Universidad
Carlos III de Madrid

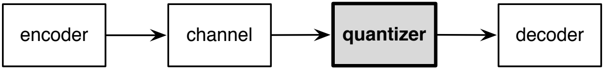
On the Dither-Quantized Gaussian Channel at Low SNR

TOBIAS KOCH
Universidad Carlos III de Madrid, Spain

2014 ITA Workshop (February 11, 2014)

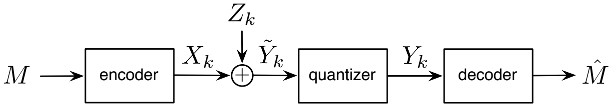
With thanks to Ram Zamir

Output Quantization



- Digital signal processing at receiver: quantize received signal.
- Quantizer: approximates output by finite number of bits (lossy).
- Loss in capacity negligible if quantizer has high resolution.
- High-resolution quantizers not feasible at high sampling rates.
- Extreme case: one-bit quantizer.

Gaussian Channel with One-Bit Quantization



- $\{Z_k\}$ is IID Gaussian noise of zero mean and variance σ^2.
- One bit quantizer characterized by quantization region $\mathcal{D} \subset \mathbb{R}$:

$$\begin{aligned} \tilde{Y}_k \in \mathcal{D} \quad &\to \quad Y_k = +1 \\ \tilde{Y}_k \notin \mathcal{D} \quad &\to \quad Y_k = -1 \end{aligned}$$

- Capacity under average-power constraint P:

$$C(\mathsf{P}) = \sup_{\substack{P_X:\ \mathsf{E}[X^2] \leq \mathsf{P} \\ \mathcal{D} \subset \mathbb{R}}} I(X;Y)$$

Low-SNR Asymptotic Capacity

- Low-SNR asymptotic capacity:

$$\dot{C}(0) \triangleq \lim_{\mathsf{P}\downarrow 0} \frac{C(\mathsf{P})}{\mathsf{P}}$$

$\Rightarrow$ slope of $C(\mathsf{P})$ at zero

$\Rightarrow$ relation to capacity per unit cost

- Free input symbol and memoryless channel:

$$\dot{C}(0) = \sup_{\xi \neq 0, \mathcal{D} \subset \mathbb{R}} \frac{D(P_{Y|X=\xi} \| P_{Y|X=0})}{\xi^2}$$

S. Verdú, "On channel capacity per unit cost," *IEEE Transactions on Information Theory*, September 1990.

2dB Power Loss for Symmetric Quantizer

- Symmetric quantizer $\mathcal{D} = \{\tilde{y} \in \mathbb{R} \colon \tilde{y} \geq 0\}$:

$$\dot{C}_{\mathsf{sym}}(0) = \frac{1}{\pi\sigma^2}, \quad \text{(achieved for } X = \pm\sqrt{\mathsf{P}} \text{ equiprobably)}$$

- Unquantized Gaussian channel:

$$\dot{C}_{\mathsf{G}}(0) = \frac{1}{2\sigma^2}, \quad \text{(achieved for } X = \pm\sqrt{\mathsf{P}} \text{ equiprobably)}$$

A. J. Viterbi, J. K. Omura. *Principles of Digital Communication and Coding*. McGraw-Hill, 1979.

J. Singh, O. Dabeer, U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Transactions on Communications*, December 2009.

C. Shannon, "A mathematical theory of communication," Bell System Technical Journal, July and October 1948.

2dB Power Loss for Symmetric Quantizer

- Symmetric quantizer $\mathcal{D} = \{\tilde{y} \in \mathbb{R} \colon \tilde{y} \geq 0\}$:

$$\dot{C}_{\text{sym}}(0) = \frac{1}{\pi\sigma^2}, \quad \text{(achieved for } X = \pm\sqrt{\mathsf{P}} \text{ equiprobably)}$$

- Unquantized Gaussian channel:

$$\dot{C}_{\text{G}}(0) = \frac{1}{2\sigma^2}, \quad \text{(achieved for } X = \pm\sqrt{\mathsf{P}} \text{ equiprobably)}$$

$$\Longrightarrow \quad \boxed{\frac{\dot{C}_{\text{sym}}(0)}{\dot{C}_{\text{G}}(0)} = \frac{2}{\pi} \approx -2\text{dB}}$$

A. J. Viterbi, J. K. Omura. *Principles of Digital Communication and Coding*. McGraw-Hill, 1979.

J. Singh, O. Dabeer, U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Transactions on Communications*, December 2009.

