A Low Complexity Estimation Architecture Based on Noisy Comparators

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

Sen Tao and Naveen Verma
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544, USA

***Abstract*—We consider a low-complexity architecture for scalar estimation using unreliable observations. A signal is observed using a number of binary comparisons for which the threshold levels can vary randomly. We analyze the statistics of this system and find a Cramér-Rao lower bound on the squared error performance of the estimator. By incorporating redundant observations and applying statistical estimation techniques, we form an estimate with error that is much smaller than the uncertainty in the threshold levels. We propose a two-stage architecture that achieves near-optimal mean square estimation error with low complexity. The performance of the architecture is evaluated using a simulated prototype.**

Keywords*—*Distributed estimation, parameter estimation, quantization, sensor networks.

I. Introduction

We consider the problem of estimating a scalar parameter using a set of uncertain one-bit observations. Consider a system that consists of a number of binary sensors, each of which compares the same input signal to a different threshold level. If the threshold levels are known, then the estimation problem reduces to scalar quantization and the accuracy of the estimate depends on the spacing between levels. Suppose, however, that due to size, power, cost, or other constraints, the sensors are unreliable; in particular, suppose that the threshold levels vary randomly about their nominal values, as shown in Figure 1. This problem framework has a number of applications in distributed estimation, sensing, and quantization.

The problem of estimating a parameter based on binary observations with uncertain thresholds is similar to distributed estimation under bandwidth constraints. In that scenario, each sensor in a network makes an observation and quantizes it before transmitting it to a fusion center. This problem has recently been studied in the context of decentralized fusion [1] and wireless sensor networks [2], [3]. In these works, the quantizer is generally considered to be deterministic while the observations are corrupted by noise. In our construction of the problem, the quantization process itself is noisy.

One important application of this work is in low-power quantization. In many applications, such as communication receivers, high-resolution analog-to-digital converters have become a significant bottleneck [4]. The size and power of these quantizers can be reduced by using smaller comparator circuits. However, at small sizes and low voltages, comparators exhibit random offsets in their reference levels due to process variation [5]. There have been some proposals in the circuits literature to use these low-power, unreliable comparators for quantization by incorporating digital logic into the architecture. If the offsets do not vary with time, then the switching levels can be measured [6], calibrated with trim currents [7], or reassigned [8]. These calibration methods are designed to suppress the uncertainty in switching levels. However, using statistical methods, that randomness could be exploited to improve performance. In [9], a quantizer is designed with redundant comparators and the outputs are corrected with fault-tolerant logic. In [10], all comparators are designed with a single nominal reference level and exhibit large offsets; the outputs are summed to produce an estimate. These last architectures achieve reasonable performance with simple digital correction schemes that require no calibration; our work considers the possibility of using such uncertain comparators in a more general context of signal estimation, whether such comparators are part of a single noisy quantizer or are distributed across many low-power sensors. As such, our work builds on their approach and seeks fundamental limits on performance.

In this work, we analyze an estimation architecture with arbitrary numbers of nominal levels and redundant unreliable observations. Rather than relying on expensive calibration procedures to reduce the impact of threshold variations, we exploit the uncertainty in switching levels to produce an estimate with mean square error that is much smaller than the level variance. In this paper, we first apply estimation theory to find a lower bound on the achievable mean square error performance of an estimator using unreliable binary observations. Next, we show that a simple linear estimator can achieve performance close to this bound. We propose an efficient two-stage estimation architecture that consists of a low resolution detector followed by a linear estimator. Finally, we show simulated performance results for an implementation of the architecture.

II. System Model and Analysis

A. Definitions

The proposed architecture, shown in Figure 2, consists of r comparators at each of n nominal reference levels $v_1, \ldots, v_n$. Denote the true level of the j^{th} comparator with nominal level v_i by the random variable $V_{i,j}$. The offsets are assumed to be independent and identically distributed with a known cumulative distribution function (CDF) $F_V(x)$. Thus, each true level has CDF $\Pr\{V_{i,j} \le x\} = F_V(x - v_i)$. The input to be

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA.

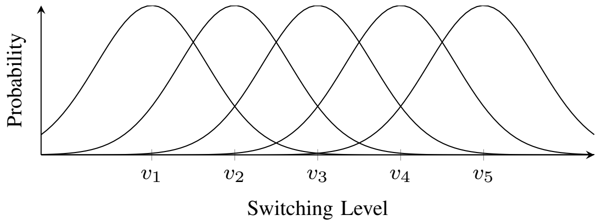

Fig. 1. If comparator levels are spaced closely together, the distributions of their offsets can overlap and cause errors.

