Finite Frames with Applications to Sigma Delta Quantization and MRI

joint work with

Matt Fickus
Alex Powell and Özgür Yılmaz
Alfredo Nava-Tudela and Stuart Fletcher
Yang B. Wang

Frames

Frames $F = \{e_n\}_{n=1}^N$ for d-dimensional Hilbert space H, e.g., $H = \mathbb{K}^d$, where $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$.

Frames

Frames $F = \{e_n\}_{n=1}^{N}$ for d-dimensional Hilbert space H, e.g., $H = \mathbb{K}^d$, where $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$.

- Any spanning set of vectors in $\mathbb{K}^d$ is a *frame* for $\mathbb{K}^d$.

Frames

Frames $F = \{e_n\}_{n=1}^N$ for d-dimensional Hilbert space H, e.g., $H = \mathbb{K}^d$, where $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$.

- Any spanning set of vectors in $\mathbb{K}^d$ is a *frame* for $\mathbb{K}^d$.

- $F \subseteq \mathbb{K}^d$ is A-tight if

$$\forall x \in \mathbb{K}^d, \ \ A\|x\|^2 = \sum_{n=1}^{N} |\langle x, e_n \rangle|^2$$

  (A=1 defines *Parseval frames*).

- F is *unit norm* if each $\|e_n\| = 1$.

Frames

Frames $F = \{e_n\}_{n=1}^N$ for d-dimensional Hilbert space H, e.g., $H = \mathbb{K}^d$, where $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$.

- Any spanning set of vectors in $\mathbb{K}^d$ is a *frame* for $\mathbb{K}^d$.

- $F \subseteq \mathbb{K}^d$ is A-tight if

$$\forall x \in \mathbb{K}^d, \;\; A\|x\|^2 = \sum_{n=1}^{N} |\langle x, e_n\rangle|^2$$

  (A=1 defines *Parseval frames*).

- F is *unit norm* if each $\|e_n\| = 1$.

- *Bessel map* – $L : H \longrightarrow \ell^2(\mathbb{Z}_N)$,

$$x \longmapsto \{< x, e_n >\}.$$

- *Frame operator* – $S = L^*L : H \longrightarrow H$.

Tight frames and applications

Theorem $\{e_n\}_{n=1}^N \subseteq \mathbb{K}^d$ is an A-tight frame for $\mathbb{K}^d \iff$

$$S = L^*L = AI : \mathbb{K}^d \longrightarrow \mathbb{K}^d.$$

Tight frames and applications

Theorem $\{e_n\}_{n=1}^{N} \subseteq \mathbb{K}^d$ is an A-tight frame for $\mathbb{K}^d \iff$

$$S = L^*L = AI : \mathbb{K}^d \longrightarrow \mathbb{K}^d.$$

- Robust transmission of data over erasure channels such as the Internet. [Casazza, Goyal, Kelner, Kovačević]

- Multiple antenna code design for wireless communications. [Hochwald, Marzetta, T. Richardson, Sweldens, Urbanke]

- Multiple description coding. [Goyal, Heath, Kovačević, Strohmer, Vetterli]

- Quantum detection. [Chandler Davis, Eldar, Forney, Oppenheim]

Finite unit norm tight frames (FUN-TFs)

- If $\{e_n\}_{n=1}^{N}$ is a finite unit norm tight frame (A-FUN-TF) for $\mathbb{K}^d$, then $A = N/d$.

Finite unit norm tight frames (FUN-TFs)

- If $\{e_n\}_{n=1}^{N}$ is a finite unit norm tight frame (A-FUN-TF) for $\mathbb{K}^d$, then $A = N/d$.

- Let $\{e_n\}$ be an A-unit norm TF for any separable Hilbert space H. $A \geq 1$, and $A = 1 \Leftrightarrow \{e_n\}$ is an ONB for H (*Vitali*).

Finite unit norm tight frames (FUN-TFs)

- If $\{e_n\}_{n=1}^N$ is a finite unit norm tight frame (A-FUN-TF) for $\mathbb{K}^d$, then $A = N/d$.

- Let $\{e_n\}$ be an A-unit norm TF for any separable Hilbert space H. $A \geq 1$, and $A = 1 \Leftrightarrow \{e_n\}$ is an ONB for H (*Vitali*).

- The geometry of finite tight frames:
  - The vertices of platonic solids are FUN-TFs.
  - Points that constitute FUN-TFs do not have to be equidistributed, e.g., Grassmanian frames.
  - FUN-TFs can be characterized as minimizers of a "frame potential function" (with Fickus) analogous to

Coulomb's Law.

Frame force and potential energy

$$F : S^{d-1} \times S^{d-1} \setminus D \longrightarrow \mathbb{R}^d$$

$$P : S^{d-1} \times S^{d-1} \setminus D \longrightarrow \mathbb{R},$$

where

$$P(a,b) = p(\|a-b\|), \qquad p'(x) = -xf(x)$$

Frame force and potential energy

$$F : S^{d-1} \times S^{d-1} \setminus D \longrightarrow \mathbb{R}^d$$

$$P : S^{d-1} \times S^{d-1} \setminus D \longrightarrow \mathbb{R},$$

where

$$P(a, b) = p(\|a - b\|), \qquad p'(x) = -xf(x)$$

- Coulomb force

$$CF(a, b) = (a-b)/\|a - b\|^3, \qquad f(x) = 1/x^3$$

Frame force and potential energy

$$F : S^{d-1} \times S^{d-1} \setminus D \longrightarrow \mathbb{R}^d$$

$$P : S^{d-1} \times S^{d-1} \setminus D \longrightarrow \mathbb{R},$$

where

$$P(a,b) = p(\|a-b\|), \qquad p'(x) = -xf(x)$$

- Coulomb force

$$CF(a,b) = (a-b)/\|a-b\|^3, \qquad f(x) = 1/x^3$$

- Frame force

$$FF(a,b) = < a, b > (a-b), \qquad f(x) = 1 - x^2/2$$

- Total potential energy for the frame force

$$TFP(\{x_n\}) = \Sigma_{m=1}^{N} \Sigma_{n=1}^{N} | < x_m, x_n > |^2$$

Characterization of FUN-TFs

For the Hilbert space $H = \mathbb{R}^d$ and N, consider

$$\{x_n\}_1^N \in S^{d-1} \times ... \times S^{d-1}$$

and

$$TFP(\{x_n\}) = \Sigma_{m=1}^N \Sigma_{n=1}^N |<x_m, x_n>|^2.$$

Theorem Let $N \leq d$. The minimum value of TFP, for the frame force and N variables, is N; and the *minimizers* are precisely the orthonormal sets of N elements for $\mathbb{R}^d$.

Theorem Let $N \geq d$. The minimum value of TFP, for the frame force and N variables, is N^2/d; and the *minimizers* are precisely the FUN-TFs of N elements for $\mathbb{R}^d$.

