# Scalar Estimation From Unreliable Binary Observations

Ryan M. Corey and Andrew C. Singer
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

*Abstract*—We consider a scalar signal estimator based on unreliable observations. The proposed architecture forms an estimate using redundant arrays of unreliable binary sensors with detection thresholds that can vary randomly about their nominal values. We analyze the achievable performance of the estimator in terms of mean square error. We also provide approximate expressions for the error of a mean square optimal estimator in terms of the degree of redundancy in the system and the distribution of the random thresholds. We show that calibration and redundancy can compensate for uncertainty in the observations to form a reliable estimate.

## I. INTRODUCTION

We consider the problem of measuring a real-valued signal using a number of binary sensors. A typical solution to such a problem is to set each sensor's threshold to a different level. The signal must then lie between the largest threshold exceeded and the smallest threshold not exceeded. These two observations are sufficient to estimate the signal. Suppose, however, that the sensors are unreliable: the true threshold levels may deviate randomly from the designed levels due to process variations, noise, or low tolerances. We wish to form an estimate of the signal using all of the available observations and knowledge of the unreliable threshold statistics.

One important application of this problem is signal quantization. In many mixed-signal systems, such as communication receivers, high-power quantizers are a performance bottleneck. Novel device technologies may improve speed and energy consumption at the cost of circuit reliability. For example, highly scaled CMOS comparators exhibit random offsets in thresholds due to device mismatch. There have been a number of proposed methods, mostly in the circuits literature, to compensate for these random offsets. For example, the circuit can be calibrated using trim currents [1], comparator reassignment [2], or comparator redundancy [3]. In [4], random thresholds are measured during calibration. In [5], the signal is estimated based on the offset distribution.

Most previously proposed methods are designed to suppress the uncertainty in the circuit to improve the input-output relation of the quantizer. In this work, we treat quantization as an estimation problem. Using statistical methods and a
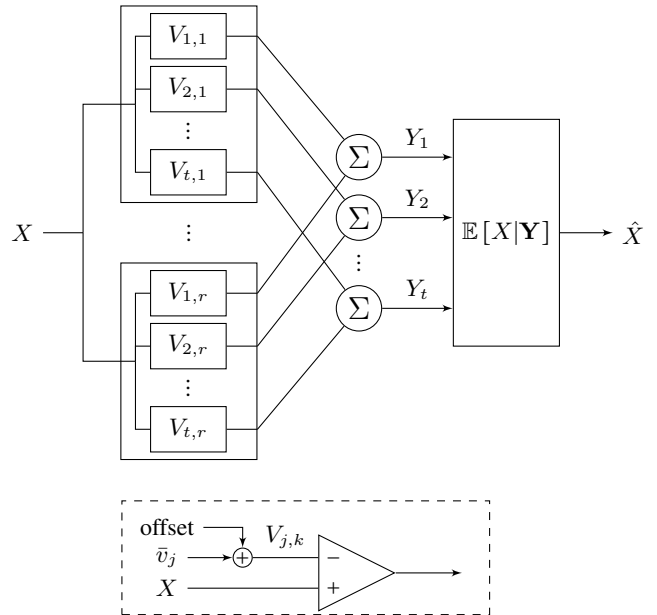
Fig. 1. The converter has $r$ binary sensors at each of $t$ nominal thresholds. The observations $Y_j$ are the number of exceedances at each level. The output is the MMSE estimate $\hat{X}$ of the input $X$. Each comparison has a random threshold $V_{j,k}$ offset from nominal threshold $\bar{v}_j$.

design that incorporates redundancy, the intrinsic uncertainty of the circuits can be exploited to improve performance. We present analytical and simulation results to show how mean square error performance scales with precision, redundancy, and threshold uncertainty.

