Scalar Quantization with Noisy Partitions and its Application to Flash ADC Design

Da Wang, Yury Polyanskiy and Gregory Wornell

Signals, Information
and Algorithms
Laboratory

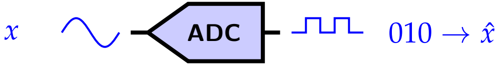
rle
AT MIT

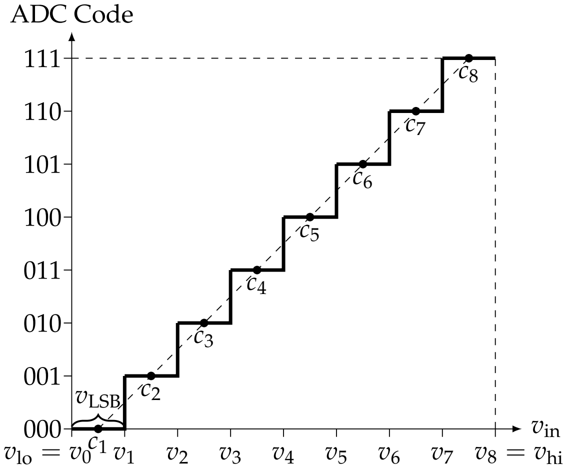
ISIT 2014, Honolulu, HI

June 30, 2014

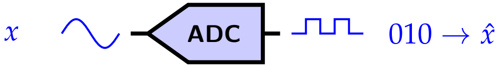
Outline

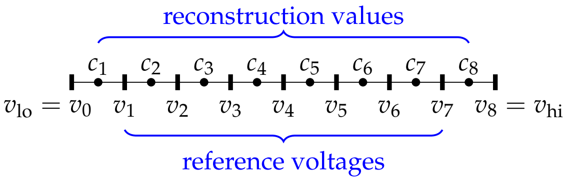
Analog-to-Digital Converter (ADC)





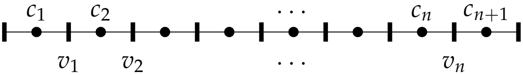
- 2^b reconstruction values
- $n = 2^b - 1$ reference voltages



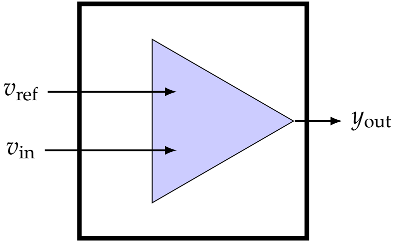


- 2^b reconstruction values
- $n = 2^b - 1$ reference voltages

ADC and its key building block: comparator

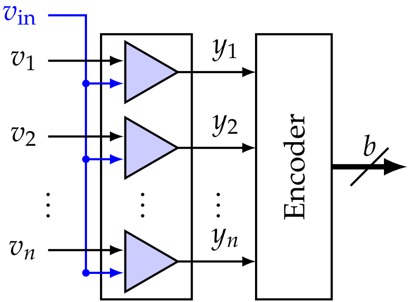


Comparator



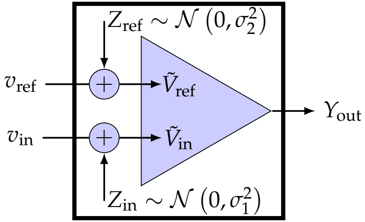
$$y_{\text{out}} = \begin{cases} 1 & v_{\text{in}} > v_{\text{ref}} \\ 0 & v_{\text{in}} \leq v_{\text{ref}} \end{cases}$$

The Flash ADC architecture



$$n = 2^b - 1$$

The imprecise comparator due to process variation



Z_{in} and Z_{ref}:

- offsets due to process variation
- variation $\nearrow$ as comparator size $\searrow$
- independent, zero-mean Gaussian distributed [Kinget 2005, Nuzzo 2008]

Note:

- fixed after fabrication
- randomness: over a collection of comparators
- aggregate variation:

$$Z = Z_{\text{ref}} - Z_{\text{in}} \sim \mathsf{N}\left(0, \sigma^2\right)$$

A call for mathematical framework

Existing theoretical error analysis (e.g., [Lundin 2005])