Characterization of FUN-TFs

For the Hilbert space $H = \mathbb{R}^d$ and N, consider

$$\{x_n\}_1^N \in S^{d-1} \times ... \times S^{d-1}$$

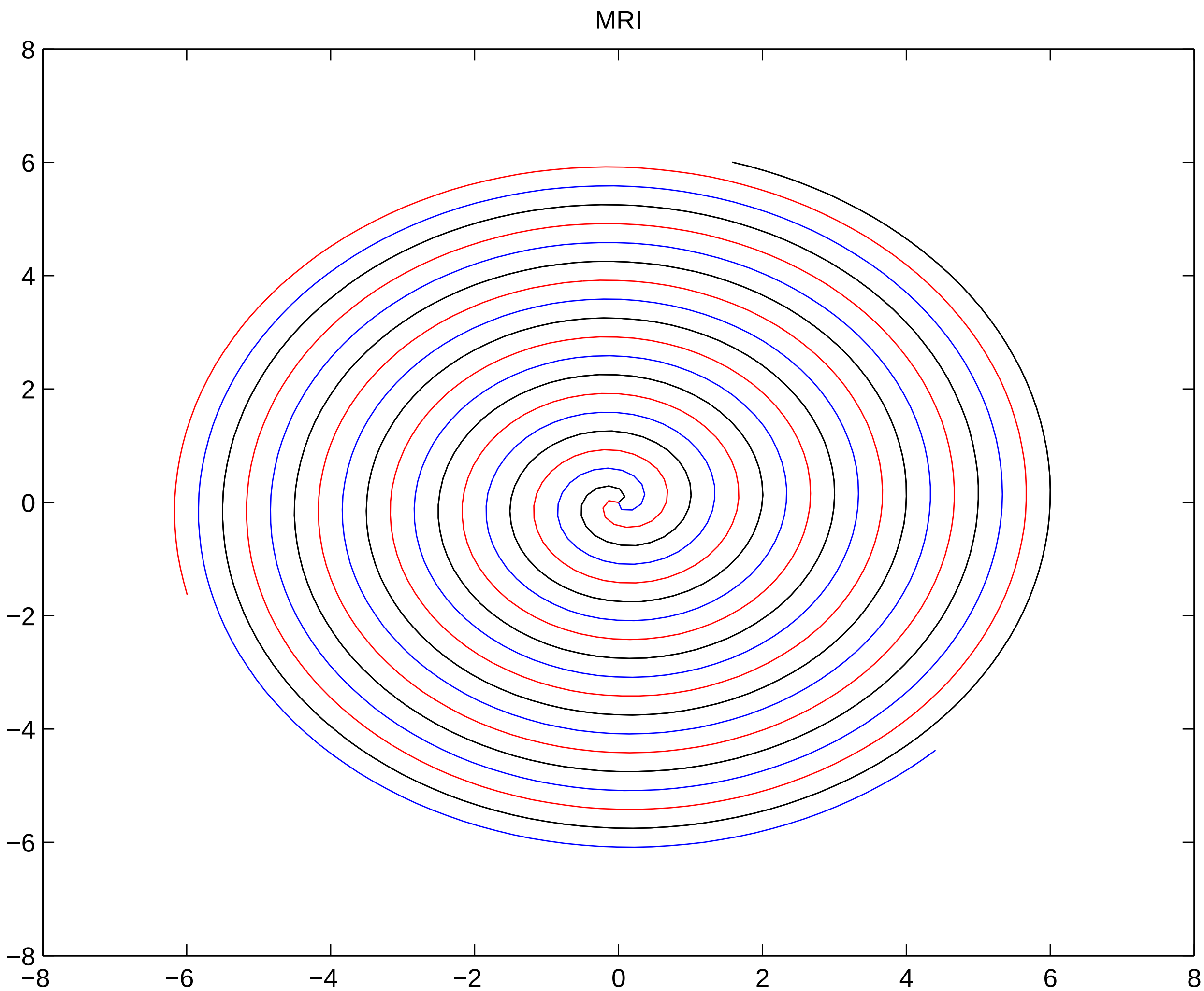
and

$$TFP(\{x_n\}) = \Sigma_{m=1}^N \Sigma_{n=1}^N | < x_m, x_n > |^2.$$

Theorem Let $N \leq d$. The minimum value of TFP, for the frame force and N variables, is N; and the *minimizers* are precisely the orthonormal sets of N elements for $\mathbb{R}^d$.

Theorem Let $N \geq d$. The minimum value of TFP, for the frame force and N variables, is N^2/d; and the *minimizers* are precisely the FUN-TFs of N elements for $\mathbb{R}^d$.

Problem Find these FUN-TFs analytically, effectively, and computationally.


MRI
8
6
4
2
0
−2
−4
−6
−8
−8
−6
−4
−2
0
2
4
6
8

Imaging equation

- The MR signal (or FID or echo) measured for imaging is

$$S(t) = S(k(t)) = S(k_x(t), k_y(t), k_z(t))$$

$$= \int\int\int \rho(x,y,z) \, \mathrm{exp}[-2\pi i < (x,y,z),$$

$$(k_x(t), k_y(t), k_z(t)) >] e^{-t/T_2} dx \; dy \; dz$$

where

$$k_x(t) = \gamma \int_0^t G_x(u) du.$$

$G_x(t)$ is an x-directional time varying gradient and ρ is the spin density function.

Imaging equation

- The MR signal (or FID or echo) measured for imaging is

$$S(t) = S(k(t)) = S(k_x(t), k_y(t), k_z(t))$$

$$= \int\int\int \rho(x,y,z) \exp[-2\pi i < (x,y,z),$$

$$(k_x(t), k_y(t), k_z(t)) >] e^{-t/T_2} dx\ dy\ dz$$

where

$$k_x(t) = \gamma \int_0^t G_x(u) du.$$

$G_x(t)$ is an x-directional time varying gradient and ρ is the spin density function.

- The imaging equation is a consequence of Bloch's equation for transverse magnetization.

- The imaging equation is a physical Fourier transformer.

Spiral echo planar imaging (SEPI)

- Design gradients G as input to the MR process resulting in the imaging equation.

- Set

$$G_x(t) = \eta \cos \xi t - \eta \xi t \sin \xi t$$

$$G_y(t) = \eta \sin \xi t + \eta \xi t \cos \xi t.$$

Then $k_x(t) = \gamma \eta t \cos \xi t$ and $k_y(t) = \gamma \eta t \sin \xi t$. k_x and k_y yield the *Archimedean spiral*

$$A_c = \{(c\theta \cos 2\pi\theta, c\theta \sin 2\pi\theta) : \theta \geq 0\} \subseteq \widehat{\mathbb{R}}^2,$$

$\theta = \theta(t) = (1/2\pi)\xi t$, $c = (1/\theta)\gamma\eta = 2\pi\gamma\eta/\xi$.

Spiral echo planar imaging (SEPI)

- Design gradients G as input to the MR process resulting in the imaging equation.

- Set

$$G_x(t) = \eta \cos \xi t - \eta \xi t \sin \xi t$$

$$G_y(t) = \eta \sin \xi t + \eta \xi t \cos \xi t.$$

  Then $k_x(t) = \gamma \eta t \cos \xi t$ and $k_y(t) = \gamma \eta t \sin \xi t$. k_x and k_y yield the *Archimedean spiral*

$$A_c = \{(c\theta \cos 2\pi\theta, c\theta \sin 2\pi\theta) : \theta \geq 0\} \subseteq \widehat{\mathbb{R}}^2,$$

  $\theta = \theta(t) = (1/2\pi)\xi t$, $c = (1/\theta)\gamma\eta = 2\pi\gamma\eta/\xi$.

- S takes values on A_c.

- Spiral scanning gives high speed data acquisition at the "Nyquist rate".

- Rectilinear scanning is expensive at the corners.

Fourier frames for $L^2(E)$

Let $E \subseteq \mathbb{R}^d$ be closed. The *Paley-Wiener* space PW_E is

$$PW_E = \left\{ \varphi \in L^2(\widehat{\mathbb{R}}^d) : \text{supp } \varphi^\vee \subseteq E \right\},$$

where

$$\varphi^\vee(x) = \int_{\widehat{\mathbb{R}}^d} \varphi(\gamma) e^{2\pi i x \cdot \gamma} d\gamma.$$

Fourier frames for $L^2(E)$

Let $E \subseteq \mathbb{R}^d$ be closed. The *Paley-Wiener* space PW_E is

$$PW_E = \left\{\varphi \in L^2(\widehat{\mathbb{R}}^d) : \text{supp } \varphi^\vee \subseteq E\right\},$$

where

$$\varphi^\vee(x) = \int_{\widehat{\mathbb{R}}^d} \varphi(\gamma) e^{2\pi i x \cdot \gamma} d\gamma.$$

Definition Given $\Lambda \subseteq \widehat{\mathbb{R}}^d$ and $E \subseteq \mathbb{R}^d$ with finite Lebesgue measure. Let $e_{-\lambda}(x) = e^{-2\pi i x \cdot \lambda}$. $\{e_{-\lambda} : \lambda \in \Lambda\}$ is a *frame* for $L^2(E)$ if

$$\forall \varphi \in PW_E, A\|\varphi\|_{L^2(\widehat{\mathbb{R}}^d)} \leq \sum_{\lambda \in \Lambda} |\varphi(\lambda)|^2 \leq B\|\varphi\|_{L^2(\widehat{\mathbb{R}}^d)}.$$

In this case we say that $\Lambda \subseteq \widehat{\mathbb{R}}^d$ is a *Fourier frame* for $L^2(E) \subseteq L^2(\mathbb{R}^d)$.

Fourier frames for $L^2(E)$

Let $E \subseteq \mathbb{R}^d$ be closed. The *Paley-Wiener* space PW_E is

$$PW_E = \left\{ \varphi \in L^2(\widehat{\mathbb{R}}^d) : \text{supp}\, \varphi^\vee \subseteq E \right\},$$

where

$$\varphi^\vee(x) = \int_{\widehat{\mathbb{R}}^d} \varphi(\gamma) e^{2\pi i x \cdot \gamma} d\gamma.$$

Definition Given $\Lambda \subseteq \widehat{\mathbb{R}}^d$ and $E \subseteq \mathbb{R}^d$ with finite Lebesgue measure. Let $e_{-\lambda}(x) = e^{-2\pi i x \cdot \lambda}$. $\{e_{-\lambda} : \lambda \in \Lambda\}$ is a *frame* for $L^2(E)$ if

$$\forall \varphi \in PW_E, A\|\varphi\|_{L^2(\widehat{\mathbb{R}}^d)} \leq \sum_{\lambda \in \Lambda} |\varphi(\lambda)|^2 \leq B\|\varphi\|_{L^2(\widehat{\mathbb{R}}^d)}.$$

In this case we say that $\Lambda \subseteq \widehat{\mathbb{R}}^d$ is a *Fourier frame* for $L^2(E) \subseteq L^2(\mathbb{R}^d)$.