C. Shannon, "A mathematical theory of communication," Bell System Technical Journal, July and October 1948.

No 2dB power loss...

Theorem: The capacity per unit-energy of the Gaussian channel with one-bit output quantization is

$$\dot{C}(0) = \frac{1}{2\sigma^2}$$

Moreover, the optimal quantizer is of the form

$$\mathcal{D} = \{\tilde{y} \in \mathbb{R} \colon \tilde{y} \geq \Upsilon\}, \quad \text{for some } \Upsilon \in \mathbb{R}.$$

T. Koch, A. Lapidoth, "Asymmetric quantizers are better at low SNR," *2011 IEEE International Symposium on Information Theory (ISIT)*, St. Petersburg, Russia.

T. Koch, A. Lapidoth, "At low SNR, asymmetric quantizers are better," *IEEE Transactions on Information Theory*, September 2013.

...but poor spectral efficiency

Theorem: Every distribution on X satisfying $\mathsf{E}\left[X^2\right] \leq \mathsf{P}$ and

$$\lim_{\mathsf{P}\downarrow 0} \frac{I(X;Y)}{\mathsf{P}} = \frac{1}{2\sigma^2}$$

must be flash signaling.

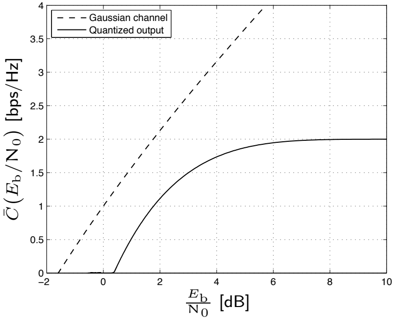
$\Rightarrow$ Flash signaling yields $\dot{C}(0) = \frac{1}{2\sigma^2}$ and $\ddot{C}(0) = -\infty$.

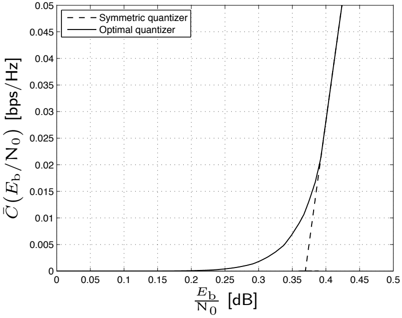
$\Rightarrow$ Flash signaling has poor spectral efficiency.

T. Koch, A. Lapidoth, "At low SNR, asymmetric quantizers are better," *IEEE Transactions on Information Theory*, September 2013.

S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, June 2002.

Flash Signaling = Bad!





T. Koch, A. Lapidoth, "At low SNR, asymmetric quantizers are better," *IEEE Transactions on Information Theory*, September 2013.

Avoid Flash Signaling

Theorem: If we replace the average-power constraint by a peak-power constraint

$$X_k^2 \leq \mathsf{P}, \quad \text{with probability one}$$

then

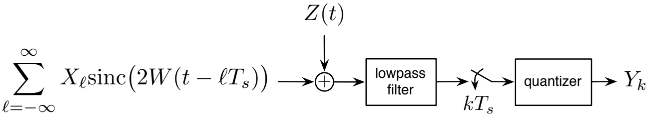
$$\dot{C}(0) = \frac{1}{\pi\sigma^2}.$$

$\Rightarrow$ Symmetric quantizer is asymptotically optimal.

$\Rightarrow$ 2dB power loss.

T. Koch, A. Lapidoth, "At low SNR, asymmetric quantizers are better," *IEEE Transactions on Information Theory*, September 2013.

Sampling and Quantizing



- Discrete-time Gaussian channel is equivalent to AWGN channel sampled at Nyquist rate.
- Sampling above Nyquist rate increases the low-SNR asymptotic capacity, even if $X_k = \pm\sqrt{\mathsf{P}}$.

E. N. Gilbert, "Increased information rate by oversampling," *IEEE Transactions on Information Theory*, July 1993.

S. Shamai (Shitz), "Information rates by oversampling the sign of a bandlimited process," *IEEE Transactions on Information Theory*, November 1993.