## II. SYSTEM MODEL

The input to the system is a real-valued random signal $X$ with known continuous distribution $f_X(x)$ and variance $\sigma_X^2$. The detector consists of $r$ redundant arrays of $t$ binary sensors with nominal threshold levels $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_t$, for a total of $n = rt$ observations, as shown in Figure 1. Each observation has a true level $V_{j,k}$ that is offset from its nominal level $\bar{v}_j$ for $j = 1, \ldots, t$ and $k = 1, \ldots, r$. The level offsets are independent and identically distributed according to a known probability distribution $f_V(v)$. Each random threshold has probability density function $f_{V_j}(v) = f_V(v - \bar{v}_j)$ with variance $\sigma_V^2$. It is convenient to normalize the offset variance by the
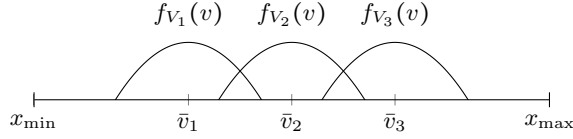
Fig. 2. The true thresholds are distributed around the desired nominal levels.

signal variance: let $\hat{\sigma}^2 \triangleq \sigma_V^2/\sigma_X^2$. The output of the detector is a length $t$ vector $\mathbf{Y}$ where $Y_j = \sum_{k=1}^{r} \mathbb{1}\{X \geq V_{j,k}\}$. These observations are used to form an estimate $\hat{X}$ of $X$.

In this paper we are concerned with the mean square error (MSE) of the estimate, defined as $\text{MSE} = \mathbb{E}_{X,\mathbf{V}}[(X - \hat{X})^2]$, where $\mathbb{E}_A[\cdot]$ denotes the expectation operation over the distribution of $A$. The estimator that minimizes the MSE is the minimum mean square error (MMSE) estimator $\hat{X} = \mathbb{E}_{X,\mathbf{V}}[X \mid \mathbf{Y}]$.

A useful formula for the MSE of the MMSE estimator is

$$\text{MSE} = \mathbb{E}_{\mathbf{Y}}\left[\text{Var}\left(X|\mathbf{Y}\right)\right] \tag{1}$$

If the thresholds are deterministic, then the quantizer output will determine which interval $[v_j, v_{j+1})$ contains the signal. If $X$ is uniform over an interval of length $\Delta x$, (1) can be written

$$\text{MSE} = \frac{1}{12\Delta x} \sum_j \left(v_{j+1} - v_j\right)^3 \tag{2}$$

For uniformly spaced levels, (2) reduces to the familiar formula for the MSE of an ideal quantizer:

$$\text{MSE} = \frac{\sigma_X^2}{(t+1)^2} \tag{3}$$

For nonuniform input distributions with large numbers of deterministic levels placed according to a density $\lambda(x)$, the MSE can be approximated by the distortion integral [6]:

$$\text{MSE} \approx \frac{1}{12}\frac{1}{t^2} \int \lambda(x)^{-2} f_X(x)\,dx \tag{4}$$

where the integral is over the support of $f_X$ and $\lambda$.

### III. PERFORMANCE WITH CALIBRATION

First, we consider the case where the random threshold levels $V_{j,k}$ are known exactly. This assumption is appropriate when noise is negligible and the offsets can be measured directly or accurately estimated.

To find an approximate expression for the MSE when the thresholds have significant variance, suppose that $n$ thresholds are placed uniformly at random across a range $\Delta x$ and ordered such that $V_{(i)} \leq V_{(i+1)}$ for $i = 1, \ldots, n$. The levels can be viewed as uniform order statistics scaled by $\Delta x$. It can be shown by integration that the third moment of an interval between uniform order statistics on $[0,1]^n$ is $\binom{n+3}{n}^{-1}$. Define $V_{(0)}$ and $V_{(n+1)}$ to be the lower and upper bounds of the interval so that $\left\{V_{(i+1)} - V_{(i)}\right\}_{i=0}^{n}$ are the intervals between
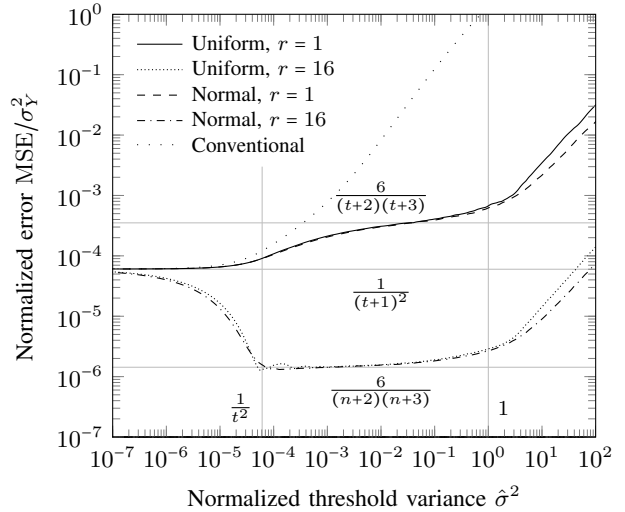