Definition $\Lambda \subseteq \widehat{\mathbb{R}}^d$ is *uniformly discrete* if there is $r > 0$ such that

$$\forall \lambda, \gamma \in \Lambda, \qquad |\lambda - \gamma| \geq r.$$

A theorem of Beurling

Theorem Let $\Lambda \subseteq \widehat{\mathbb{R}}^d$ be uniformly discrete, and define

$$\rho = \rho(\Lambda) = \sup_{\zeta \in \widehat{\mathbb{R}}^d} \text{dist}(\zeta, \Lambda),$$

where $\text{dist}(\zeta, \Lambda)$ is Euclidean distance between ζ and Λ, and $B(0, R) \subseteq \mathbb{R}^d$ is closed ball centered at $0 \in \mathbb{R}^d$ with radius R. If $R\rho < 1/4$, then Λ is a Fourier frame for $PW_{B(0,R)}$.

A theorem of Beurling

Theorem Let $\Lambda \subseteq \widehat{\mathbb{R}}^d$ be uniformly discrete, and define

$$\rho = \rho(\Lambda) = \sup_{\zeta \in \widehat{\mathbb{R}}^d} \text{dist}(\zeta, \Lambda),$$

where $\text{dist}(\zeta, \Lambda)$ is Euclidean distance between ζ and Λ, and $B(0, R) \subseteq \mathbb{R}^d$ is closed ball centered at $0 \in \mathbb{R}^d$ with radius R. If $R\rho < 1/4$, then Λ is a Fourier frame for $PW_{B(0,R)}$.

Remark The assertion of Beurling's theorem implies

$$\forall f \in L^2(B(0,R)), \qquad f(x) = \sum_{\lambda \in \Lambda} a_\lambda(f) e^{2\pi i x \cdot \lambda}$$

in $L^2(B(0,R))$, where

$$\sum_{\lambda \in \Lambda} |a_\lambda(f)|^2 < \infty.$$

An MRI problem and mathematical solution

Given *any* $R > 0$ and $c > 0$. Consider the Archimedean spiral A_c.

We can show how to construct a finite interleaving set $B = \cup_{k=1}^{M-1} A_k$ of spirals

$$A_k = \left\{ c\theta e^{2\pi i(\theta - k/M)} : \theta \geq 0 \right\}, \quad k = 0, 1, \ldots, M-1,$$

and a uniformly discrete set $\Lambda_R \subseteq B$ such that Λ_R is a Fourier frame for $PW_{B(0,R)}$. Thus, all of the elements of $L^2(B(0,R))$ will have a decomposition in terms of the Fourier frame $\{e_\lambda : \lambda \in \Lambda_R\}$.

An MRI problem and mathematical solution

Given *any* $R > 0$ and $c > 0$. Consider the Archimedean spiral A_c.

We can show how to construct a finite interleaving set $B = \cup_{k=1}^{M-1} A_k$ of spirals

$$A_k = \left\{ c\theta e^{2\pi i(\theta - k/M)} : \theta \geq 0 \right\}, \quad k = 0, 1, \ldots, M-1,$$

and a uniformly discrete set $\Lambda_R \subseteq B$ such that Λ_R is a Fourier frame for $PW_{B(0,R)}$. Thus, all of the elements of $L^2(B(0,R))$ will have a decomposition in terms of the Fourier frame $\{e_\lambda : \lambda \in \Lambda_R\}$.

Method Combine Beurling's theorem and trigonometry.

Problem Although Λ_R is constructible, this mathematical solution must be effectively finitized and implemented to be of any use.

A first algorithm for implementation

- Given $N > 0$, e.g., $N = 256$, let $f \in L^2([0,1]^2)$.

- Assume f is piecewise constant (from pixel information) on

  $[m/N, (m+1)/(N+1)) \times [n/N, (n+1)/N)$,

  where $m, n \in \{0, 1, \ldots, N-1\}$.

- Write f lexicographically as $\{f_{a_k}\}$.

A first algorithm for implementation

- Given $N > 0$, e.g., $N = 256$, let $f \in L^2([0,1]^2)$.

- Assume f is piecewise constant on

  $[m/N, (m+1)/(N+1)) \times [n/N, (n+1)/N)$,

  where $m, n \in \{0, 1, \ldots, N-1\}$.

- Write f lexicographically as $\{f_{a_k}\}$.

- The Fourier transform of $f \in L^2(\mathbb{R}^2)$ is

$$\hat{f} = \sum_{k=0}^{N^2-1} f_{a_k} H_{a_k},$$

  where $e(\lambda) = e^{-2\pi i \lambda}$ and

$$H_{m,n}(\lambda, \gamma) = \frac{-1}{4\pi^2 \lambda\gamma} e(\frac{m\lambda + n\gamma}{N})(e(\frac{\lambda}{N})-1)(e(\frac{\gamma}{N})-1).$$

Problem Reconstruct $\{f_{a_k}\}$ from given $\hat{f}_{\alpha_m}$, where $\alpha_m = (\lambda_m, \gamma_m)$, $k = 0, 1, \ldots, N^2-1$, $m = 0, 1, \ldots, M-1$, and $M \geq N^2$.

The role of finite frames in implementation

Let $H = \mathbb{K}^{N^2}$ and take $M \geq N^2$. Given $\{a_k : k = 0, 1, \ldots N^2 - 1\}$, let $\alpha = (\lambda, \gamma) \in \widehat{\mathbb{R}}^2$, and choose α_m, $m = 0, 1, 2, \ldots, M - 1$.

The role of finite frames in implementation

Let $H = \mathbb{K}^{N^2}$ and take $M \geq N^2$. Given $\{a_k : k = 0, 1, \ldots N^2 - 1\}$, let $\alpha = (\lambda, \gamma) \in \widehat{\mathbb{R}}^2$, and choose α_m, $m = 0, 1, 2, \ldots, M - 1$.

- Let $x_m = (H_{a_0}(\alpha_m), H_{a_1}(\alpha_m), \ldots, H_{a_{N^2-1}}(\alpha_m))$.

- Define $L : H \longrightarrow \ell^2(\mathbb{Z}_M) = \mathbb{K}^M$,

$$f \longmapsto \{< f, x_m >\}_{m=0}^{M-1}.$$

- Equivalently, $L = (H_{a_k}(\alpha_m))$, $M \times N^2$.

The role of finite frames in implementation

Let $H = \mathbb{K}^{N^2}$ and take $M \geq N^2$. Given $\{a_k : k = 0, 1, \ldots N^2 - 1\}$, let $\alpha = (\lambda, \gamma) \in \widehat{\mathbb{R}}^2$, and choose α_m, $m = 0, 1, 2, \ldots, M-1$.

- Let $x_m = (H_{a_0}(\alpha_m), H_{a_1}(\alpha_m), \ldots, H_{a_{N^2-1}}(\alpha_m))$.

- Define $L : H \longrightarrow \ell^2(\mathbb{Z}_M) = \mathbb{K}^M$,
$$f \longmapsto \{< f, x_m >\}_{m=0}^{M-1}.$$

- Equivalently, $L = (H_{a_k}(\alpha_m))$, $M \times N^2$.

- $\{x_m\}_{m=0}^{M-1}$ frame for H implies L is a Bessel map and $S = L^* L$, an $N^2 \times N^2$ matrix, where L^* is the adjoint of L. S is the frame operator.

- S "reduces" dimensionality since $M \geq N^2$.

- The finite frame decomposition of f is
$$f = S^{-1} L^* (Lf).$$

Logic for empirical evaluation of algorithm

- Given high resolution image I, e.g., 1024×1024.
- Downsample I (room for "creativity") to I_N, $N \times N$, e.g., $N = 128, 256$.
- Therefore, I_N is the optimal, available image at $N \times N$ level for comparison purposes.

Logic for empirical evaluation of algorithm

- Given high resolution image I, e.g., 1024×1024.
- Downsample I (room for "creativity") to I_N, $N \times N$, e.g., $N = 128, 256$.
- Therefore, I_N is the optimal, available image at $N \times N$ level for comparison purposes.
- Calculate $\hat{I} = \sum I_{a_k} H_{a_k}$ (10^6 terms per α_m).
- Take $\hat{I}(\alpha_m)$, $m = 0, 1, \ldots, M-1 \geq N^2 - 1$.
- Set $LI = \hat{I}$, $M \times 1$
- Implementation gives $\tilde{I} = S^{-1} L^* \hat{I}$.
- Compare $I_N - \tilde{I}$, $N \times N$.

Optimal $N \times N$ approximant

Let $Q = [0,1]^2$ and set

$$\mathcal{S}_N = \left\{ f(x,y) \in L^2(Q) : f \sim \{f_{a_k} : k = 0, \ldots, N^2 - 1\} \right\}$$

Problem Find the optimal $\mathcal{S}_N$ approximant for $f \in L^2(Q)$.

Optimal $N \times N$ approximant

Let $Q = [0, 1]^2$ and set

$$\mathcal{S}_N = \left\{ f(x, y) \in L^2(Q) : f \sim \{f_{a_k} : k = 0, \ldots, N^2 - 1\} \right\}$$

Problem Find the optimal $\mathcal{S}_N$ approximant for $f \in L^2(Q)$.

Solution The minimizer of $\|f - g\|_2, g \in \mathcal{S}_N$ is

$$f_a(x, y) = \sum A(f)_{a_k} \mathbf{1}_{a_k}(x, y),$$

where $A(f)_{a_k}$ is the average of f over the a_k square, $k = 0, 1, \cdots, N^2 - 1$.