- assumes small process variation
- does not attempt to change the design

ADC design with imprecise comparators

Practice

- ADC with redundancy [Flynn *et al.*, 2003]
- ADC with redundancy, calibration and reconfiguration [Daly *et al.*, 2008]

A call for mathematical framework

Existing theoretical error analysis (e.g., [Lundin 2005])

- assumes small process variation
- does not attempt to change the design

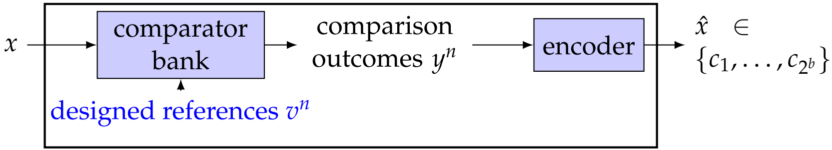
ADC design with imprecise comparators

Practice
- ADC with redundancy [Flynn *et al.*, 2003]
- ADC with redundancy, calibration and reconfiguration [Daly *et al.*, 2008]

Theory
- Little prior work
- Related: scalar quantizer with random thresholds for uniform input [Goyal 2011]

b-bit ADC

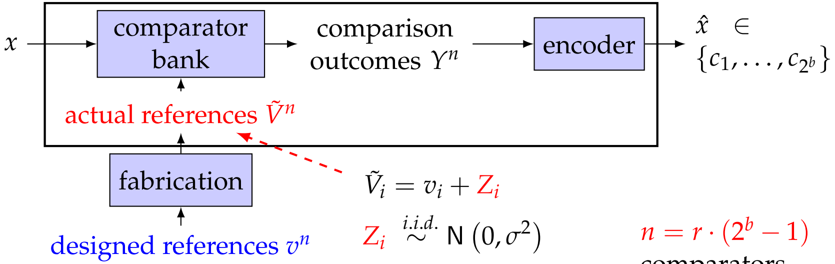
classical



$n = 2^b - 1$ comparators

b-bit ADC

with redundancy



$\tilde{V}_i = v_i + Z_i$

$Z_i \overset{i.i.d.}{\sim} \mathsf{N}\left(0, \sigma^2\right)$

$i = 1, 2, \ldots, n$

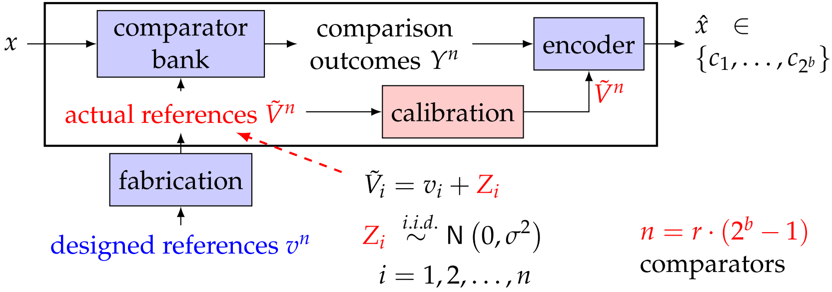
$n = r \cdot (2^b - 1)$ comparators

- "references" = "partition points"
- r: redundancy factor

System model: Scalar Quantization with Noisy Partition Points

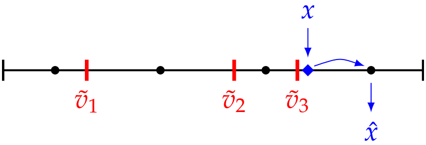
b-bit ADC

with redundancy and calibration



- "references" = "partition points"
- r: redundancy factor

Performance measures of ADC



error function

$$e(x) = x - \hat{x}$$

mean-square error

$$\mathrm{MSE} = \mathbb{E}_{X,\tilde{V}^n}\left[e(X)^2\right]$$

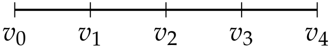
v^n — $\tilde{V}^n$ → analyze MSE

→ Given input distribution f_X, how to design optimal $v_1, v_2, \ldots, v_n$?

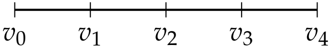
→ Is scaling down the size of comparators actually beneficial?