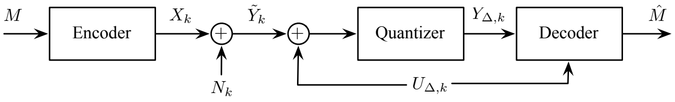
T. Koch, A. Lapidoth, "Increased capacity per unit-cost by oversampling," *2010 IEEE 26th Conv. on Electrical & Electronics Eng. in Israel*, Eilat, Israel.

Now What?

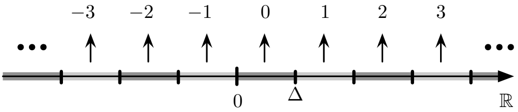
- Low-resolution quantizers with more than one bit?
- Ultimate tradeoff between quantizer resolution and sampling rate?
- Beyond low-SNR asymptotics?
- ...

$\Rightarrow$ **Analysis becomes intractable very quickly.**

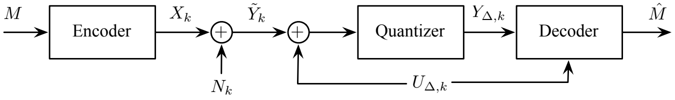
The Dither-Quantized Gaussian Channel



- Uniform, infinite-level, quantizer of step size Δ: $q_\Delta(x) = \left\lfloor \frac{x}{\Delta} \right\rfloor$
- Δ determines quantizer resolution.



The Dither-Quantized Gaussian Channel



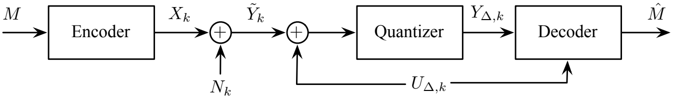
- Uniform, infinite-level, quantizer of step size Δ: $q_\Delta(x) = \left\lfloor \frac{x}{\Delta} \right\rfloor$
- Δ determines quantizer resolution.
- $Y_{\Delta,k} = q_\Delta(\tilde{Y}_k + U_{\Delta,k})$, where $\{U_{\Delta,k}\} \sim$ IID $\mathcal{U}([-\Delta/2, \Delta/2])$

  $\Rightarrow$ "dither"

- Average- and peak-power constraint:

$$\frac{1}{n}\sum_{k=1}^{n} x_k^2 \leq \mathsf{P} \quad \text{and} \quad |x_k| \leq \mathsf{A}$$

Channel Capacity



- Channel capacity:

$$C_\Delta(\mathsf{P},\mathsf{A}) = \sup_{\substack{P_X:\ \mathsf{E}[X^2]\leq \mathsf{P} \\ |X_k|\leq \mathsf{A} \text{ a.s.}}} I(X;Y_\Delta|U_\Delta)$$

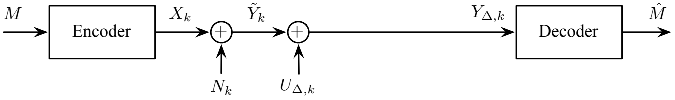
- Dithered quantizer is equivalent to additive-noise channel:

$$I(X;Y_\Delta|U_\Delta) = I(X;X+N+U_\Delta).$$

- Denote additive noise by $Z_\Delta \triangleq N + U_\Delta$.

R. Zamir, M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Transactions on Information Theory*, March 1992.

Channel Capacity



- Channel capacity:

$$C_\Delta(\mathsf{P},\mathsf{A}) = \sup_{\substack{P_X:\, \mathsf{E}[X^2]\leq \mathsf{P} \\ |X_k|\leq \mathsf{A} \text{ a.s.}}} I(X; Y_\Delta | U_\Delta)$$

- Dithered quantizer is equivalent to additive-noise channel:

$$I(X; Y_\Delta|U_\Delta) = I(X; X + N + U_\Delta).$$

- Denote additive noise by $Z_\Delta \triangleq N + U_\Delta$.