Asymptotic evaluation of algorithm

- Given $f \in L^2(Q)$, fix N $(N = 128, 256)$, and assume we know $\widehat{f}$ in k-space.

- Recall $\mathcal{S}_N = \Big\{g \in L^2(Q) : g \sim \{g_{a_k}\}\Big\}$.

- Take K, M, and

$$\{\alpha_m : m = 0, \ldots, M-1\} \subseteq [-K, K]^2 \subseteq \widehat{\mathbb{R}}^2.$$

- Denote N^2 data reconstructed by the algorithm from $\{\widehat{f}(\alpha_m)\}$ by

$$\tilde{f} = \tilde{f}_{M,K,\{\alpha_m\}} \in \mathcal{S}_N.$$

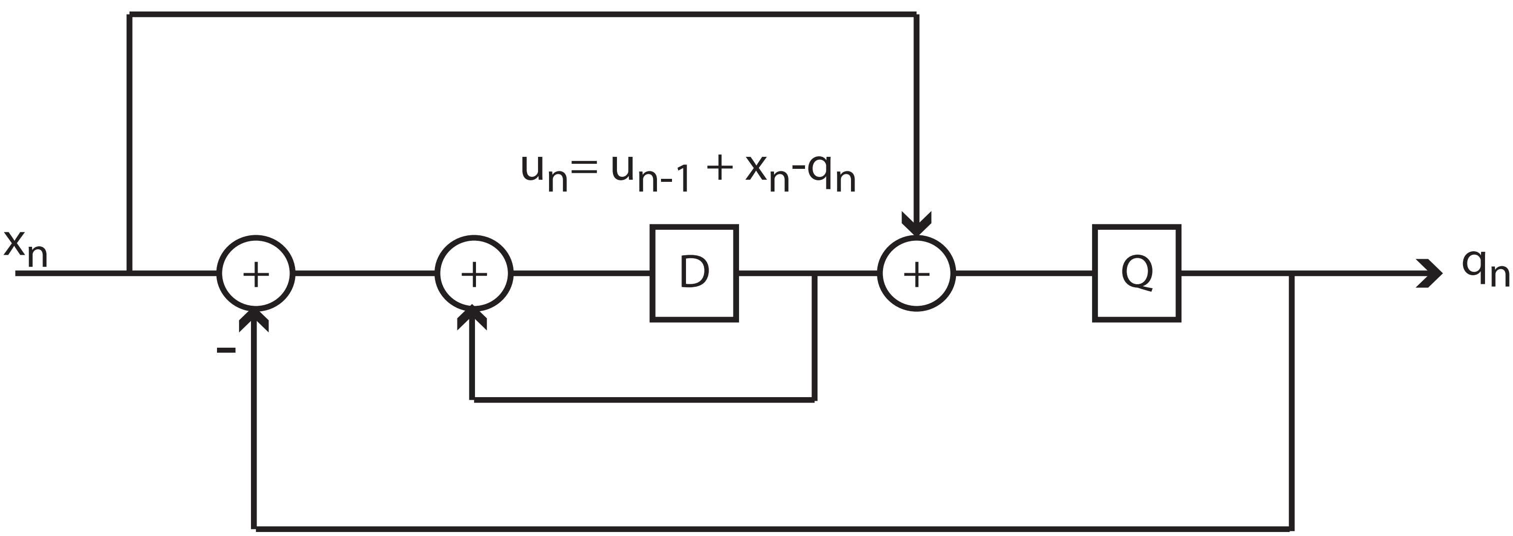
- $\lim \tilde{f} = \sum_1^{N^2-1} A(f)_{a_k} \mathbf{1}_{a_k}$.

- The limit as $M, K \longrightarrow \infty$ must be explained.

- Implementation of the Fourier frame algorithm approaches optimal $\mathcal{S}_N$ approximant.

Given u_0 and $\{x_n\}_{n=1}$

$$u_n = u_{n-1} + x_n - q_n$$
$$q_n = Q(u_{n-1} + x_n)$$



First Order ΣΔ

A quantization problem

Qualitative Problem Obtain *digital* representations for class X, suitable for storage, transmission, recovery.

A quantization problem

Qualitative Problem Obtain *digital* representations for class X, suitable for storage, transmission, recovery.

Quantitative Problem Find dictionary $\{e_n\} \subseteq X$:

1. Sampling

   $$\forall x \in X, \quad x = \sum x_n e_n, \quad x_n \in \mathbb{K} \ (\mathbb{R} \text{ or } \mathbb{C})$$

   [Continuous range $\mathbb{K}$ is not digital.]

2. Quantization. Construct finite alphabet $\mathcal{A}$ and

   $$Q: \ X \to \{\sum q_n e_n : \ q_n \in \mathcal{A} \subseteq \mathbb{K}\}$$

   such that $|x_n - q_n|$ and/or $\|x - Qx\|$ small.

A quantization problem

Qualitative Problem Obtain *digital* representations for class X, suitable for storage, transmission, recovery.

Quantitative Problem Find dictionary $\{e_n\} \subseteq X$:

1. Sampling

   $$\forall x \in X, \quad x = \sum x_n e_n, \quad x_n \in \mathbb{K} \text{ } (\mathbb{R} \text{ or } \mathbb{C})$$

   [Continuous range $\mathbb{K}$ is not digital.]

2. Quantization. Construct finite alphabet $\mathcal{A}$ and

   $$Q: \ X \to \{\sum q_n e_n : \ q_n \in \mathcal{A} \subseteq \mathbb{K}\}$$

   such that $|x_n - q_n|$ and/or $\|x - Qx\|$ small.

Methods

1. Fine quantization, e.g., PCM.
   Take $q_n \in \mathcal{A}$ close to given x_n. Reasonable in 16-bit (65,536 levels) digital audio.

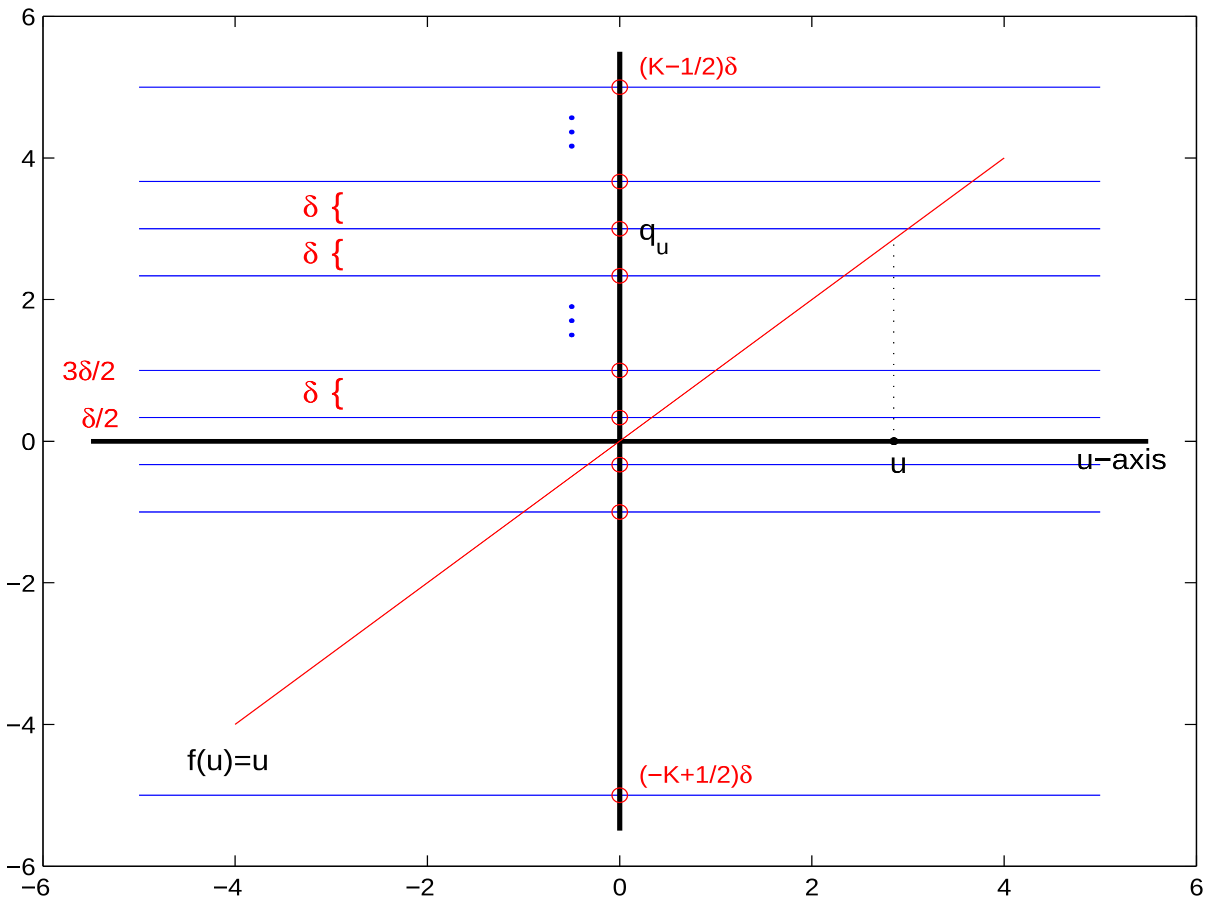
2. Coarse quantization, e.g., $\Sigma\Delta$.
   Use fewer bits to exploit redundancy of $\{e_n\}$ when sampling expansion is not unique.