Challenge: randomness in partition points

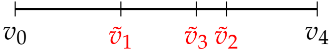
e.g., design:



e.g., design:

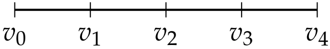


After fabrication:

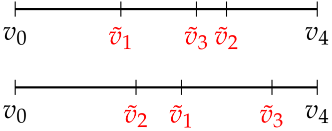


Challenge: randomness in partition points

e.g., design:

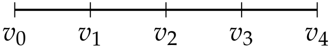


After fabrication:

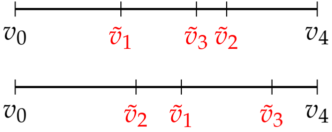


or …

e.g., design:



After fabrication:



or ...

Observations

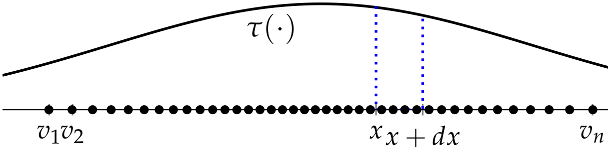
- Ordering may change $\rightarrow$ order statistics
- Random interval sizes $\rightarrow$?

interval sizes $\longleftrightarrow$ density of partition points $\longleftrightarrow$ high resolution approximation

Assume $n \to \infty$

- Represent v^n by point density functions $\tau(x)$

$$\tau(x)\,dx \approx \frac{\text{number of } v^n \text{ in } [x, x+dx]}{n}$$

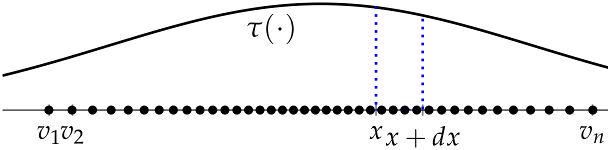


High resolution approximation

Assume $n \to \infty$

- Represent v^n by point density functions $\tau(x)$

$$\tau(x)\,dx \approx \frac{\text{number of } v^n \text{ in } [x, x+dx]}{n}$$

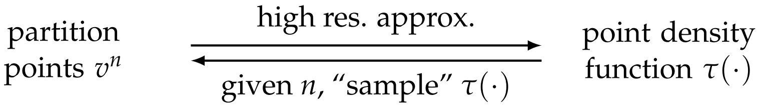


- $\tilde{V}^n$: point density functions $\lambda(x)$

$$\lambda(x)\,dx \approx \frac{\mathbb{E}\left[\text{number of } \tilde{V}^n \text{ in } [x, x+dx]\right]}{n}$$

- Point density function simplifies analysis!

Point density function guides partition point design



Examples

- $\tau \sim \mathsf{Unif}\left([-1,1]\right)$
- v^n: n-point evenly-spaced grid on [-1, 1]

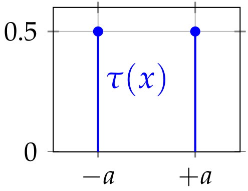
Point density function guides partition point design

partition points v^n —— high res. approx. ——→ point density function $\tau(\cdot)$

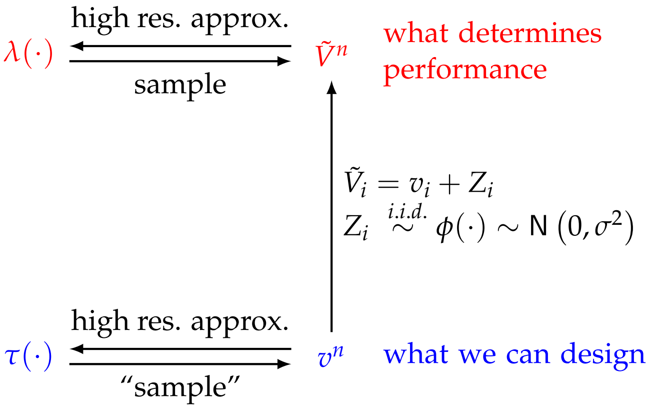
←—— given n, "sample" $\tau(\cdot)$ ——

Examples

- $\tau(x) = 0.5 \cdot \delta(x - a) + 0.5 \cdot \delta(x + a)$
- v^n:
  - $n/2$ points at $+a$
  - $n/2$ points at $-a$