R. Zamir, M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Transactions on Information Theory*, March 1992.

Low-SNR Asymptotic Capacity

- No peak-power constraint ($\mathsf{A} = \infty$):

$$\dot{C}_\Delta^{(\infty)}(0) \triangleq \lim_{\mathsf{P}\downarrow 0} \frac{C_\Delta(\mathsf{P},\infty)}{\mathsf{P}} = \sup_{\xi\neq 0} \frac{D(P_{X+Z_\Delta|X=\xi} \| P_{X+Z_\Delta|X=0})}{\xi^2}$$

- Finite peak-to-average-power ratio $\mathsf{K} = \frac{\mathsf{A}^2}{\mathsf{P}}$:

$$\dot{C}_\Delta^{(\mathsf{K})}(0) \triangleq \lim_{\mathsf{P}\downarrow 0} \frac{C_\Delta(\mathsf{P},\sqrt{\mathsf{KP}})}{\mathsf{P}} = \frac{1}{2} I(0)$$

where $I(x) \triangleq \int_{-\infty}^{\infty} \frac{[\frac{\partial}{\partial x} f_{Z_\Delta}(y-x)]^2}{f_{Z_\Delta}(y-x)}\,\mathrm{d}y$ denotes the Fisher information.

S. Verdú, "On channel capacity per unit cost," *IEEE Transactions on Information Theory*, September 1990.

I.A. Ibragimov, R.Z. Khas'minskii, "Weak signal transmission in a memoryless channel," *Problemy Peredachi Informatsii*, October-December 1972.

V. Prelov, E. van der Meulen, "An asymptotic expression for the information and capacity of a multidimensional channel with weak input signals," *IEEE Transactions on Information Theory*, September 1993.

No Peak-Power Constraint

Theorem: When the peak-power constraint is absent ($\mathsf{A} = \infty$), irrespective of Δ,

$$\dot{C}_{\Delta}^{(\infty)}(0) = \frac{1}{2\sigma^2}.$$

- No peak-power constraint = no power loss.
- Is this surprising?

No Peak-Power Constraint

Theorem: When the peak-power constraint is absent ($\mathsf{A} = \infty$), irrespective of Δ,

$$\dot{C}_{\Delta}^{(\infty)}(0) = \frac{1}{2\sigma^2}.$$

- No peak-power constraint = no power loss.
- Is this surprising?

  $\Rightarrow$ Uniform quantizer performs not worse than one-bit quantizer.

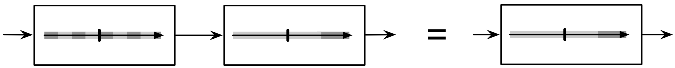
No Peak-Power Constraint

Theorem: When the peak-power constraint is absent ($\mathsf{A} = \infty$), irrespective of Δ,

$$\dot{C}_{\Delta}^{(\infty)}(0) = \frac{1}{2\sigma^2}.$$

- No peak-power constraint = no power loss.
- Is this surprising?

  $\Rightarrow$ Uniform quantizer performs not worse than one-bit quantizer.



No Peak-Power Constraint

Theorem: When the peak-power constraint is absent ($\mathsf{A} = \infty$), irrespective of Δ,

$$\dot{C}_{\Delta}^{(\infty)}(0) = \frac{1}{2\sigma^2}.$$

- No peak-power constraint = no power loss.
- Is this surprising?
  - ⇒ Uniform quantizer performs not worse than one-bit quantizer.
  - ⇒ But dither may reduce capacity.

A Glimpse of the Proof

- Show that

$$\sup_{\xi \neq 0} \frac{D(P_{X+Z_\Delta|X=\xi} \| P_{X+Z_\Delta|X=0})}{\xi^2} \geq \frac{1}{2\sigma^2}$$

- Let $V \triangleq \mathrm{I}\{X + Z_\Delta \geq \Delta\ell_0 - \delta\}$ for some ℓ_0 and δ. By the data processing inequality for relative entropy

$$D(P_{X+Z_\Delta|X=\xi} \| P_{X+Z_\Delta|X=0}) \geq D(P_{V|X=\xi} \| P_{V|X=0})$$

- V can be viewed as output of one-bit quantizer with threshold $\Delta\ell_0 - \delta$ and input $X + Z_\Delta$

  $\Rightarrow$ replicate steps of one-bit quantized case.

T. Koch, A. Lapidoth, "At low SNR, asymmetric quantizers are better," *IEEE Transactions on Information Theory*, September 2013.

