-\psi_{1,n} & \cdots & -\psi_{m,n} & 0 & & \sum_t \psi_{t,n}
\end{bmatrix}. & (11)
\end{aligned}$$

Since we have a prior distribution on $\mathbf{\Theta}$, the total Fisher information is

$$\begin{aligned}
\mathcal{I}_{\mathbf{\Theta}}^{\text{tot}} &= -\mathbb{E}\left[\nabla^2 \ln f_{\mathbf{\Theta},\mathbf{Y}}\left(\mathbf{\Theta}, \mathbf{Y}\right)\right] & (12)\\
&= -\mathbb{E}\left[\nabla^2 \ln f_{\mathbf{\Theta}}\left(\mathbf{\Theta}\right)\right] - \mathbb{E}\left[\nabla^2 \ln p_{\mathbf{Y}|\mathbf{\Theta}}\left(\mathbf{Y} \mid \mathbf{\Theta}\right)\right] & (13)\\
&= \mathcal{I}_{\mathbf{\Theta}}^{\text{prior}} + \mathbb{E}\left[J^T \mathcal{I}_{\mathbf{D}}\left(J\mathbf{\Theta}\right) J\right]. & (14)
\end{aligned}$$

The prior information is necessary to ensure that $\mathcal{I}_{\mathbf{\Theta}}^{\text{tot}}$ is invertible. The inverse Fisher information gives a lower bound on the error when $\mathbf{Y}$ is used to estimate $\mathbf{\Theta}$ [10]:

$$\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}}\left(\mathbf{Y}\right) - \mathbf{\Theta}\right)\left(\hat{\boldsymbol{\theta}}\left(\mathbf{Y}\right) - \mathbf{\Theta}\right)^T\right] \succeq \left(\mathcal{I}_{\mathbf{\Theta}}^{\text{tot}}\right)^{-1}. \tag{15}$$

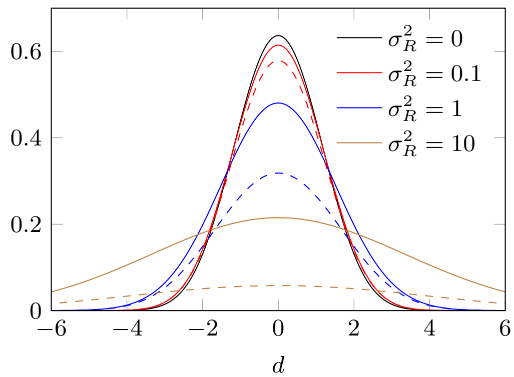

Figure 3. Fisher information provided by a single Y about the corresponding D. The solid curves show $\mathbb{E}\left[\psi_Z(D)\right]$ where $Z \sim \mathcal{N}(0,1)$ and $D \sim \mathcal{N}\left(d, \sigma_R^2\right)$. The dashed curves show $\psi_{\tilde{Z}}(d)$, where $\tilde{Z} \sim \mathcal{N}\left(0, 1+\sigma_R^2\right)$. The offsets are normally distributed with variance ranging from $\sigma_R^2 = 0.1$ (narrowest curves) to $\sigma_R^2 = 10$ (widest curves).

We can use this bound to quantify the uncertainty due to the stochastic converter. For now, suppose that we do not have a prior for $\mathbf{x}$. Suppose further that the comparator offsets are independent so that the prior information matrix is diagonal. Then the estimation error bound for a single sample x_t out of the length-m block is

$$\mathbb{E}\left[\left(\hat{x}_t(\mathbf{Y}) - x_t\right)^2\right] \geq \left(\sum_{i=1}^{n} \bar{\psi}_{t,i} - \frac{\bar{\psi}_{t,i}^2}{\mathcal{I}_{\mathbf{R}_i}^{\text{prior}} + \sum_{\tau=1}^{m} \bar{\psi}_{\tau,i}}\right)^{-1}, \tag{16}$$

where $\bar{\psi}_{t,i} = \mathbb{E}\left[\psi_Z\left(x_t - R_i\right)\right]$.