Fig. 3. Simulated MSE with perfect calibration for uniformly and normally distributed offsets with $t = 128$. The horizontal gray lines are the expressions from (3) and (8).

levels. Taking the expectation of (2) with respect to $\mathbf{V}$, the MSE is

$$\text{MSE} = \frac{1}{12\Delta x} \sum_{i=0}^{n} \mathbb{E}_{\mathbf{V}}\left[\left(V_{(i+1)} - V_{(i)}\right)^3\right] \tag{5}$$

$$= \frac{1}{12\Delta x} \sum_{i=0}^{n} (\Delta x)^3 \frac{3!n!}{(n+3)!} \tag{6}$$

$$= \frac{(\Delta x)^2}{2(n+2)(n+3)} \tag{7}$$

If the nominal levels and the input signal are uniformly distributed over the same range, then the MSE is

$$\text{MSE} = \frac{6}{(n+2)(n+3)}\sigma_X^2 \tag{8}$$

which is roughly six times the error of the deterministic quantizer from (3). Similarly, for a large number of levels with random offsets, (7) can be used to approximate the mean square error for signals near each point in an interval $\Delta x$ based on the density of nearby levels:

$$\Delta \text{MSE} \approx \frac{1}{2}\left(\frac{\Delta n}{\Delta x}\right)^{-2} \tag{9}$$

If the levels have nominal density $\lambda(x)$ and offset distribution $f_V(x)$, then the true levels will have density $(\lambda \star f_V)(x)$, where $\star$ denotes convolution. Substituting this level density in (9) and integrating with the input distribution gives

$$\text{MSE} \approx \frac{1}{2}\frac{1}{t^2} \int (\lambda \star f_V)(x)^{-2} f_X(x)\,dx \tag{10}$$

which is also six times the error of the offset-free quantizer from (4). Thus, there is a performance penalty of roughly 8 dB using sensors with random offsets.

This approximation is most useful when $\sigma_V^2$ is comparable to $\sigma_X^2/t^2$ so that the threshold distributions overlap significantly. When $\sigma_V^2$ is much smaller, the offsets are negligible

compared to the quantization error. When the offset power is comparable to the signal power, the range of the thresholds exceeds the range of the input signal and many sensors provide little useful information. Notice that the offset distribution affects the MSE in (10) only by filtering the level density. As long as the offsets are small relative to the signal range, performance does not depend strongly on $\sigma_V^2$ or on the specific shape of $f_V$.

These approximate expressions were verified by simulations for both uniform and normal offset distributions. The $t = 128$ nominal levels were spaced evenly across the range of a uniformly distributed input signal. At each nominal level, $r$ true thresholds were randomly selected. There were 600 randomly generated level sets for each distribution. For each of 512 random inputs, the MMSE estimate was computed from the comparator outputs. The MSE shown is the average of the squared error for each estimate. The variance of the offsets was varied from $\hat{\sigma}^2 = 10^{-7}$ to $10^2$. For comparison, a conventional quantizer was simulated that estimates the signal based on the largest nominal level of the activated sensors.

Figure 3 shows simulation results in which the MMSE estimator is given the true thresholds. For small offsets, the MSE is close to that of the ideal quantizer (3). When the offsets are comparable to the nominal spacing, the MSE is well approximated by (8). For large offsets, many thresholds fall outside the signal range and the error grows in proportion to the variance. The performance remains relatively constant for $1/t^2 < \hat{\sigma}^2 < 1$, particularly for the system that incorporates redundant sensors. The performance for uniformly and normally distributed offsets is similar; the statistics of the random thresholds appear not to have a significant impact on performance.