Quantization

$$\mathcal{A}_K^{\delta} = \{(-K+1/2)\delta, (-K+3/2)\delta, \ldots, (-1/2)\delta, (1/2)\delta, \ldots, (K-1/2)\delta\}$$

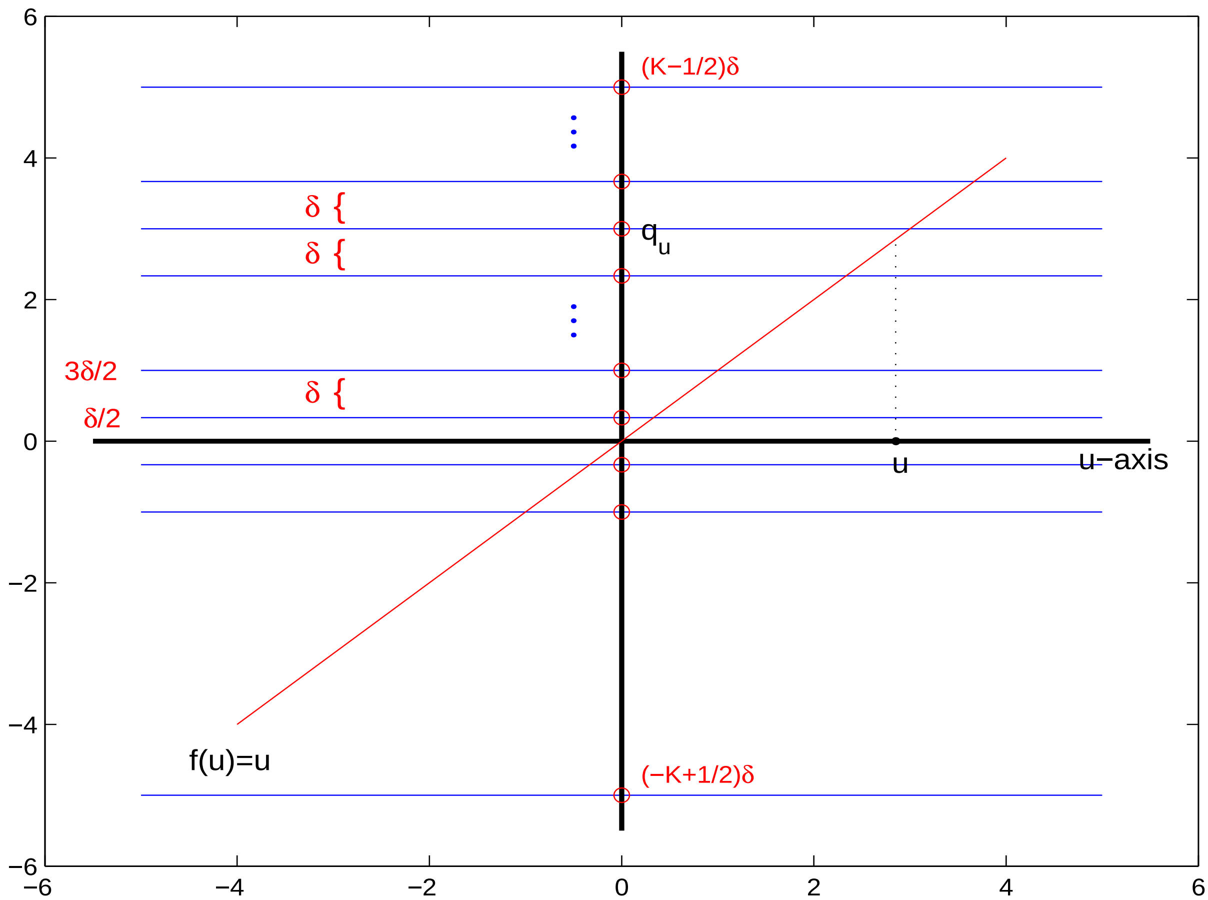
Quantization

$$\mathcal{A}_K^\delta = \{(-K+1/2)\delta, (-K+3/2)\delta, \ldots, (-1/2)\delta, (1/2)\delta, \ldots, (K-1/2)\delta\}$$



Quantization

$$\mathcal{A}_K^\delta = \{(-K+1/2)\delta, (-K+3/2)\delta, \ldots, (-1/2)\delta, (1/2)\delta, \ldots, (K-1/2)\delta\}$$



$$Q(u) = \arg\min_{q \in \mathcal{A}_K^\delta} |u - q| = q_u$$

Setting

Let $x \in \mathbb{R}^d$, $\|x\| \leq 1$. Suppose $F = \{e_n\}_{n=1}^N$ is a unit norm tight frame for $\mathbb{R}^d$. Thus, we have

$$x = \frac{d}{N} \sum_{n=1}^{N} x_n e_n$$

with $x_n = \langle x, e_n \rangle$. Note: $A = N/d$, and $|x_n| \leq 1$.

Goal Find a “good” quantizer, given

$$\mathcal{A}_K^\delta = \{(-K + \frac{1}{2})\delta, (-K + \frac{3}{2})\delta, \ldots, (K - \frac{1}{2})\delta\}.$$

Setting

Let $x \in \mathbb{R}^d$, $\|x\| \leq 1$. Suppose $F = \{e_n\}_{n=1}^N$ is a unit norm tight frame for $\mathbb{R}^d$. Thus, we have

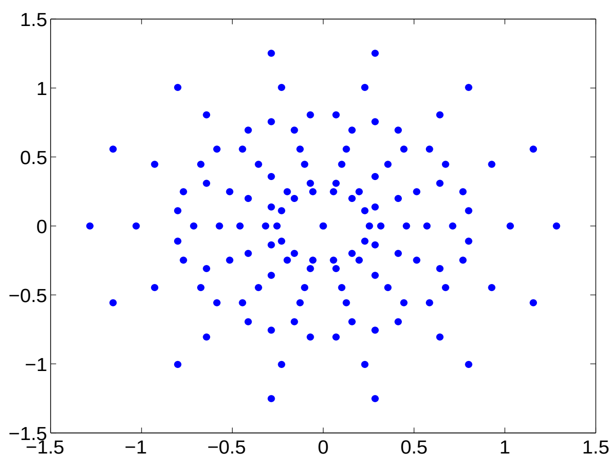
$$x = \frac{d}{N} \sum_{n=1}^{N} x_n e_n$$

with $x_n = \langle x, e_n \rangle$. Note: $A = N/d$, and $|x_n| \leq 1$.

Goal Find a "good" quantizer, given

$$\mathcal{A}_K^\delta = \{(-K + \frac{1}{2})\delta, (-K + \frac{3}{2})\delta, \ldots, (K - \frac{1}{2})\delta\}.$$

Example Consider the alphabet $\mathcal{A}_1^2 = \{-1, 1\}$, and $E_7 = \{e_n\}_{n=1}^7$, with $e_n = (\cos(\frac{2n\pi}{7}), \sin(\frac{2n\pi}{7}))$.



$\Gamma_{\mathcal{A}_1^2}(E_7) = \{\frac{2}{7}\sum_{n=1}^7 q_n e_n : \ q_n \in \mathcal{A}_1^2\}$

PCM

Replace $x_n \leftrightarrow q_n = \arg\min_{q \in \mathcal{A}_K^\delta} |x_n - q|$.

Then $\tilde{x} = \frac{d}{N} \sum_{n=1}^{N} q_n e_n$ satisfies

$$\begin{aligned} \|x - \tilde{x}\| &\leq \frac{d}{N} \| \sum_{n=1}^{N} (x_n - q_n) e_n \| \\ &\leq \frac{d}{N} \frac{\delta}{2} \sum_{n=1}^{N} \|e_n\| = \frac{d}{2} \delta \end{aligned}$$

Not good!

PCM

Replace $x_n \leftrightarrow q_n = \arg\min_{q \in \mathcal{A}_K^\delta} |x_n - q|$.

Then $\tilde{x} = \frac{d}{N} \sum_{n=1}^{N} q_n e_n$ satisfies

$$\begin{aligned} \|x - \tilde{x}\| &\leq \frac{d}{N} \| \sum_{n=1}^{N} (x_n - q_n) e_n \| \\ &\leq \frac{d}{N} \frac{\delta}{2} \sum_{n=1}^{N} \|e_n\| = \frac{d}{2} \delta \end{aligned}$$

Not good! **Bennett's "white noise assumption"**

Assume that $(\eta_n) = (x_n - q_n)$ is a sequence of independent, identically distributed random variables with mean 0 and variance $\frac{\delta^2}{12}$. Then the mean square error (MSE) satisfies

$$\text{MSE} = E\|x - \tilde{x}\|^2 \leq \frac{d}{12A} \delta^2 = \frac{(d\delta)^2}{12N}$$

Remarks

1. Bennett's "white noise assumption" is not rigorous, and not true in certain cases.

2. The MSE behaves like C/A. In the case of $\Sigma\Delta$ quantization of bandlimited functions, the MSE is $O(A^{-3})$ (Gray, Güntürk and Thao, Bin Han and Chen). PCM does not utilize redundancy efficiently.