B. Performance and Block Length

The per-sample error bound (16) can be used to characterize the converter performance in terms of the block length m. First consider estimating a single sample x. In this case, we need not distinguish between the static reference offsets and time-varying internal noise. Let $\tilde{\mathbf{Z}}$ represent the combined offsets and noise and let $\mathbf{r}$ be the deterministic vector of reference levels. The Fisher information for x is then

$$\mathcal{I}_X(x) = \sum_{i=1}^{n} \psi_{\tilde{Z}}\left(x - r_i\right). \tag{17}$$

Next consider very large m: if the input statistics are known, then over time the estimator can learn $\mathbf{R}$ with arbitrary accuracy. Treating $\mathbf{R}$ as an observation instead of a parameter, the Fisher information for each x_t is

$$\mathcal{I}_X(x_t) = \sum_{i=1}^{n} \mathbb{E}\left[\psi_Z\left(x_t - R_i\right)\right]. \tag{18}$$

The two expressions (17) and (18) have very similar forms but differ in their treatment of the offset distribution. Figure 3 compares $\psi_{\tilde{Z}}(d)$ and $\mathbb{E}\left[\psi_Z(D)\right]$ for normally distributed offsets and noise. The observations provide more information when the references can be measured. In both cases, the information is maximized at $d = 0$, that is, when the input signal is close to the nominal reference. In the next section, we will show that under suitable conditions, (17) and (18) can both be approximated by simple ratios.

C. High Resolution Approximation

The Fisher information is most useful for high-resolution converters that have many closely spaced levels covering a wide dynamic range. In particular, consider a converter with nominal levels $\bar{\mathbf{r}}$ uniformly spaced distance Δr apart from $\bar{r}_1$ to $\bar{r}_n$, where Δr is small compared to the standard deviation of the noise and $\bar{r}_1$ and $\bar{r}_n$ are many standard deviations away from the point of interest. In this regime, we can make several simplifications: first, (17) and (18) are nearly constant as a function of x; second, the sums in those expressions can be approximated by improper integrals; and third, the uncertainty introduced by the converter resembles additive Gaussian noise. Furthermore, the Fisher information bound is most accurate in the high-resolution limit; see [8] for more details on the high-resolution approximation in the single-sample case.

If $\psi_Z(\cdot)$ has finite area and if f_Z is a density with scale parameter σ_Z of the form $f_Z(z) = \frac{1}{\sigma_Z} f_0\left(\frac{z}{\sigma_Z}\right)$ for some f_0, then the area under ψ_Z is given by

$$\int_{-\infty}^{\infty} \psi_Z(z)\,\mathrm{d}z = \int_{-\infty}^{\infty} \frac{\sigma_Z^{-2} f_0\left(z/\sigma_Z\right)^2}{F_0\left(z/\sigma_Z\right)\bar{F}_0\left(z/\sigma_Z\right)}\,\mathrm{d}z \tag{19}$$

$$= \frac{1}{\sigma_Z}\int_{-\infty}^{\infty} \frac{f_0(u)^2}{F_0(u)\bar{F}_0(u)}\,\mathrm{d}u \tag{20}$$

$$= \frac{c_f}{\sigma_Z}, \tag{21}$$

where c_f is a constant that depends on the distribution. For the normal distribution, $c_f \approx 1.806$ by numerical integration.[1]

For a high-resolution converter, the single-sample Fisher information (17) is

$$\mathcal{I}_X(x) = \frac{1}{\Delta r}\sum_i \psi_{\tilde{Z}}\left(x - r_i\right)\Delta r \tag{22}$$

$$= \frac{1}{\Delta r}\int_{-\infty}^{\infty} \psi_{\tilde{Z}}(x - r)\,\mathrm{d}r + \varepsilon \tag{23}$$

$$= \frac{c_f}{\sigma_{\tilde{Z}}\Delta r} + \varepsilon, \tag{24}$$

where ε is the approximation error due to finite resolution and dynamic range. For (18), if the offsets are identically distributed (denote them by $U_i = R_i - \bar{r}_i$, $i = 1, \ldots, n$), then by linearity of expectation we have

$$\mathcal{I}_X(x_t) = \frac{1}{\Delta r}\sum_i \mathbb{E}\left[\psi_Z\left(x_t - \bar{r}_i - U_i\right)\right]\Delta r \tag{25}$$

$$= \mathbb{E}\left[\frac{1}{\Delta r}\int_{-\infty}^{\infty} \psi_Z\left(x_t - \bar{r} - U\right)\mathrm{d}\bar{r} + \varepsilon\right] \tag{26}$$

$$= \frac{c_f}{\sigma_Z \Delta r} + \varepsilon. \tag{27}$$

[1] Incidentally, the integral has a closed form for the logistic distribution, with $c_f = \pi/\sqrt{3}$.