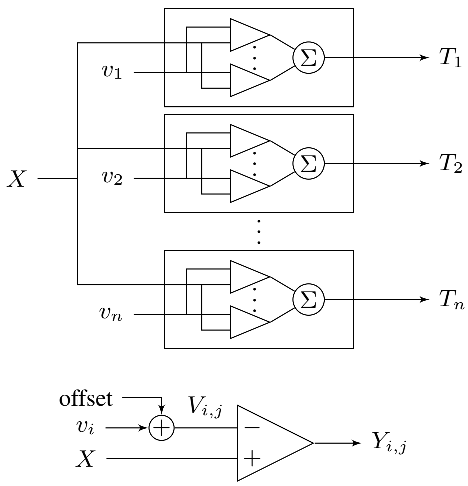

Fig. 2. A quantization system with unreliable comparators which exhibit random offsets in their switching levels.

estimated is a scalar signal $X \in \mathcal{X}$ where $\mathcal{X} = [x_{\min}, x_{\max}]$ is a finite interval of the real line. The observations are binary comparisons $Y_{i,j} = 1_{\{X \geq V_{i,j}\}}$. For a given $X = x$, $Y_{i,j}$ has a Bernoulli probability mass function (pmf):

$$p_{Y_{i,j}|X}(1 \mid x) = \Pr\{x \geq V_{i,j}\} \quad (1)$$
$$= F_V(x - v_i) \quad (2)$$

For brevity, let $F_i(x) = F_V(x - v_i)$ and let $\bar{F}_i(x) = 1 - F_V(x - v_i)$. When F_V is differentiable at x, denote the probability density function (pdf) by $f_i(x) = \frac{\partial}{\partial x} F_V(x - v_i)$.

In our analysis, we will consider two specific offset distributions. The logistic distribution, which induces a convenient pmf on the observations, is defined by the CDF

$$L(z) = \frac{\exp\{(z-\mu)/\beta\}}{1 + \exp\{(z-\mu)/\beta\}} \quad (3)$$

where μ is the mean and $\beta > 0$ is a scale parameter. The logistic distribution has variance $\sigma^2 = \beta^2\pi^2/3$. The normal distribution, which is often used to model offsets in comparator circuits [11] and other physical devices, has the CDF

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(u-\mu)^2}{2\sigma^2}\right\} \mathrm{d}u \quad (4)$$

where μ is the mean and σ^2 is the variance.

B. Sufficient Statistics

Because the offsets are independent and identically distributed, the rn observations can be reduced to the n sums $T_i = \sum_{j=1}^{r} Y_{i,j}$ for $i = 1, \ldots, n$. For a given x, each T_i has a binomial pmf

$$p_{T_i|X}(t_i \mid x) = \binom{r}{t_i} F_i(x)^{t_i} \bar{F}_i(x)^{r-t_i} \quad (5)$$

with mean $rF_i(x)$ and variance $rF_i(x)\bar{F}_i(x)$. On the subset of $\mathcal{X}$ where $F_i(x) \in (0,1)$ and F_i is differentiable, i.e., where $f_i(x) > 0$, (5) forms an exponential family of distributions [12]. The conditional pmf of T_i can be expressed in exponential form as

$$p_{T_i|X}(t_i \mid x) = h_i(t_i) \exp\{t_i \eta_i(x) - A_i(x)\} \quad (6)$$

where

$$\eta_i(x) = \ln \frac{F_i(x)}{\bar{F}_i(x)} \quad (7)$$

is the natural parameter for x, $A_i(x) = -r \ln \bar{F}_i(x)$, and $h_i(t_i) = \binom{r}{t_i}$. The conditional distribution of the full vector of observations is the product of the individual observation pmfs and has the exponential form

$$p_{\mathbf{T}|X}(\mathbf{t} \mid x) = h(\mathbf{t}) \exp\left\{\sum_{i=1}^{n} t_i \eta_i(x) - \sum_{i=1}^{n} A_i(x)\right\} \quad (8)$$

where $h(\mathbf{t}) = \prod_{i=1}^{n} h_i(t_i)$. If the sum is restricted to only those observations for which $f_i(x) > 0$, $i = 1, \ldots, k$, $k \leq n$, then (8) forms a curved exponential family. Because the space of natural parameters has a higher dimension than that of the true parameter, the sufficient statistic $\mathbf{T}$ may not be minimal; that is, there may exist a lower-dimensional statistic that is also sufficient [13].

If the offsets have a logistic distribution so that $F_i(x) = L(x - v_i)$, then

$$\eta_i(x) = (x - v_i)/\beta \quad (9)$$

and (8) can be written

$$p_{\mathbf{T}|X}(\mathbf{t} \mid x) = h(\mathbf{t}) \prod_{i=1}^{n} \left(e^{-\frac{t_i v_i}{\beta}}\right) e^{x \sum_{i=1}^{n} \frac{t_i}{\beta} - \sum_{i=1}^{n} A_i(x)} \quad (10)$$

which has the form of a one-dimensional exponential family. By the completeness theorem for exponential families [12], the sum

$$S(\mathbf{T}) = \sum_{i=1}^{N} T_i/\beta \quad (11)$$

is a complete sufficient statistic for the family of (10). Thus, it can be used to form a simple estimator that achieves equality in the Cramér-Rao lower bound, described in the next section.