3. The MSE only tells us about the average performance of a quantizer.

Remarks

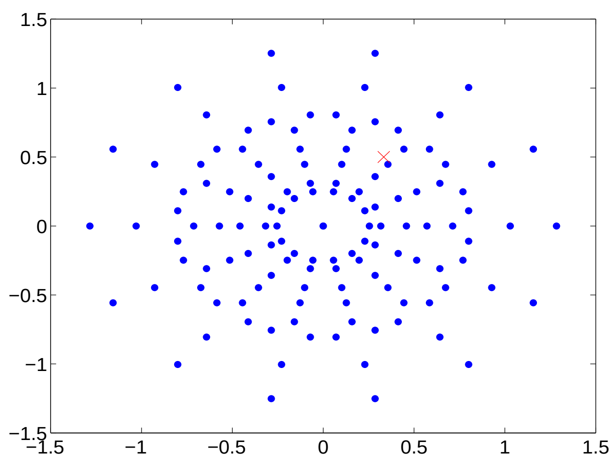
1. Bennett's "white noise assumption" is not rigorous, and not true in certain cases.

2. The MSE behaves like C/A. In the case of $\Sigma\Delta$ quantization of bandlimited functions, the MSE is $O(A^{-3})$ (Gray, Güntürk and Thao, Bin Han and Chen). PCM does not utilize redundancy efficiently.

3. The MSE only tells us about the average performance of a quantizer.

Example (continued):

Let $x = (\frac{1}{3}, \frac{1}{2})$, $E_7 = \{(\cos(\frac{2n\pi}{7}), \sin(\frac{2n\pi}{7}))\}_{n=1}^{7}$. Consider quantizers with $\mathcal{A} = \{-1, 1\}$.



Remarks

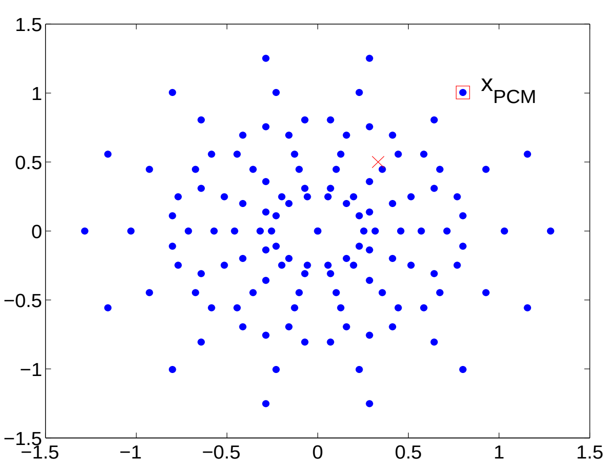
1. Bennett's "white noise assumption" is not rigorous, and not true in certain cases.

2. The MSE behaves like C/A. In the case of $\Sigma\Delta$ quantization of bandlimited functions, the MSE is $O(A^{-3})$ (Gray, Güntürk and Thao, Bin Han and Chen). PCM does not utilize redundancy efficiently.

3. The MSE only tells us about the average performance of a quantizer.

Example (continued):

Let $x = (\frac{1}{3}, \frac{1}{2})$, $E_7 = \{(\cos(\frac{2n\pi}{7}), \sin(\frac{2n\pi}{7}))\}_{n=1}^{7}$. Consider quantizers with $\mathcal{A} = \{-1, 1\}$.



Remarks

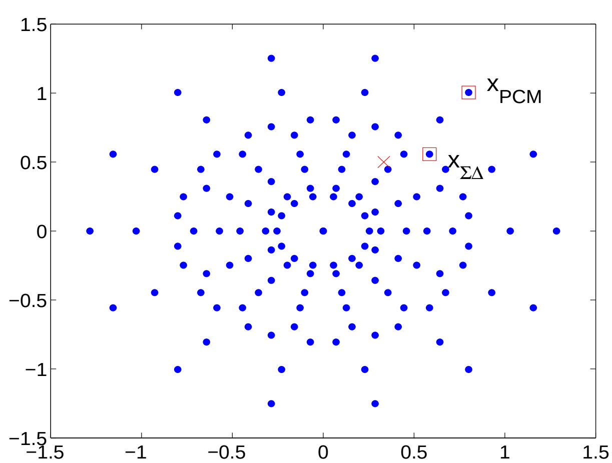
1. Bennett's "white noise assumption" is not rigorous, and not true in certain cases.

2. The MSE behaves like C/A. In the case of $\Sigma\Delta$ quantization of bandlimited functions, the MSE is $O(A^{-3})$ (Gray, Güntürk and Thao, Bin Han and Chen). PCM does not utilize redundancy efficiently.

3. The MSE only tells us about the average performance of a quantizer.

Example (continued):

Let $x = (\frac{1}{3}, \frac{1}{2})$, $E_7 = \{(\cos(\frac{2n\pi}{7}), \sin(\frac{2n\pi}{7}))\}_{n=1}^{7}$. Consider quantizers with $\mathcal{A} = \{-1, 1\}$.



$\Sigma\Delta$ quantizers for finite frames

Let $F = \{e_n\}_{n=1}^N$ be a frame for $\mathbb{R}^d$, $x \in \mathbb{R}^d$.

Define $x_n = \langle x, e_n \rangle$.

Fix the ordering p, a permutation of $\{1, 2, \ldots, N\}$.

Quantizer alphabet $\mathcal{A}_K^\delta$

Quantizer function $Q(u) = \underset{q \in \mathcal{A}_K^\delta}{\arg\min}\ |u - q|$

$\Sigma\Delta$ quantizers for finite frames

Let $F = \{e_n\}_{n=1}^{N}$ be a frame for $\mathbb{R}^d$, $x \in \mathbb{R}^d$.

Define $x_n = \langle x, e_n \rangle$.

Fix the ordering p, a permutation of $\{1, 2, \ldots, N\}$.

Quantizer alphabet $\mathcal{A}_K^\delta$

Quantizer function $Q(u) = \arg\min_{q \in \mathcal{A}_K^\delta} |u - q|$

Define the *first-order* $\Sigma\Delta$ *quantizer* with ordering p and with the quantizer alphabet $\mathcal{A}_K^\delta$ by means of the following recursion.

$$\begin{aligned} u_n - u_{n-1} &= x_{p(n)} - q_n \\ q_n &= Q(u_{n-1} + x_{p(n)}) \end{aligned}$$

where $u_0 = 0$ and $n = 1, 2, \ldots, N$.

Stability

The following stability result is used to prove error estimates.

Proposition If the frame coefficients $\{x_n\}_{n=1}^N$ satisfy

$$|x_n| \le (K - 1/2)\delta, \quad n = 1, \cdots, N,$$

then the state sequence $\{u_n\}_{n=0}^N$ generated by the first-order $\Sigma\Delta$ quantizer with alphabet $\mathcal{A}_K^\delta$ satisfies

$$|u_n| \le \frac{\delta}{2}, \quad n = 1, \cdots, N.$$

Stability

The following stability result is used to prove error estimates.

Proposition If the frame coefficients $\{x_n\}_{n=1}^N$ satisfy

$$|x_n| \leq (K - 1/2)\delta, \quad n = 1, \cdots, N,$$

then the state sequence $\{u_n\}_{n=0}^N$ generated by the first-order $\Sigma\Delta$ quantizer with alphabet $\mathcal{A}_K^\delta$ satisfies

$$|u_n| \leq \frac{\delta}{2}, \quad n = 1, \cdots, N.$$

- The first-order $\Sigma\Delta$ scheme is equivalent to

$$u_n = \sum_{j=1}^{n} x_{p(j)} - \sum_{j=1}^{n} q_j, \quad n = 1, \cdots, N.$$

- Stability results lead to tiling problems for higher order schemes.