Finite Peak-to-Average-Power Ratios

Theorem: For finite peak-to-average-power ratios $\mathsf{K} = \mathsf{A}^2/\mathsf{P}$, irrespective of K,

$$\dot{C}_{\Delta}^{(\mathsf{K})}(0) = \frac{1}{\Delta}\frac{1}{4\pi\sigma^2}\int_{-\infty}^{\infty}\frac{\left[e^{-\frac{(y-\Delta/2)^2}{2\sigma^2}} - e^{-\frac{(y+\Delta/2)^2}{2\sigma^2}}\right]^2}{Q\left(\frac{y-\Delta/2}{\sigma}\right) - Q\left(\frac{y+\Delta/2}{\sigma}\right)}\,\mathrm{d}y$$

Finite Peak-to-Average-Power Ratios

Theorem: For finite peak-to-average-power ratios $\mathsf{K} = \mathsf{A}^2/\mathsf{P}$, irrespective of K,

$$\dot{C}_{\Delta}^{(\mathsf{K})}(0) = \frac{1}{\Delta}\frac{1}{4\pi\sigma^2}\int_{-\infty}^{\infty} \frac{\left[e^{-\frac{(y-\Delta/2)^2}{2\sigma^2}} - e^{-\frac{(y+\Delta/2)^2}{2\sigma^2}}\right]^2}{Q\left(\frac{y-\Delta/2}{\sigma}\right) - Q\left(\frac{y+\Delta/2}{\sigma}\right)}\,\mathrm{d}y$$

Corollary: For every $\mathsf{K} = \mathsf{A}^2/\mathsf{P}$,

$$\lim_{\Delta\downarrow 0} \dot{C}_{\Delta}^{(\mathsf{K})}(0) = \frac{1}{2\sigma^2}$$

$$\lim_{\Delta\to\infty} \dot{C}_{\Delta}^{(\mathsf{K})}(0) = 0.$$

A Glimpse of the Proof

- For finite peak-to-average-power ratios,

$$\dot{C}_{\Delta}^{(\mathsf{K})}(0) = \frac{1}{2} I(0)$$

provided that a number of technical conditions are satisfied.

- $I(0) = \dfrac{1}{\Delta} \dfrac{1}{2\pi\sigma^2} \displaystyle\int_{-\infty}^{\infty} \frac{\left[e^{-\frac{(y-\Delta/2)^2}{2\sigma^2}} - e^{-\frac{(y+\Delta/2)^2}{2\sigma^2}} \right]^2}{Q\left(\frac{y-\Delta/2}{\sigma}\right) - Q\left(\frac{y+\Delta/2}{\sigma}\right)} \, \mathrm{d}y$

- Main work is to show that the required conditions hold.

I.A. Ibragimov, R.Z. Khas'minskii, "Weak signal transmission in a memoryless channel," *Problemy Peredachi Informatsii*, October-December 1972.

V. Prelov, E. van der Meulen, "An asymptotic expression for the information and capacity of a multidimensional channel with weak input signals," *IEEE Transactions on Information Theory*, September 1993.

Dithered vs. One-Bit Quantization

- Same low-SNR behavior as one-bit quantized Gaussian channel if peak-power constraint is relaxed ($A = \infty$).
- Same low-SNR behavior as (unquantized) Gaussian channel in the high-resolution limit ($\Delta \downarrow 0$).
- Different low-SNR behavior than one-bit quantized Gaussian channel in the low-resolution limit ($\Delta \to \infty$):

one-bit quantization:

$$\dot{C}(0) = \frac{1}{\pi\sigma^2}$$

dithered quantization:

$$\lim_{\Delta\to\infty} \dot{C}_{\Delta}^{(K)}(0) = 0$$

$\Longrightarrow$ **For finite peak-to-average-power ratios and in the low-resolution limit ($\Delta \to \infty$), dither is detrimental!**

Hopes and Worries

Hopes:

- Capacity of dither-quantized channel easier to analyze.
- Better understanding of loss due to low-resolution quantization:
  - $\Rightarrow$ Beyond one-bit quantization.
  - $\Rightarrow$ Beyond low-SNR asymptotics.
  - $\Rightarrow$ Ultimate tradeoff between quantizer resolution and sampling rate.

Worries:

- Dither detrimental when quantizer resolution is small.
- Overload region of infinite-level quantizer.