With process variation, fabricated references matters



- Performance characterization in $\lambda(\cdot)$
- Want to find the optimal $\tau(\cdot)$

With process variation, fabricated references matters

$\lambda(\cdot)$ ⟵ high res. approx. / sample ⟶ $\tilde{V}^n$ — what determines performance

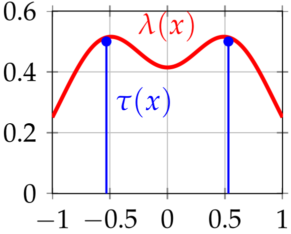
$\lambda(x) = (\tau * \phi)(x)$

($*$: convolution)

$\tilde{V}_i = v_i + Z_i$

$Z_i \overset{i.i.d.}{\sim} \phi(\cdot) \sim \mathsf{N}\left(0, \sigma^2\right)$

$\tau(\cdot)$ ⟵ high res. approx. / “sample” ⟶ v^n — what we can design



- Performance characterization in $\lambda(\cdot)$
- Want to find the optimal $\tau(\cdot)$

Process variation increases MSE 6-fold

Input $X \sim f_X(\cdot)$,

$$\text{MSE} = \mathbb{E}_X\left[e(X)^2\right]$$

classical case [Bennett 1948, Panter & Dite 1951]

$$\text{MSE} \simeq \frac{1}{12n^2} \int \frac{f_X(x)}{\lambda^2(x)}\, dx \qquad \lambda = \tau$$

with process variations

$$\text{MSE} \simeq \frac{1}{2n^2} \int \frac{f_X(x)}{\lambda^2(x)}\, dx \qquad \lambda = \tau * \phi$$

Why 6 times?

deterministic grid vs. random division of an interval (a topic in order statistics)

Optimal τ

- a necessary and sufficient condition

Optimal partition point density

Key function:

$$R(\tau) = \int f_X(x)(\tau * \phi)^{-2}(x)\, dx$$

Theorem

τ minimizes $R(\tau)$ if and only if

$$\sup_{x \in \mathcal{A}} \left[\frac{f_X}{(\tau * \phi)^3} * \phi \right](x) \leq \left\langle f_X, \frac{1}{(\tau * \phi)^2} \right\rangle .$$

In particular, if there exists τ^* such that

$$\tau^* * \phi \propto f_X^{1/3},$$

then τ^* minimizes $R(\tau)$ and

$$R(\tau^*) = \left(\int f_X^{1/3}(x) d\, x \right)^3 .$$

MSE-optimal designs can be quite different
Gaussian input distribution

Complete characterization of optimal τ

When

$$f_X \sim \mathsf{N}\left(0, \sigma_X^2\right),$$

then

$$\tau^* \sim \begin{cases} \mathsf{N}\left(0, 3\sigma_X^2 - \sigma^2\right) & \text{when } 3\sigma_X^2 > \sigma^2 \\ \delta(x) & \text{when } 3\sigma_X^2 \leq \sigma^2 \end{cases},$$

and

$$R(\tau^*) = \begin{cases} 6\sqrt{3}\pi\sigma_X^2 & \text{when } 3\sigma_X^2 > \sigma^2 \\ 2\pi\sigma^3 / \sqrt{\sigma^2 - 2\sigma_X^2} & \text{when } 3\sigma_X^2 \leq \sigma^2 \end{cases}.$$