C. Cramér-Rao Lower Bound

We can quantify the information an observation provides about X by its contribution to the Fisher information for X. Let $\mathbb{E}_x[A]$ and $\mathrm{Var}_x(A)$ denote the expectation and variance, respectively, of a random variable A given $X = x$, let $I(x)$ denote the Fisher information of $\mathbf{T}$ for $X = x$ and let $I_i(x)$ denote the Fisher information contributed by T_i for $X = x$.

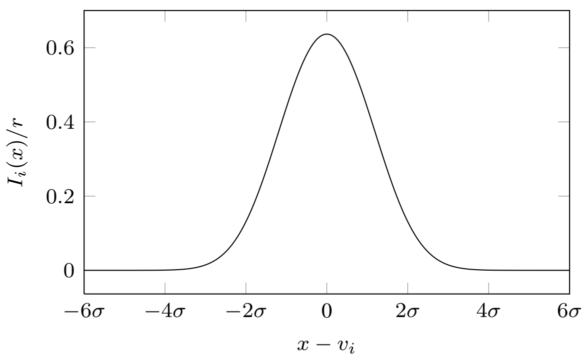

Fig. 3. Fisher information $I_i(x)$ provided by a single statistic T_i with normally distributed offsets. The information is maximized when $x = v_i$.

If $F_i(x)$ is differentiable at x and $f_i(x) > 0$, then the Fisher information is given by

$$\begin{aligned} I_i(x) &= \mathbb{E}_x\left[\left(\frac{\partial}{\partial x}\ln p_{T_i|X}(T_i \mid x)\right)^2\right] & (12)\\ &= (\eta_i'(x))^2\,\mathrm{Var}_x(T_i) & (13)\\ &= r\frac{f_i(x)^2}{F_i(x)\,\bar{F}_i(x)} & (14)\end{aligned}$$

where

$$\eta_i'(x) = \frac{f_i(x)}{F_i(x)\,\bar{F}_i(x)} \qquad (15)$$

is the derivative of the natural parameter. Figure 3 shows the Fisher information of a single statistic T_i as a function of x for normally distributed offsets with mean zero and variance σ^2. The comparators provide the most information about signals near their nominal level and little information about signals far from their level. Because the observations are independent, the overall Fisher information is the sum of the contributions from each observation that has support and a density at x:

$$I(x) = \sum_{i:f_i(x)>0} r\frac{f_i(x)^2}{F_i(x)\,\bar{F}_i(x)} \qquad (16)$$

To assess the Fisher information more concretely, suppose that the nominal levels are uniformly spaced distance Δv apart on $[x_{\min}, x_{\max}]$ and that all f_i have support at x. If x is far from $x_{\min}$ and $x_{\max}$ and n grows large so that $\Delta v \ll \sigma$, the sum in (16) can be approximated by an improper integral:

$$I(x) \approx \frac{r}{\Delta v}\int_{-\infty}^{\infty}\frac{f_V(x-v)^2}{F_V(x-v)\,(1-F_V(x-v))}\,\mathrm{d}v \qquad (17)$$

If the offsets have a logistic distribution with mean 0 and variance σ^2, then $I(x) \approx \frac{r}{\beta\Delta v} = \frac{\pi}{\sqrt{3}}\frac{r}{\sigma\Delta v} = 1.814\frac{r}{\sigma\Delta v}$. If the offsets are normally distributed with mean 0 and variance σ^2, then $I(x) \approx 1.806\frac{r}{\sigma\Delta v}$. Note that, except near the boundaries of $\mathcal{X}$, the Fisher information is not a function of x.

Using the Fisher information, we can bound the achievable performance of an estimator based on these observations. An estimator $\hat{X}(\mathbf{T})$ is called unbiased if $\mathbb{E}_x[\hat{X}(\mathbf{T})] = x$. The Cramér-Rao lower bound [12] (CRLB) on the variance of an unbiased estimator $\hat{X}$ is

$$\mathrm{Var}_x\left(\hat{X}(\mathbf{T})\right) \geq I(x)^{-1} \qquad (18)$$

Thus, the smallest achievable variance of an unbiased estimator with uniform levels and normally distributed offsets is

$$\mathrm{Var}_x\left(\hat{X}(\mathbf{T})\right) \geq 0.554\frac{\sigma\Delta v}{r} \qquad (19)$$

An unbiased estimator that achieves the CRLB is said to be efficient. In the next section, we derive a simple estimator for which the variance closely approaches the CRLB.