Error estimate

Definition Let $F = \{e_n\}_{n=1}^N$ be a frame for $\mathbb{R}^d$, and let p be a permutation of $\{1, 2, \ldots, N\}$. We define the *variation* σ of F with respect to p by

$$\sigma(F, p) = \sum_{n=1}^{N-1} \|e_{p(n)} - e_{p(n+1)}\|.$$

Error estimate

Definition Let $F = \{e_n\}_{n=1}^N$ be a frame for $\mathbb{R}^d$, and let p be a permutation of $\{1, 2, \ldots, N\}$. We define the *variation* σ of F with respect to p by

$$\sigma(F, p) = \sum_{n=1}^{N-1} \|e_{p(n)} - e_{p(n+1)}\|.$$

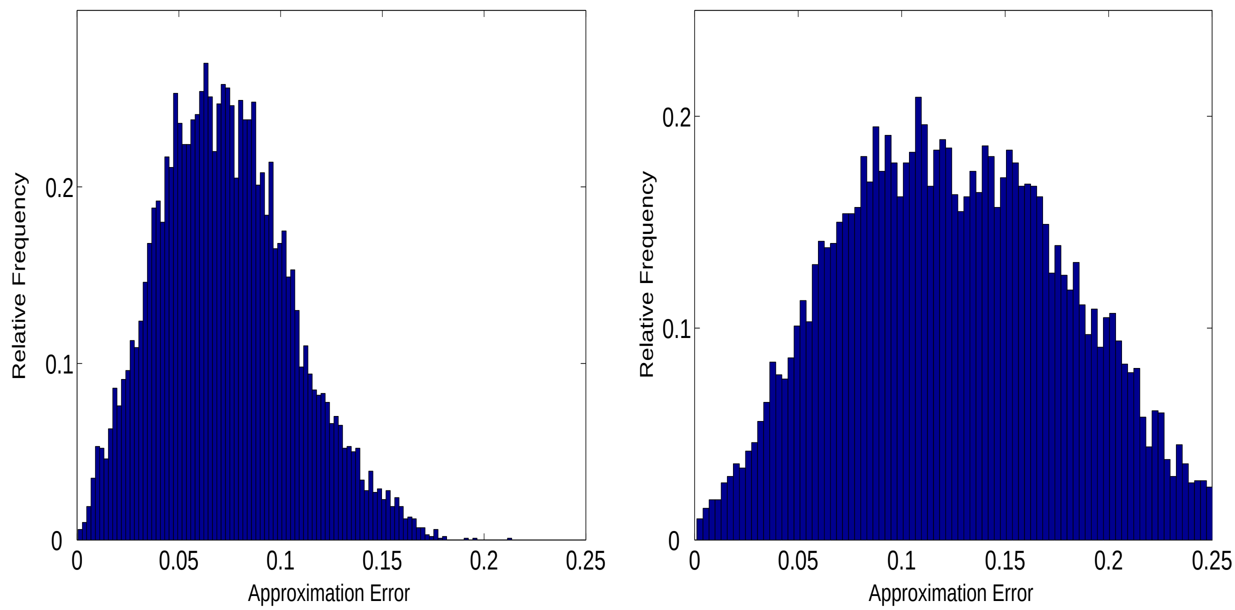
Theorem Let $F = \{e_n\}_{n=1}^N$ be an A-FNTF for $\mathbb{R}^d$. The approximation

$$\tilde{x} = \frac{d}{N} \sum_{n=1}^{N} q_n e_{p(n)}$$

generated by the first-order ΣΔ quantizer with ordering p and with the quantizer alphabet $\mathcal{A}_K^\delta$ satisfies

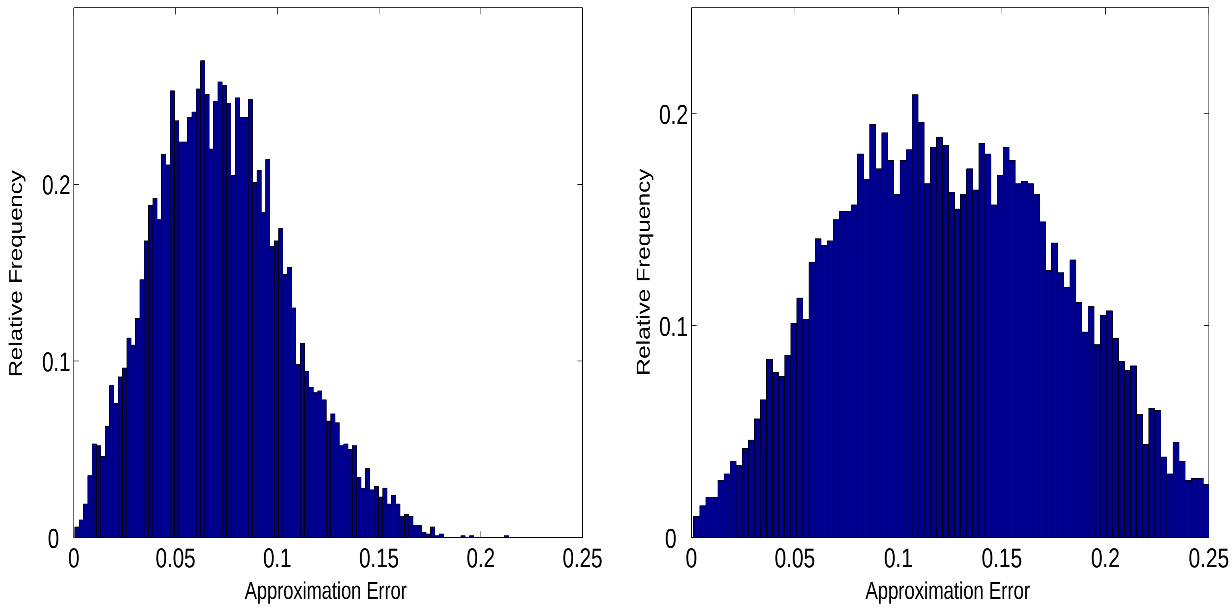
$$\|x - \tilde{x}\| \leq \frac{(\sigma(E, p) + 1)d}{N} \frac{\delta}{2}.$$

Order is important



Let E_7 be the FUN-TF for $\mathbb{R}^2$ given by the 7th roots of unity. Randomly select 10,000 points in the unit ball of $\mathbb{R}^2$. Quantize each point using the $\Sigma\Delta$ scheme with alphabet $\mathcal{A}_4^{1/4}$.

Order is important



Let E_7 be the FUN-TF for $\mathbb{R}^2$ given by the 7th roots of unity. Randomly select 10,000 points in the unit ball of $\mathbb{R}^2$. Quantize each point using the $\Sigma\Delta$ scheme with alphabet $\mathcal{A}_4^{1/4}$.

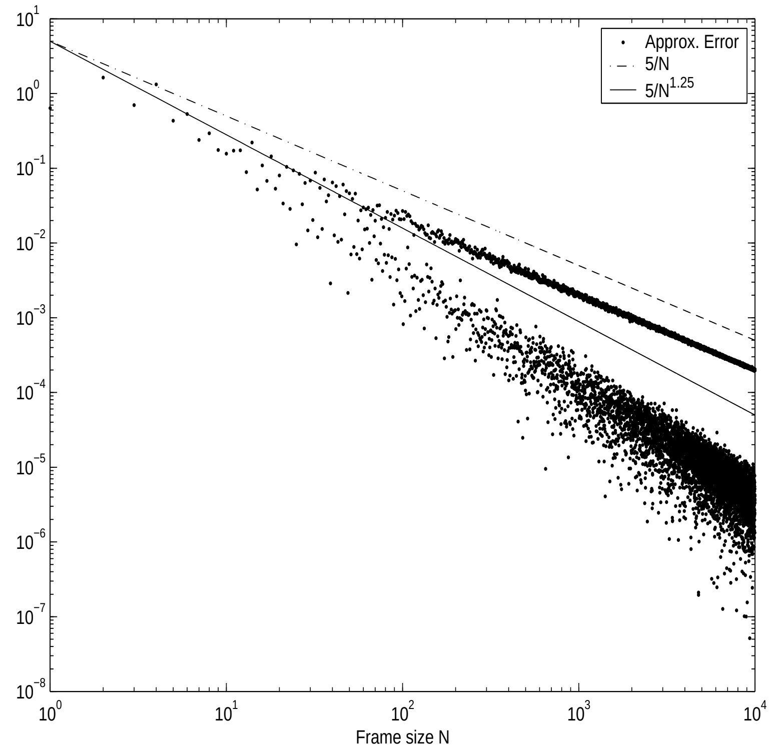
The figures show histograms for $||x - \tilde{x}||$ when the frame coefficients are quantized in their natural order (histogram on left)

$$x_1, x_2, x_3, x_4, x_5, x_6, x_7$$

and in the order (histogram on right) given by

$$x_1, x_4, x_7, x_3, x_6, x_2, x_5.$$

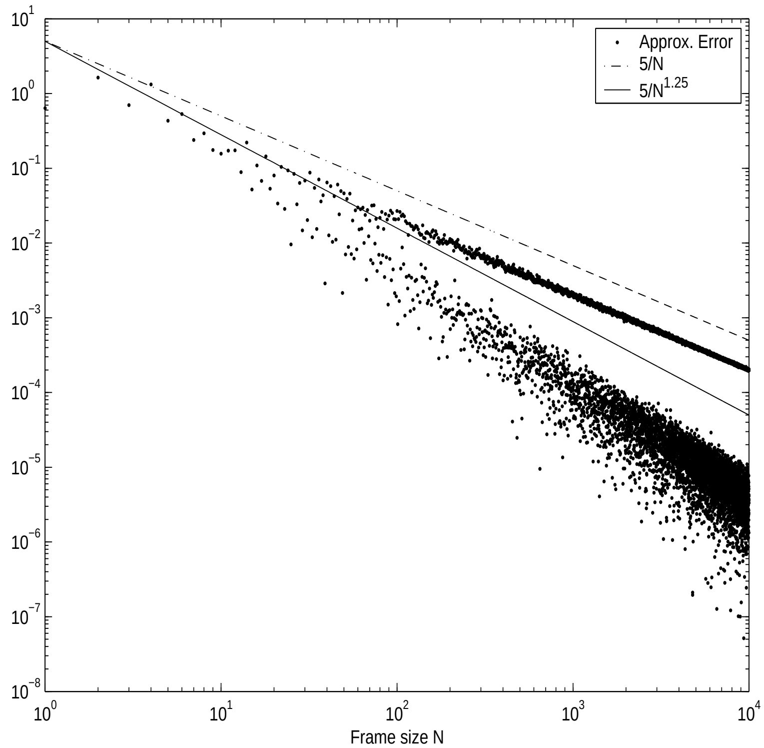
Even – odd



$E_N = \{e_n^N\}_{n=1}^N, e_n^N = (\cos(2\pi n/N), \sin(2\pi n/N))$.
Let $x = (\frac{1}{\pi}, \sqrt{\frac{3}{17}})$.