MSE-optimal designs can be quite different

Uniform input distribution

$$f_X \sim \mathsf{Unif}\left([-1,1]\right)$$

$\sigma_0 \approx 0.7228$

$\sigma < \sigma_0$

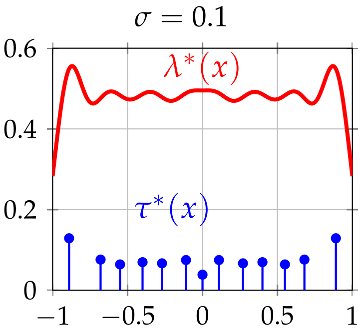
iterative optimization

$\Rightarrow$ locally optimal $\tau^*(x)$

$\sigma \geq \sigma_0$

the necessary and sufficient condition

$\Rightarrow \tau^*(x) = \delta(x)$



MSE-optimal designs can be quite different

Uniform input distribution

$$f_X \sim \mathsf{Unif}\left([-1, 1]\right)$$

$\sigma_0 \approx 0.7228$

$\sigma < \sigma_0$

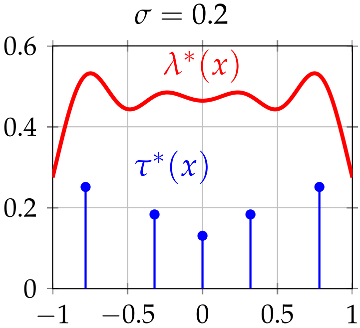
iterative optimization

$\Rightarrow$ locally optimal $\tau^*(x)$

$\sigma \geq \sigma_0$

the necessary and sufficient condition

$\Rightarrow \tau^*(x) = \delta(x)$



MSE-optimal designs can be quite different

Uniform input distribution

$$f_X \sim \mathsf{Unif}\left([-1,1]\right)$$

$\sigma_0 \approx 0.7228$

$\sigma < \sigma_0$

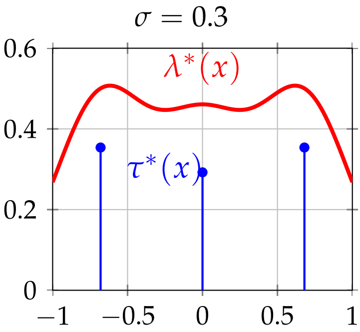
iterative optimization

$\Rightarrow$ locally optimal $\tau^*(x)$

$\sigma \geq \sigma_0$

the necessary and sufficient condition

$\Rightarrow \tau^*(x) = \delta(x)$



MSE-optimal designs can be quite different

Uniform input distribution

$$f_X \sim \mathsf{Unif}\left([-1,1]\right)$$

$\sigma_0 \approx 0.7228$

$\sigma < \sigma_0$

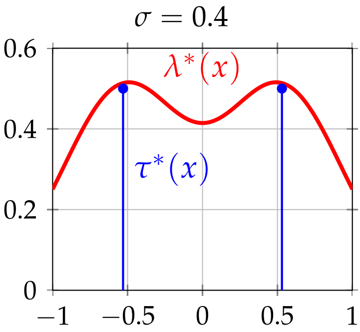
iterative optimization

$\Rightarrow$ locally optimal $\tau^*(x)$

$\sigma \geq \sigma_0$

the necessary and sufficient condition

$\Rightarrow \tau^*(x) = \delta(x)$



MSE-optimal designs can be quite different
Uniform input distribution

$$f_X \sim \mathsf{Unif}\left([-1,1]\right)$$

$\sigma_0 \approx 0.7228$

$\sigma < \sigma_0$

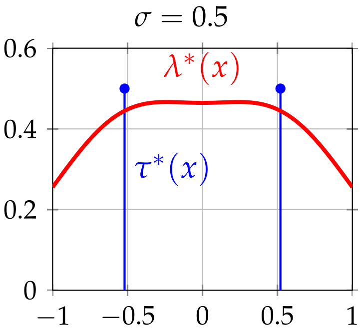
iterative optimization

$\Rightarrow$ locally optimal $\tau^*(x)$

$\sigma \geq \sigma_0$

the necessary and sufficient condition

$\Rightarrow \tau^*(x) = \delta(x)$



MSE-optimal designs can be quite different

Uniform input distribution

$$f_X \sim \mathsf{Unif}\left([-1,1]\right)$$

$\sigma_0 \approx 0.7228$

$\sigma < \sigma_0$

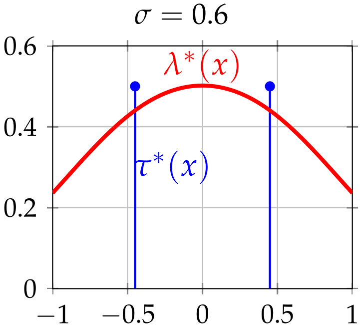
iterative optimization

$\Rightarrow$ locally optimal $\tau^*(x)$

$\sigma \geq \sigma_0$

the necessary and sufficient condition

$\Rightarrow \tau^*(x) = \delta(x)$



MSE-optimal designs can be quite different

Uniform input distribution

$$f_X \sim \mathsf{Unif}\left([-1,1]\right)$$