III. Localized Linear Estimator

A. Linear Estimation of the Input Signal

Because the distribution of $\mathbf{T}$ is parametrized by the generally nonlinear function $\eta(x)$, a minimum variance unbiased estimator may be computationally complex. We seek a low complexity suboptimal estimator $\hat{X}(\mathbf{T})$ for X that is approximately unbiased and has variance close to the CRLB. We first consider level offsets with the logistic distribution, which imposes a scalar complete sufficient statistic. If the nominal levels are uniformly spaced across $[x_{\min}, x_{\max}]$, x is far from the boundaries of the level range, and the levels are closely spaced so that the sum can be approximated by an integral, then $S(\mathbf{T})$ has mean

$$\begin{aligned} \mathbb{E}_x[S(\mathbf{T})] &= \sum_{i=1}^{n}\frac{r}{\beta}F_V(x-v_i) & (20)\\ &\approx \frac{r}{\beta\Delta v}\int_{x_{\min}}^{x_{\max}} L(x-v)\,\mathrm{d}v & (21)\\ &\approx \frac{r}{\beta\Delta v}(x-x_{\min}) & (22)\end{aligned}$$

An unbiased estimator is therefore

$$\begin{aligned} \hat{X}(\mathbf{T}) &= x_{\min} + \frac{\beta\Delta v}{r}S(\mathbf{T}) & (23)\\ &= x_{\min} + I(x)^{-1}S(\mathbf{T}) & (24)\end{aligned}$$

Because $\mathrm{Var}_x(S) = I(x)$, this estimator achieves the CRLB with equality. Furthermore, when the approximation in (22) holds, it can be shown from (10) that (24) is the maximum likelihood estimator for logistic offsets.

For other offset distributions, $\eta_i(x)$ is nonlinear and there is no scalar sufficient statistic. However, a linear estimator of the form $\hat{X}(\mathbf{T}) = \mathbf{a}^{\intercal}\mathbf{T} + b$ similar to (24) may perform well. Suppose that x is restricted to an interval centered around a known point x_0. Then we can locally approximate $\eta_i(x)$ by $\eta_i'(x_0)(x-v_i)$, which has a form similar to (9) and imposes a one-dimensional distribution in the form of (10) on $\mathbf{T}$. We also approximate $F_i(x)$ by $F_i(x_0) + f_i(x_0)(x-x_0)$. Let $S_{x_0}(\mathbf{T}) = \sum_{i=1}^{n}\eta_i'(x_0)\,T_i$. This statistic has mean

$$\begin{aligned} \mathbb{E}_x[S_{x_0}(\mathbf{T})] &= \sum_{i=1}^{n}\eta_i'(x_0)\,rF_i(x) & (25)\\ &\approx \sum_{i=1}^{n}\eta_i'(x_0)\,r\,(F_i(x_0) + f_i(x_0)(x-x_0)) & (26)\\ &= \sum_{i=1}^{n}\eta_i'(x_0)\,rF_i(x_0) \\ &\quad + \sum_{i=1}^{n}\frac{rf_i(x_0)^2(x-x_0)}{F_i(x_0)\,\bar{F}(x_0)} & (27)\\ &= \mathbb{E}_{x_0}[S_{x_0}(\mathbf{T})] + I(x_0)(x-x_0) & (28)\end{aligned}$$

Based on this relationship between S_{x_0} and x, we propose the following linear estimator for X:

$$\hat{X}(\mathbf{T}, x_0) = x_0 + \sum_{i=1}^{n} \frac{\eta_i'(x_0)}{I(x_0)} (T_i - rF_i(x_0)) \quad (29)$$

If $x_0 = x$, then $\hat{X}(\mathbf{T}, x_0)$ is unbiased and achieves the CRLB for unbiased estimators based on $\mathbf{T}$ with equality. Otherwise, the estimator has higher variance and may be biased. We characterize its performance in Section III-C.

B. Localized Estimator

The proposed estimator (29) is a function of the full vector $\mathbf{T}$. However, only a subset of the statistics contribute significantly to the estimate. For normally distributed offsets, as shown by Figure 3, more than 99% of the Fisher information is contributed by comparators with nominal levels within 3σ of x. By considering only a subset of the observations, we can compute a low-complexity suboptimal estimate that performs nearly as well as (29). We start with a low-resolution estimate X_0 of X, to be discussed in Section III-D. We retain the n^* statistics in $\mathbf{T}$ whose nominal levels are closest to X_0 and discard the rest. Denote this subset by $\mathbf{T}^*$. For the remainder of this section let T_k denote a statistic in $\mathbf{T}^*$ for $k = 1, \ldots, n^*$ and let v_k, F_k, $\bar{F}_k$, f_k, η_k, and I_k denote the corresponding variables. The Fisher information for this subset of statistics is $I^*(x) = \sum_{k=1}^{n^*} I_k(x)$. The localized linear estimator is

$$\hat{X}(\mathbf{T}^*, X_0) = X_0 + \sum_{k=1}^{n^*} \frac{\eta_k'(X_0)}{I^*(X_0)} (T_k - rF_k(X_0)) \quad (30)$$