## IV. Performance without calibration

Next, we consider the error performance when the true thresholds are not known to the estimator. This is the case when the detector is not calibrated or when there is significant noise, and is similar to some problems in distributed estimation [7], [8]. The MSE performance depends on the size of the offsets relative to the spacing between nominal levels. We restrict our attention to the case where the offsets are comparable to the spacing: $1/t^2 < \hat{\sigma}^2 < 1$. Much smaller offsets would have little effect; much larger offsets would be impractical.

The best achievable MSE performance comes from the MMSE estimator. To find an an expression for the MMSE performance, we use a linear estimator to locally approximate the nonlinear MMSE estimator. Consider an input signal $X$ in the neighborhood of $x_0$. The probability that a sensor with nominal level $\bar{v}_j$ is activated by $X$ is $P\{X \geq V_{j,k}\} = F_{V_j}(X)$ where $F_{V_j}$ is the cumulative distribution function of the random threshold. Taking a linear expansion about $x_0$, $F_{V_j}(X) \approx F_{V_j}(x_0) + f_{V_j}(x_0)(X - x_0)$. The observation $Y_j$ has a binomial distribution with parameters $(r, F_{V_j}(X))$. The covariance of $X$ and $Y_j$ can be shown to be the same as that of $X$ and the conditional mean $rF_{V_j}(X)$:

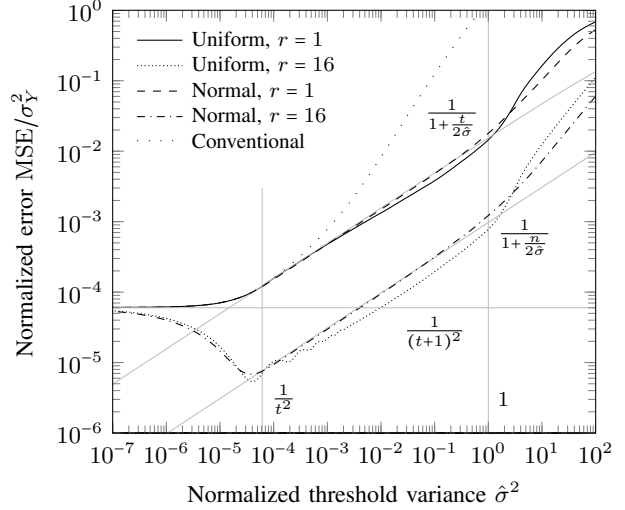$$\Sigma_{XY_j} = \mathrm{Cov}\left(X, rF_{V_j}(X)\right) \approx rf_{V_j}(x_0)\sigma_X^2 \tag{11}$$



Fig. 4. Simulated MSE with no calibration for uniformly and normally distributed offsets and $t = 128$. The gray lines show the theoretical error from (16).

By similar reasoning, the covariances of the observations are

$$\begin{aligned}\Sigma_{Y_iY_j} \approx{}& r^2 f_{V_i}(x_0) f_{V_j}(x_0)\sigma_X^2 \\ &+ \mathbb{E}_X[F_{V_i}(X)(1 - F_{V_j}(X))]\mathbb{1}_{\{i=j\}}\end{aligned} \tag{12}$$

Let $\mathbf{f} = f_{\mathbf{V}}(x_0)$ and $\mathbf{F} = F_{\mathbf{V}}(X)$. From the formula for the error of a linear MMSE estimator [9],

$$\begin{aligned}\mathrm{MSE} ={}& \Sigma_{XX} - \Sigma_{X\mathbf{Y}}\Sigma_{\mathbf{YY}}^{-1}\Sigma_{\mathbf{Y}X} \tag{13} \\ \approx{}& \sigma_X^2 - r^2\sigma_X^4\mathbf{f}^\mathsf{T}\left(r^2\sigma_X^2\mathbf{ff}^\mathsf{T}\right. \\ &+ \left. r\,\mathrm{diag}\left(\mathbb{E}[\mathbf{F}(1-\mathbf{F})^\mathsf{T}]\right)\right)^{-1}\mathbf{f} \tag{14}\end{aligned}$$