Note that (27) is identical to (24) except that σ_Z replaces $\sigma_{\tilde{Z}}$ in the denominator. For a high-resolution converter with random but known references, where time-varying noise is the dominant source of uncertainty, the distribution of the references is irrelavant to the converter's performance. Note that this would not be the case if quantization error dominated [11], [12].

The Fisher information is traditionally applied in regression problems; here, we are interested in detection. It can be shown by a central limit theorem argument [13] that in the high resolution limit, the distortion introduced by the converter can be well approximated by additive Gaussian noise with power $\mathcal{I}^{-1}$. We will use this additive-noise approximation to predict the converter performance for sequence detection.

IV. Sequence Detection

A. Single Symbol Classification

First consider the detection of a single symbol $w \in \mathcal{W}$. The maximum likelihood estimator is

$$w_{\mathrm{ML}} = \arg\max_{w\in\mathcal{W}} p_{\mathbf{Y}|W}\left(\mathbf{y} \mid w\right) \tag{28}$$

$$= \arg\max_{w\in\mathcal{W}} \int_{-\infty}^{\infty} f_{X|W}\left(x \mid w\right) p_{\mathbf{Y}|X}\left(\mathbf{y} \mid x\right) \mathrm{d}x. \tag{29}$$

Suppose that X is normally distributed with mean w and variance σ_X^2 and that the symbols in $\mathcal{W}$ are uniformly spaced distance Δw apart. If the estimator were allowed to observe X directly, the symbol error probability would be

$$\mathrm{P}_{\mathrm{e,ideal}} = 2\left(1 - \frac{1}{|\mathcal{W}|}\right) Q\left(\frac{\Delta w}{2\sigma_X}\right), \tag{30}$$

where $Q(\cdot)$ is the complementary standard normal cumulative distribution function. If we approximate the converter uncertainty by additive noise with power $\mathcal{I}^{-1}$, the "effective noise" of the overall system is $\sigma_X^2 + \mathcal{I}^{-1}$ and we expect an error probability of around

$$\mathrm{P}_{\mathrm{e,predicted}} = 2\left(1 - \frac{1}{|\mathcal{W}|}\right) Q\left(\frac{\Delta w}{2\sqrt{\sigma_X^2 + \mathcal{I}^{-1}}}\right). \tag{31}$$

To test this prediction, a communication channel was simulated with an alphabet of eight symbols spaced $\Delta w = 0.375$ apart. The symbols were drawn independently with uniform probability. The nominal references were evenly spaced $\Delta v = 0.05$ apart across the input range. The channel noise, offsets, and internal noise are normally distributed with variances σ_X^2, σ_R^2, and $\sigma_Z^2 = 0.05$ respectively. The experiment was repeated for ten thousand sets of references, each measuring 64 inputs. Each symbol was estimated independently using (29). Figure 4 shows the symbol error rate $\mathrm{P_e}$ as a function of the channel signal-to-noise ratio (SNR) and σ_R^2. Because the references are not learned over time, there is an error floor that depends on the uncertainty of the references. Figure 5 shows a waterfall plot of $\mathrm{P_e}$ against the effective SNR, $\mathcal{E}_s/\left(\sigma_X^2 + \mathcal{I}^{-1}\right)$, where $\mathcal{I}$ was computed using (17). The empirical error rates closely match the predicted error (31) for the single-sample case.

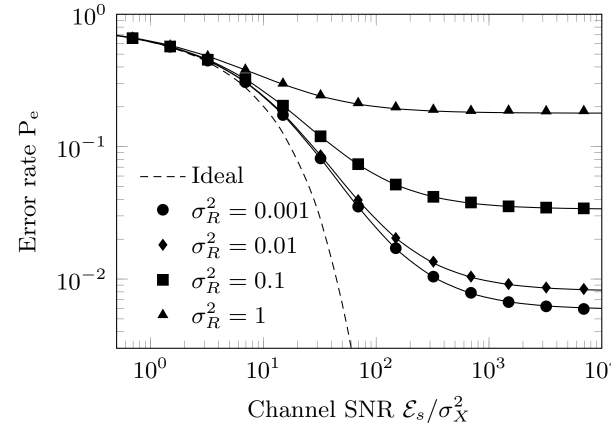

Figure 4. Simulated performance for detecting a single symbol, i.e., without learning the references. The solid curves show the predicted error rates based on (31) and (17) and the markers show the experimental error rates. The dashed curve shows the ideal error rate (30).