This estimator would be especially simple to implement if X_0 were drawn from a finite subset of $\mathcal{X}$. In particular, suppose that X_0 is selected from the set of midpoints between nominal levels and that n^* is even. That is, $X_0 = \frac{1}{2}\left(v_{n^*/2} + v_{n^*/2+1}\right)$. Because $I^*(X_0)$ is approximately constant for X_0 far from the boundaries of the level range and $\eta_k'(X_0)$ is a symmetric function of $X_0 - v_k$, the coefficients of (30) are symmetric and do not depend on X_0. Thus, except near the boundaries of the level range, the linear estimator has a single set of symmetric coefficients for all X_0.

C. Performance Analysis

The performance of the localized estimator depends on the accuracy of X_0. We will now characterize the bias and efficiency of $\hat{X}$ in terms of X_0. Assume that f_k has support around X_0 and x for all $k = 1, \ldots, n^*$. If the offsets have a logistic distribution, the estimator will be unbiased regardless of the choice of X_0. More generally, if the offset distribution has an even pdf and the levels are symmetric about X_0, then it can be shown that the Taylor series about X_0 of the conditional mean includes only even derivatives of f_k:

$$\begin{aligned} &\mathbb{E}_x\left[\hat{X} \mid X_0\right] \\ &= X_0 + \sum_{k=1}^{n^*} \frac{\eta_k'(X_0)}{I^*(X_0)} (rF_k(x) - rF_k(X_0)) \quad (31) \\ &= x + r\sum_{k=1}^{n^*} \frac{\eta_k'(X_0)}{I^*(X_0)} \sum_{m=1}^{\infty} \frac{(x - X_0)^{2m+1}}{(2m+1)!} f_k^{(2m)}(X_0) \quad (32) \end{aligned}$$

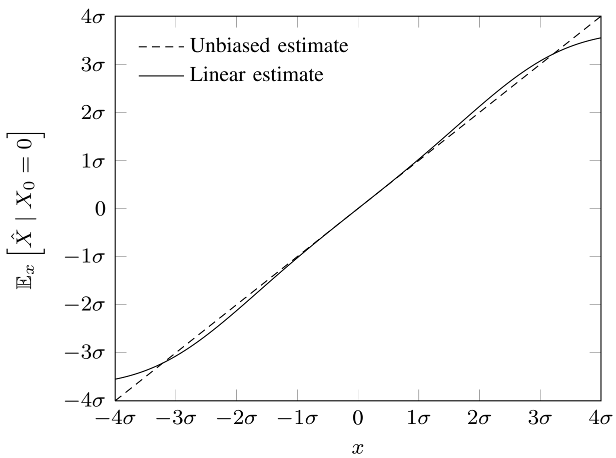

Fig. 4. Mean of the linear estimate (32) for $X_0 = 0$ with closely spaced uniform levels and normally distributed $(0, \sigma^2)$ level offsets. The dashed line represents an unbiased estimator.

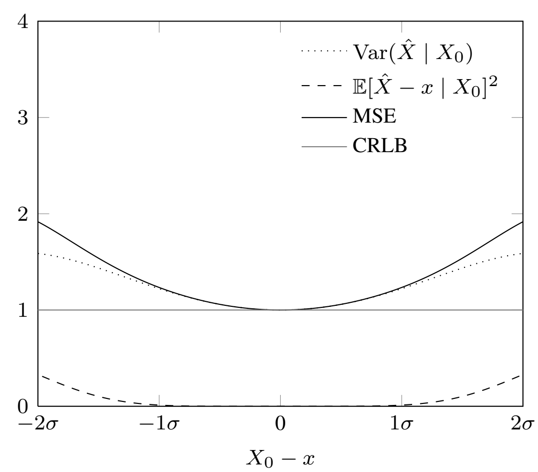

Fig. 5. Mean square error of the linear estimator (36) as a function of the initial estimate X_0 for normally distributed $(0, \sigma^2)$ offsets and $\Delta v = 0.1\sigma$, relative to the CRLB.