$\sigma_0 \approx 0.7228$

$\sigma < \sigma_0$

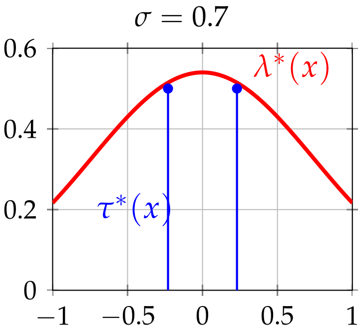
iterative optimization

$\Rightarrow$ locally optimal $\tau^*(x)$

$\sigma \geq \sigma_0$

the necessary and sufficient condition

$\Rightarrow \tau^*(x) = \delta(x)$



MSE-optimal designs can be quite different

Uniform input distribution

$$f_X \sim \mathsf{Unif}([-1, 1])$$

$\sigma_0 \approx 0.7228$

$\sigma < \sigma_0$

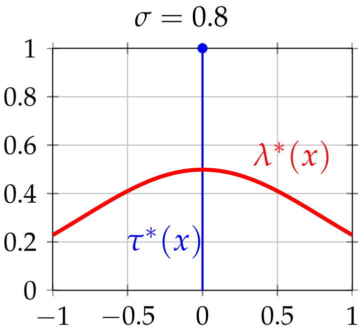
iterative optimization

$\Rightarrow$ locally optimal $\tau^*(x)$

$\sigma \geq \sigma_0$

the necessary and sufficient condition

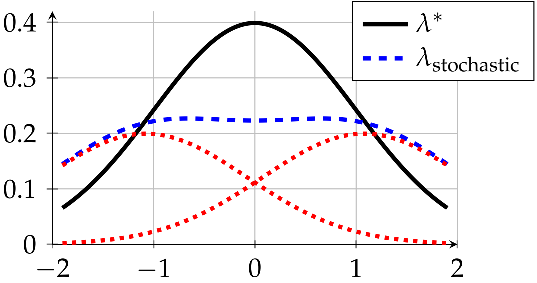
$\Rightarrow \tau^*(x) = \delta(x)$



ADC design implications

Stochastic ADC [Weaver 2010] is suboptimal

- Uniform input distribution over [-1, 1], $\sigma = 1.0$



Weaver:
$\tau(x) = \delta(x - 1.078\sigma) + \delta(x + 1.078\sigma)$

- Flatter $\lambda(x)$, but larger MSE
- Many points out of input range

Optimal: $\tau(x) = \delta(x)$

- Minimum MSE

$$\text{MSE}_{\text{stochastic}} / \text{MSE}^* \approx 2.15!$$

Scaling down the size of comparators is beneficial

For circuit fabrication [Kinget 2005, Nuzzo 2008],

$$\text{process variation} \qquad \sigma^2 \propto \frac{1}{\text{component area}}$$

Given a fixed silicon area,

$$\text{\# components} \qquad n \propto \frac{1}{\text{component area}}$$

Uniform input distribution, when $\sigma \geq \sigma_0$,

$$\text{MSE} \approx 2\pi\sigma^2/n^2 \quad \xRightarrow{\sigma^2 \propto n} \quad \text{MSE} = \Theta\left(1/n\right)$$

Scaling down the size of comparators is beneficial

For circuit fabrication [Kinget 2005, Nuzzo 2008],

$$\text{process variation} \quad \sigma^2 \propto \frac{1}{\text{component area}}$$

Given a fixed silicon area,

$$\text{\# components} \quad n \propto \frac{1}{\text{component area}}$$

Uniform input distribution, when $\sigma \geq \sigma_0$,

$$\text{MSE} \approx 2\pi\sigma^2/n^2 \quad \overset{\sigma^2 \propto n}{\Longrightarrow} \quad \text{MSE} = \Theta\left(1/n\right)$$

Building an ADC with more smaller but less precise comparators improves accuracy!

Recap

- Scalar quantization with noisy partitions
- High resolution analysis of MSE
- Optimal partition point designs difference from the classical case

Work in progress

- More error metrics: maximum quantization error, …
- Partial-calibration or no-calibration
- Take power consumption of ADC into account