Applying the Woodbury identity [10], we find

$$\mathrm{MSE} \approx \frac{\sigma_X^2}{1 + r\sum_{j=1}^{t}\frac{f_{V_j}(x_0)^2\sigma_X^2}{\mathbb{E}_X\left[F_{V_j}(X)\left(1-F_{V_j}(X)\right)\right]}} \tag{15}$$

For distributions that are approximately uniform near the mean and have decreasing probability away from the mean, the ratio in the sum is proportional to $1/\hat{\sigma}^2$ and the number of levels for which $f_{V_j}(x_0)$ is non-negligible is proportional to $t\hat{\sigma}$. The MSE is therefore

$$\mathrm{MSE} \approx \sigma_Y^2\left(1 + \frac{rt}{\gamma\hat{\sigma}}\right)^{-1} \tag{16}$$

where $\gamma$ is a constant that depends on the distribution. Empirically, for uniform, normal, and similar distributions, $\gamma \approx 2$.

Figure 4 shows the simulated MSE performance of the MMSE estimator using the known distribution of the offsets. The simulations were conducted in the same manner as in Section III except that the true thresholds are unknown to the estimator. The simulated performance closely matches the theoretical error expression in (16) for offsets that are comparable to or larger than the spacing between levels but small relative to the signal range. The performance is similar for uniformly and normally distributed offsets.
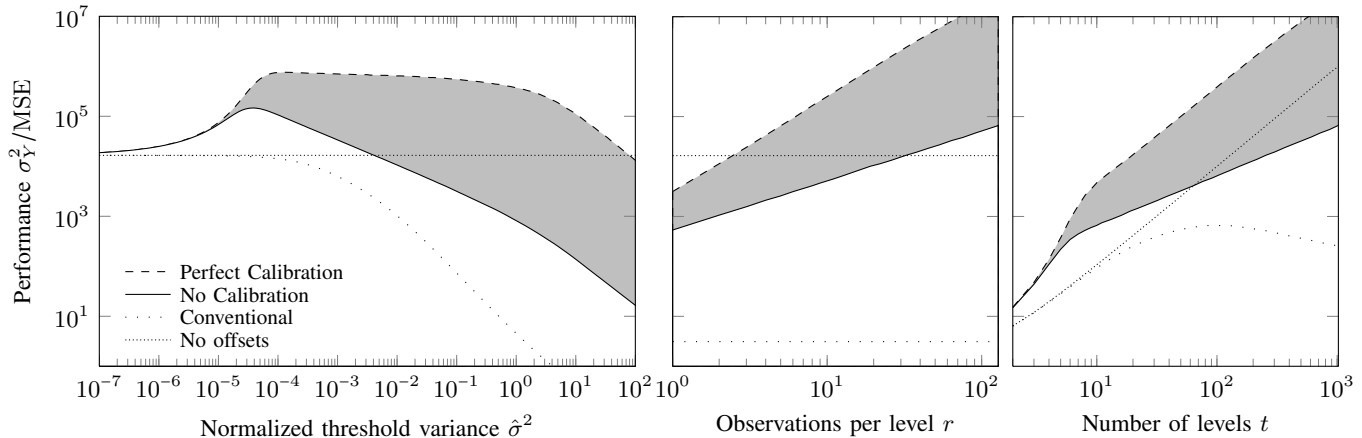
Fig. 5. Simulated performance as the system parameters are varied. The fixed parameters are $t = 128$, $r = 16$, and $\hat{\sigma}^2 = 1/64$. The offsets have a normal distribution. The shaded regions show achievable performance with imperfect calibration.