The sum includes only odd powers of $(x - X_0)$. Therefore, if X_0 is itself an unbiased estimate of X with $\mathbb{E}[(X_0 - X)^{2m+1}] = 0$ for all m, then $\mathbb{E}_x[\hat{X}(\mathbf{T}^*, X_0)] = x$ so $\hat{X}$ is also unbiased. If the offsets are normally distributed, then the integral approximation to the sum over k in (32) gives the conditional mean

$$\begin{aligned} \mathbb{E}_x\left[\hat{X} \mid X_0\right] = x &+ 3.696 \times 10^{-2} \frac{(x - X_0)^3}{\sigma^2} \\ &- 4.157 \times 10^{-4} \frac{(x - X_0)^5}{\sigma^4} + \cdots \quad (33) \end{aligned}$$

The mean for uniformly spaced levels with normally distributed offsets is shown in Figure 4.

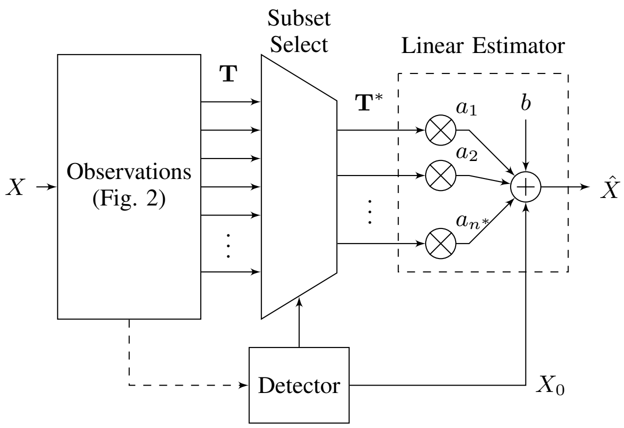

Fig. 6. The two stage estimation architecture includes a low-resolution detector and a linear estimator.

The conditional variance of the estimate is

$$\mathrm{Var}_x\left(\hat{X} \mid X_0\right) = \sum_{k=1}^{n^*} \frac{\left(\eta_k'\left(X_0\right)\right)^2}{I^*\left(X_0\right)^2} r F_k\left(x\right) \bar{F}_k\left(x\right) \quad (34)$$

$$= \sum_{k=1}^{n^*} \frac{I_k\left(X_0\right)}{I^*\left(X_0\right)^2} \frac{F_k\left(x\right) \bar{F}_k\left(x\right)}{F_k\left(X_0\right) \bar{F}_k\left(X_0\right)} \quad (35)$$

The variance approaches the CRLB as $X_0 \to x$.

Combining the variance and bias, the overall mean square error (MSE) of $\hat{X}$ is

$$\mathrm{MSE}_x\left(\hat{X} \mid X_0\right) = \mathrm{Var}_x\left(\hat{X} \mid X_0\right) + \left(\mathbb{E}_x\left[\hat{X} - x \mid X_0\right]\right)^2 \quad (36)$$

The conditional variance, squared bias, and overall MSE for normally distributed offsets are shown in Figure 5. For the estimator shown in the figure, if X_0 is accurate within σ, the variance of $\hat{X}$ will be within 1 dB of the CRLB.

D. Low Resolution Detection

To ensure that the performance of $\hat{X}$ is close to the CRLB, the error of X_0 must be small compared to the offset deviation but can be much larger than the desired error of $\hat{X}$. This motivates the two-stage estimation architecture shown in Figure 6: a low complexity, low resolution detector chooses the initial estimate X_0, which is used to select the subset of observations $\mathbf{T}^*$ that is used in the linear estimator. We wish to find a detector with the lowest possible complexity that achieves acceptable performance. Because the detector need not have high resolution, there are many possible solutions. The complexity of the detector will depend upon its particular implementation in hardware; here, we consider one example of a detector that is conceptually and analytically simple.

Let $Q\left(\mathbf{Y}\right) = \sum_{i=1}^{n} Y_{i,1}$ be the sum of one comparator output from each level. The decision rule simply selects the subset of $\mathbf{T}$ indexed by Q. The corresponding low resolution estimate resembles that from (24):

$$X_0\left(Q\left(\mathbf{Y}\right)\right) = \frac{1}{2}\left(v_Q + v_{Q+1}\right) \quad (37)$$