$$x = \frac{d}{N} \sum_{n=1}^{N} x_n^N e_n^N, \quad x_n^N = \langle x, e_n^N \rangle.$$

Even – odd



$E_N = \{e_n^N\}_{n=1}^N, e_n^N = (\cos(2\pi n/N), \sin(2\pi n/N))$.
Let $x = (\frac{1}{\pi}, \sqrt{\frac{3}{17}})$.

$$x = \frac{d}{N} \sum_{n=1}^{N} x_n^N e_n^N, \quad x_n^N = \langle x, e_n^N \rangle.$$

Let $\widetilde{x}_N$ be the approximation given by the 1st order $\Sigma\Delta$ quantizer using the alphabet $\{-1, 1\}$ and the natural ordering p.

The figure shows a log-log plot of $||x - \widetilde{x}_N||$.

Improved estimates

$E_N = \{e_n^N\}_{n=1}^N$, Nth roots of unity FUN-TFs for $\mathbb{R}^2$.

Let $x \in \mathbb{R}^d$, $||x|| \leq (K - 1/2)\delta$.

Quantize $$x = \frac{d}{N} \sum_{n=1}^{N} x_n^N e_n^N, \quad x_n^N = \langle x, e_n^N \rangle$$

using 1st order $\Sigma\Delta$ scheme with alphabet $\mathcal{A}_K^\delta$.

Improved estimates

$E_N = \{e_n^N\}_{n=1}^N$, Nth roots of unity FUN-TFs for $\mathbb{R}^2$.

Let $x \in \mathbb{R}^d$, $||x|| \leq (K - 1/2)\delta$.

Quantize $$x = \frac{d}{N} \sum_{n=1}^{N} x_n^N e_n^N, \quad x_n^N = \langle x, e_n^N \rangle$$

using 1st order $\Sigma\Delta$ scheme with alphabet $\mathcal{A}_K^\delta$.

Theorem If N is even and large then

$$||x - \tilde{x}|| \lesssim \frac{\delta \log N}{N^{5/4}}.$$

If N is odd and large then

$$\frac{\delta}{N} \lesssim ||x - \tilde{x}|| \leq \frac{(2\pi + 1)d}{N} \frac{\delta}{2}.$$

Improved estimates

$E_N = \{e_n^N\}_{n=1}^N$, Nth roots of unity FUN-TFs for $\mathbb{R}^2$.

Let $x \in \mathbb{R}^d$, $||x|| \leq (K - 1/2)\delta$.

Quantize $$x = \frac{d}{N} \sum_{n=1}^{N} x_n^N e_n^N, \quad x_n^N = \langle x, e_n^N \rangle$$

using 1st order $\Sigma\Delta$ scheme with alphabet $\mathcal{A}_K^\delta$.

Theorem If N is even and large then

$$||x - \tilde{x}|| \lesssim \frac{\delta \log N}{N^{5/4}}.$$

If N is odd and large then

$$\frac{\delta}{N} \lesssim ||x - \tilde{x}|| \leq \frac{(2\pi + 1)d}{N} \frac{\delta}{2}.$$

- The proof uses the analytic number theory approach developed by Sinan Güntürk.
- The theorem is true more generally, but additional technical assumptions are needed.

Harmonic frames

Zimmermann and Goyal, Kelner, Kovačević, Thao, Vetterli.

- $H = \mathbb{C}^d$. An *harmonic frame* $\{e_n\}_{n=1}^N$ for H is defined by the rows of the Bessel map L which is the complex N-DFT $N \times d$ matrix with $N - d$ columns removed.

Harmonic frames

Zimmermann and Goyal, Kelner, Kovačević, Thao, Vetterli.

- $H = \mathbb{C}^d$. An *harmonic frame* $\{e_n\}_{n=1}^N$ for H is defined by the rows of the Bessel map L which is the complex N-DFT $N \times d$ matrix with $N - d$ columns removed.

- $H = \mathbb{R}^d$, d even. The harmonic frame $\{e_n\}_{n=1}^N$ is defined by the Bessel map L which is the $N \times d$ matrix whose nth row is

$$e_n^N = \sqrt{\frac{2}{d}}\Big(\cos(\frac{2\pi n}{N}), \sin(\frac{2\pi n}{N}), \ldots, \cos(\frac{2\pi(d/2)n}{N}), \sin(\frac{2\pi(d/2)n}{N})\Big)$$

- Harmonic frames are FUN-TFs.

Harmonic frames

Zimmermann and Goyal, Kelner, Kovačević, Thao, Vetterli.

- $H = \mathbb{C}^d$. An *harmonic frame* $\{e_n\}_{n=1}^N$ for H is defined by the rows of the Bessel map L which is the complex N-DFT $N \times d$ matrix with $N - d$ columns removed.

- $H = \mathbb{R}^d$, d even. The harmonic frame $\{e_n\}_{n=1}^N$ is defined by the Bessel map L which is the $N \times d$ matrix whose nth row is

$$e_n^N = \sqrt{\frac{2}{d}}\Big(\cos(\frac{2\pi n}{N}), \sin(\frac{2\pi n}{N}), \ldots, \cos(\frac{2\pi(d/2)n}{N}), \sin(\frac{2\pi(d/2)n}{N})\Big)$$

- Harmonic frames are FUN-TFs.

- Let E_N be the harmonic frame for $\mathbb{R}^d$ and let p_N be the identity permutation. Then

$$\forall N, \ \sigma(E_N, p_N) \leq \pi d(d+1).$$

Error estimate for harmonic frames

Theorem Let E_N be the harmonic frame for $\mathbb{R}^d$ with frame bound N/d. Consider $x \in \mathbb{R}^d$, $\|x\| \leq 1$, and suppose the approximation $\tilde{x}$ of x is generated by a first-order $\Sigma\Delta$ quantizer as before. Then

$$\|x - \tilde{x}\| \leq \frac{d^2(d+1)+d}{N} \, \frac{\delta}{2}.$$

Error estimate for harmonic frames

Theorem Let E_N be the harmonic frame for $\mathbb{R}^d$ with frame bound N/d. Consider $x \in \mathbb{R}^d$, $\|x\| \leq 1$, and suppose the approximation $\tilde{x}$ of x is generated by a first-order $\Sigma\Delta$ quantizer as before. Then

$$\|x - \tilde{x}\| \leq \frac{d^2(d+1)+d}{N} \, \frac{\delta}{2}.$$

- Hence, for harmonic frames (and all those with bounded variation),

$$\mathrm{MSE}_{\Sigma\Delta} \leq \frac{C_d}{N^2} \, \delta^2.$$

Error estimate for harmonic frames

Theorem Let E_N be the harmonic frame for $\mathbb{R}^d$ with frame bound N/d. Consider $x \in \mathbb{R}^d$, $\|x\| \leq 1$, and suppose the approximation $\tilde{x}$ of x is generated by a first-order $\Sigma\Delta$ quantizer as before. Then

$$\|x - \tilde{x}\| \leq \frac{d^2(d+1) + d}{N} \frac{\delta}{2}.$$

- Hence, for harmonic frames (and all those with bounded variation),

$$\text{MSE}_{\Sigma\Delta} \leq \frac{C_d}{N^2} \delta^2.$$

- This bound is clearly superior asymptotically to

$$\text{MSE}_{\text{PCM}} = \frac{(d\delta)^2}{12N}$$

$\Sigma\Delta$ and "optimal" PCM

The digital encoding

$$\text{MSE}_{\text{PCM}} = \frac{(d\delta)^2}{12N}$$

in PCM format leaves open the possibility that decoding (reconstruction) could lead to

$$\text{“MSE}^{\text{opt}}_{\text{PCM}}\text{”} \ll O(\frac{1}{N}).$$

Goyal, Vetterli, Thao (1998) proved

$$\text{“MSE}^{\text{opt}}_{\text{PCM}}\text{”} \sim \frac{\tilde{C}_d}{N^2}\delta^2.$$

Theorem The first order $\Sigma\Delta$ scheme achieves the asymptotically optimal MSE_{PCM} for harmonic frames.

Thats all folks!