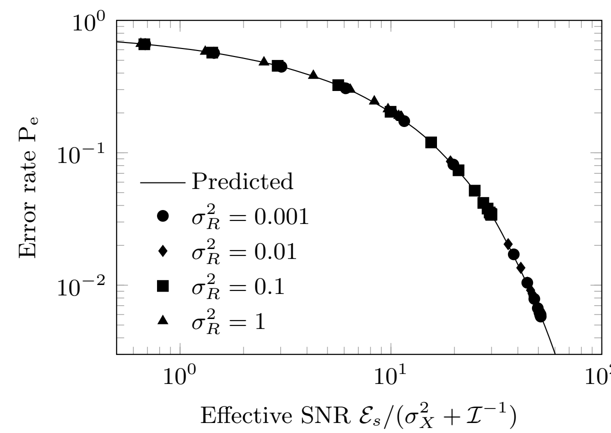

Figure 5. Simulated symbol error rate as a function of the effective SNR $\mathcal{E}_s/\left(\sigma_X^2 + \mathcal{I}^{-1}\right)$ for the single-sample estimator. The solid curve is the predicted error (31).

B. Joint Symbol Classification

The performance of the converter should improve for larger blocks of symbols. Unfortunately, it is more difficult to compute the maximum likelihood estimate for a sequence because the likelihood function (6) requires a high-dimensional integral. Future work will explore efficient algorithms for joint symbol detection. In this work, we instead separate the estimation and detection stages. First, we use Newton's algorithm to compute the joint maximum a priori estimate of the channel input and references:

$$\hat{\boldsymbol{\theta}}_{\mathrm{MAP}}\left(\mathbf{y}\right) = \arg\max_{\boldsymbol{\theta}} \ln f_{\boldsymbol{\Theta}}\left(\boldsymbol{\theta}\right) + \ln p_{\mathbf{Y}|\boldsymbol{\Theta}}\left(\mathbf{y} \mid \boldsymbol{\theta}\right). \tag{32}$$

Next, we choose the most likely sequence of symbols based on this estimate:

$$\hat{\mathbf{w}}\left(\mathbf{y}\right) = \arg\max_{\mathbf{w}} \ln f_{\mathbf{X}|\mathbf{W}}\left(\hat{\mathbf{x}}_{\mathrm{MAP}}\left(\mathbf{y}\right) \mid \mathbf{w}\right). \tag{33}$$

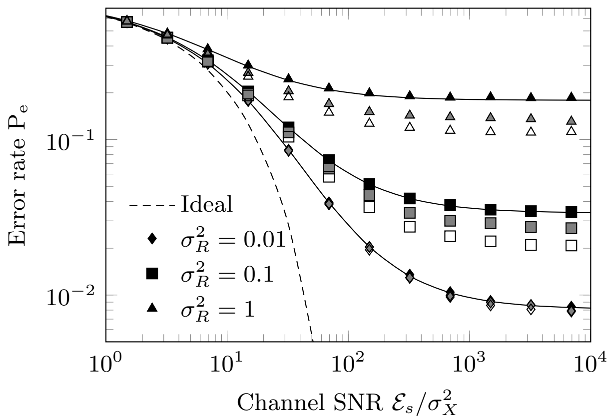

Figure 6. Simulated performance for detecting a sequence of symbols. The solid curves show the single-sample predicted error and the solid markers show the single-sample error rates. The gray and white markers show the error rates for $m = 16$ and 64 symbols per block, respectively.

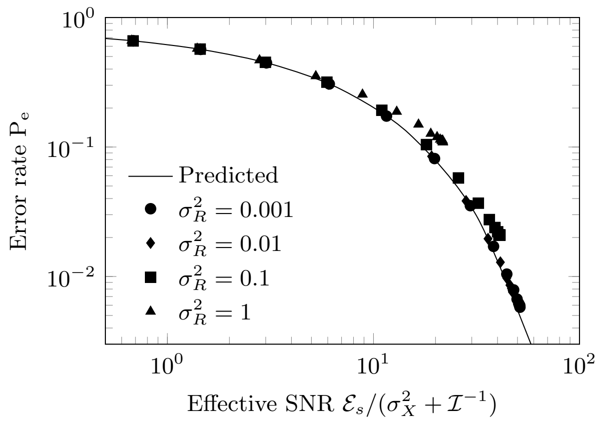

Figure 7. Simulated symbol error rate as a function of the effective SNR $\mathcal{E}_s/\left(\sigma_X^2 + \mathcal{I}^{-1}\right)$ for sequences of length $m = 64$.