$$= x_{\min} + \Delta v Q\left(\mathbf{Y}\right) \quad (38)$$

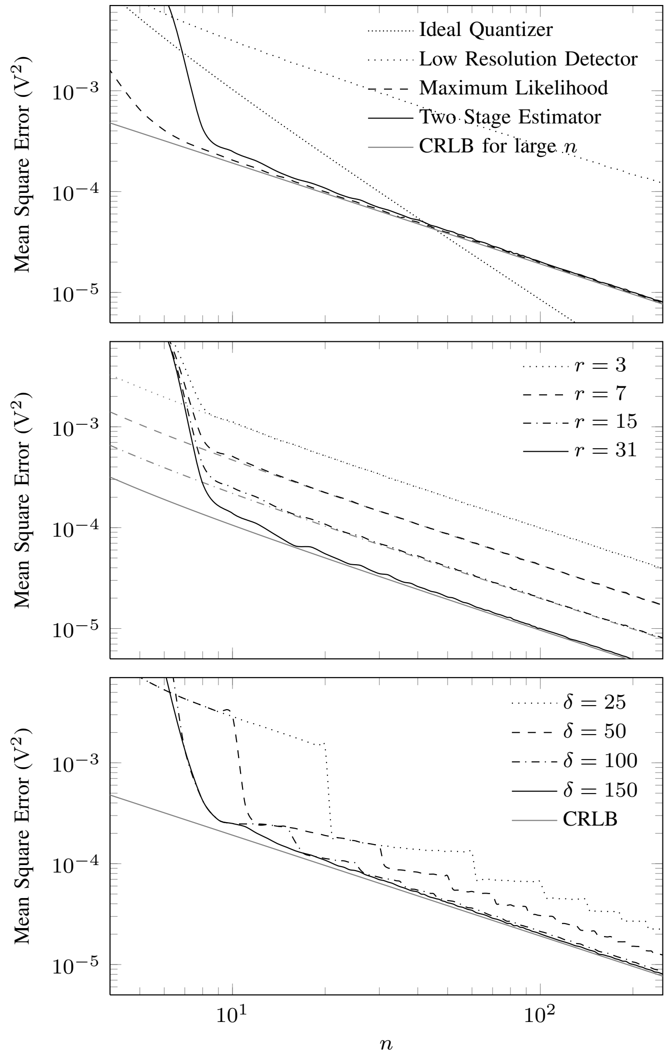

Fig. 7. Simulated performance of a prototype architecture with $\mathcal{X} = [0\,\mathrm{V}, 1\,\mathrm{V}]$ and $\sigma = 50\,\mathrm{mV}$. Top: $r = 15$ and $\delta = 150\,\mathrm{mV}$. Middle: Two stage estimator with $\delta = 150\,\mathrm{mV}$. Bottom: Two stage estimator with $r = 15$. Gray curves show the CRLB for large n from (19).

For x far from the boundaries of the level range, the variance of the low resolution estimate is approximately

$$\mathrm{Var}_x\left(X_0\right) = \Delta v^2 \sum_{i=1}^{n} F_i\left(x\right) \bar{F}_i\left(x\right) \quad (39)$$

$$\approx \Delta v \int_{-\infty}^{\infty} F_V\left(x - v\right)\left(1 - F_V\left(x - v\right)\right) \mathrm{d}v \quad (40)$$

For normally distributed offsets, the variance is $0.564\sigma\Delta v$. If the nominal levels are placed with spacing $\Delta v \ll \sigma$, then the variance of X_0 is small compared to σ^2 and the local estimator performance is close to the CRLB.

IV. Scalar Estimation Performance

A. Prototype Architecture

We will demonstrate the performance of the proposed estimation architecture using a simulated example. We wish to measure a voltage signal between 0 V and 1 V using a number of comparator circuits. The n nominal reference voltages are spaced $\Delta v = (1\,\mathrm{V})/(n-1)$ apart from 0 V to 1 V. The comparators are subject to process variations that cause normally distributed offsets with mean 0 and deviation $\sigma = 50\,\mathrm{mV}$. The low resolution detector is that described in Section III-D. The localized estimator acts on outputs from nominal levels within a range parameter $\pm\delta$ of the initial estimate. If the levels are spaced more than 2δ apart, then the two levels closest to X_0 are included. If the initial estimate is less than δ or greater than $1\,\mathrm{V} - \delta$, fewer levels are included and the weights are normalized accordingly.

A uniform quantizer with nonrandom switching levels would achieve mean square error $1\,\mathrm{V}^2/12(n-1)^2$. From (16), the CRLBs of the estimator for large n at the center and edges of the signal range are $I\left(0.5\,\mathrm{V}\right)^{-1} = 0.027\,\mathrm{V}^2/(n-1)r$ and $I\left(0\,\mathrm{V}\right)^{-1} = I\left(1\,\mathrm{V}\right)^{-1} = 0.055\,\mathrm{V}^2/(n-1)r$. The mean CRLB is $0.0295\,\mathrm{V}^2/(n-1)r$.