## V. CONCLUSIONS

The analysis of the previous sections shows that the performance of the estimator depends strongly on the degree of randomness in the observation thresholds. Suppose that the performance metric is the signal to noise ratio $\sigma_X^2/\mathrm{MSE}$. Figure 5 shows how the estimator's performance scales as a function of the three design parameters, $\sigma_V^2$, $t$, and $r$, for normally distributed offsets with and without calibration. The shaded region represents the achievable performance with partial calibration. Based on the analytical and simulation results of the previous sections, the performance appears to scale asymptotically as follows:

| Offsets | Conventional | Uncalibrated | Calibrated |
|---------|--------------|--------------|------------|
| $\hat{\sigma} \ll \frac{1}{t}$ | $t^2$ | $t^2$ | $t^2$ |
| $\frac{1}{t} < \hat{\sigma} < 1$ | $\frac{1}{\hat{\sigma}^2}$ | $\frac{n}{2\hat{\sigma}}$ | $\frac{n^2}{6}$ |
| $\hat{\sigma} \gg 1$ | $< 1$ | $\frac{n}{\frac{\pi}{2}\hat{\sigma}^2}$ | $\frac{n^2}{6\hat{\sigma}^2}$ |

Notice that for $\hat{\sigma} > 1/t$, these expressions depend on the total number of observations $n$; as long as the nominal levels are closely spaced, adding redundant arrays is equivalent to adding more levels to each array. For a perfectly calibrated system with offsets that are small compared to the signal power but large compared to the nominal spacing, the performance is within about 8 dB of that of an ideal detector and is not sensitive to the offset variance. Even without calibration, the performance scales at least as well as order $n/\sigma_V$.

In many applications, it may be more cost effective to use a large number of low-precision observations rather than a few high-precision observations. For example, consider designing a quantizer for an analog voltage signal using CMOS comparator circuits. Because transistor threshold variance is inversely proportional to circuit area [11], the total area of the quantizer is proportional to $n/\sigma_V^2$. Since the performance scales as $n/\sigma_V$ or better, it is asymptotically more cost-effective in terms of area to add redundancy than to use more reliable circuits. The statistical estimator allows the system designer to leverage

redundancy to improve performance without using more costly components.

We have shown that it is possible to reliably estimate a signal with unreliable observations by incorporating redundancy and statistical estimation. Without calibration, the achievable performance depends on the total number of observations and the uncertainty in thresholds; with calibration, it depends only on the number of observations. In both cases, the particular statistics of the random thresholds do not significantly impact performance. To build a reliable system, the designer must choose a sufficient number of observations and distribute the thresholds appropriately, but is free to choose unreliable sensors with arbitrary offset statistics. By embracing the stochastic behavior of next-generation devices, we can push performance and efficiency past the limits of deterministic architectures.

## REFERENCES

[1] S. Park, Y. Palaskas, and M. Flynn, "A 4-GS/s 4-bit flash ADC in 0.18-$\mu$m CMOS," *IEEE J. Solid-State Circuits*, vol. 42, no. 9, pp. 1865–1872, 2007.

[2] M. Flynn, C. Donovan, and L. Sattler, "Digital calibration incorporating redundancy of flash ADCs," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 50, no. 5, pp. 205–213, 2003.

[3] C. Donovan and M. Flynn, "A 'Digital' 6-bit ADC in 0.25 um CMOS," *Proc. Cust. Integr. Circuits Conf.*, vol. 37, no. 3, pp. 432–437, 2002.

[4] 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.

[5] 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.

[6] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Springer, 1992.

[7] A. Ribeiro and G. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-part i: Gaussian case," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131–1143, March 2006.

[8] T. Aysal and K. Barner, "Constrained decentralized estimation over noisy channels for sensor networks," *IEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1398–1410, April 2008.

[9] H. Stark and J. Woods, *Probability and Random Processes With Applications to Signal Processing*. Prentice Hall, 2002.

[10] J. E. Gentle, *Matrix Algebra: Theory, Computations, and Applications in Statistics*, ser. Springer Texts in Statistics. Springer, 2007.

[11] R. W. Keyes, "The effect of randomness in the distribution of impurity atoms on FET thresholds," *Appl. Phys.*, vol. 8, no. 3, pp. 251–259, 1975.