The predicted error rate for this setup is also given by (31), but $\mathcal{I}^{-1}$ is computed using the per-sample bound (16).

The simulation described in the previous section was repeated for sequences of length $m = 16$ and 64 samples using the joint estimator (33). The symbol error rate for each data point was averaged over ten thousand random sequences and converter realizations. Figure 6 compares the simulated performance for each sequence length. Longer sequences improve the accuracy of the estimator significantly when the reference variance is large and only slightly when the reference variance is small. Figure 7 shows the experimental data in a waterfall plot of $\mathrm{P_e}$ against the effective SNR. The additive noise approximation is not as reliable for sequences as it is for single samples (Figure 5), but it is still useful for predicting system performance.

V. Conclusions

By conventional quantization metrics such as input-to-output linearity, the stochastic converter is essentially worthless: its input-to-output mapping is unknown and time-varying. Nevertheless, we have demonstrated that such a converter can perform well by application-level metrics, in this case symbol error rate. The stochastic converter requires more comparator circuits to achieve the same error rate as a conventional quantizer, but each comparator can be smaller and noisier. Designers can therefore trade off device performance for potential improvements in size, power, speed, or cost, allowing greater flexibility in system design.

To build systems with unreliable components, we must incorporate the device uncertainty into the overall statistical model of the system. The communication receiver is well suited to this statistical approach because the input signal is already stochastic. In fact, we have shown that we can predict the error rate of the system using a simple additive noise model. The uncertainty introduced by the converter depends on the variance of the reference levels as well as the spacing between them. We can improve the performance of the converter by learning the reference levels over time using past inputs. Future work will explore efficient approximate algorithms for jointly estimating sequences of inputs and the random reference levels.

References

[1] P. Kinget, "Device mismatch and tradeoffs in the design of analog circuits," *IEEE J. Solid-State Circuits*, vol. 40, no. 6, pp. 1212–1224, 2005.

[2] B. Razavi and B. Wooley, "Design techniques for high-speed, high-resolution comparators," *IEEE J. Solid-State Circuits*, vol. 27, pp. 1916–1926, Dec. 1992.

[3] C. Donovan and M. Flynn, "A 'Digital' 6-bit ADC in 0.25 μm CMOS," *IEEE J. Solid-State Circuits*, vol. 37, pp. 432–437, May 2002.

[4] D. Daly and A. Chandrakasan, "A 6-bit, 0.2 V to 0.9 V highly digital flash ADC with comparator redundancy," *IEEE J. Solid-State Circuits*, vol. 44, no. 11, pp. 3030–3038, 2009.

[5] T. Sundstrom and A. Alvandpour, "Utilizing process variations for reference generation in a flash ADC," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 56, no. 5, pp. 364–368, 2009.

[6] S. Weaver, B. Hershberg, P. Kurahashi, D. Knierim, and U.-k. Moon, "Stochastic Flash Analog-to-Digital Conversion," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 11, pp. 2825–2833, 2010.

[7] S. Tao, E. Abbe, and N. Verma, "Statistical adc enhanced by pipelining and sub-ranging," *IEEE Trans. Circuits and Syst. II: Exp. Briefs*, vol. PP, no. 99, pp. 1–1, 2015.

[8] R. M. Corey and A. C. Singer, "A multistage architecture for statistical inference with stochastic signal acquisition," *J. Signal Process. Syst.*, Jul. 2015.

[9] H. Poor, *An Introduction to Signal Detection and Estimation*. Springer, 1994.

[10] H. L. Van Trees, *Detection, estimation, and modulation theory*. John Wiley & Sons, 2004.

[11] R. Corey and A. Singer, "Scalar estimation from unreliable binary observations," in *IEEE Sensor Array and Multichannel Signal Process. Workshop (SAM)*, pp. 145–148, June 2014.

[12] D. Wang, Y. Polyanskiy, and G. Wornell, "Scalar quantization with noisy partitions and its application to flash adc design," in *IEEE Intl. Symp. Inf. Theory*, pp. 196–200, IEEE, 2014.

[13] R. M. Corey, *Statistical Inference with Unreliable Binary Observations*. PhD thesis, University of Illinois at Urbana-Champaign, Dec. 2014.