B. Simulations

The estimator performance was assessed using Monte Carlo simulations. The number of nominal levels was varied from $n = 4$ to 250, the number of comparators per level was varied from $r = 3$ to 31, and the range parameter was varied from $\delta = 25\,\mathrm{mV}$ to $150\,\mathrm{mV}$. For each configuration, 2000 input signals were drawn randomly from a uniform distribution on $[0\,\mathrm{V}, 1\,\mathrm{V}]$ and 2000 sets of switching levels were drawn from the normal distribution with $\sigma = 50\,\mathrm{mV}$. For comparison, each signal was also estimated with a maximum likelihood (ML) estimator $\hat{X}_{\mathrm{ML}}(\mathbf{t}) = \arg\max_x p_{\mathbf{T}|X}(\mathbf{t} \mid x)$ using the full vector of observations. The results, shown in Figure 7, represent the average MSE performance over all estimators.

C. Performance and Design

The simulation results confirm that the two-stage estimator performs nearly as well as the ML estimator and close to the CRLB as long as the spacing between nominal levels is small compared to the offset deviation. For $r = 15$ and $n = 21$, the levels are spaced σ apart and the average MSE is about 3% above the CRLB. At $n = 100$, it is about 1% higher. If the levels are closely spaced compared to σ, then the average MSE is inversely proportional to both n and r. The system designer can therefore improve performance by adding more redundant observations at each level without changing the low resolution detector or the size of the local estimator.

The size of the local estimator is a tradeoff between complexity and performance. If the range is too narrow, more observations will be required to achieve the desired level of performance and the estimator will be more sensitive to error in X_0. On the other hand, there is little benefit to including observations for which the levels have negligible probability density near X_0. Figure 7 shows that performance is quite poor when $\delta \leq \sigma$ but good for $\delta > \sigma$.

This architecture is best suited to systems where the offset deviation is small compared to the input range but large compared to the desired error. Then the linear estimator performs nearly as well as a more complex ML estimator. The choice of design parameters depends on the estimate of the offset deviation σ. If the assumed value of σ is too high, the levels may be spaced too far apart. If it is too low, the coefficients will be too large. Thus, it is best to use a conservative estimate of σ to choose the level spacing and a generous estimate to set the coefficients. If the true value of σ is smaller than the assumed value, the estimator variance will be worse than the achievable performance but better than the designed performance.

V. Conclusions

The analytical results in Section III and the simulation results in Section IV suggest that the proposed architecture achieves performance close to that of an optimal estimator in terms of mean square error. The signal to noise ratio of the estimate is directly proportional to total number of observations and is inversely proportional to the spacing between levels and to the square root of the offset variance. The estimator can achieve accuracy much finer than the spacing between levels or the deviation in level offsets by leveraging independent redundant observations. Thus, as long as the offsets remain independent, there is no fundamental limit on performance.

The two-stage structure has a simple implementation requiring a sum unit, a subset selector, and a weighted sum. It requires no calibration or measurement. To design the estimator, the system designer does not need to know the precise statistics of the input signal, only a rough estimate of the offset variance. By leveraging the uncertainty in the observations, the estimation architecture achieves strong performance and robustness with low complexity.

References

[1] H. Papadopoulos, G. W. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Trans. Inf. Theory*, vol. 47, pp. 978–1002, Mar. 2001.

[2] A. Ribeiro and G. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks-Part I: Gaussian case," *IEEE Trans. Signal Process.*, vol. 54, pp. 1131–1143, Mar. 2006.

[3] T. Aysal and K. Barner, "Constrained decentralized estimation over noisy channels for sensor networks," *IEEE Trans. Signal Process.*, vol. 56, pp. 1398–1410, Apr. 2008.

[4] B. Murmann, "A/D converter trends: Power dissipation, scaling and digitally assisted architectures," *2008 IEEE Cust. Integr. Circuits Conf.*, pp. 105–112, Sep. 2008.

[5] P. Kinget, "Device mismatch and tradeoffs in the design of analog circuits," *IEEE J. Solid-State Circuits*, vol. 40, pp. 1212–1224, 2005.

[6] T. Sundstrom and A. Alvandpour, "Utilizing process variations for reference generation in a flash ADC," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 56, pp. 364–368, 2009.

[7] S. Park, Y. Palaskas, and M. Flynn, "A 4-GS/s 4-bit flash ADC in 0.18-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 42, pp. 1865–1872, 2007.

[8] C. Donovan and M. Flynn, "A 'Digital'6-bit ADC in 0.25 um CMOS," *Proc. Cust. Integr. Circuits Conf.*, vol. 37, pp. 432–437, 2002.

[9] C. Paulus *et al.*, "A 4GS/s 6b flash ADC in 0.13 μm CMOS," *VLSI Circuits*, pp. 420–423, 2004.

[10] 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, pp. 2825–2833, 2010.

[11] R. W. Keyes, "The effect of randomness in the distribution of impurity atoms on FET thresholds," *Appl. Phys.*, vol. 8, pp. 251–259, Nov. 1975.

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

[13] E. Lehmann and G. Casella, *Theory of Point Estimation*. Springer, 1998.