Quantization Methods
SIGMA-DELTA QUANTIZATION

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

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

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

First Order $\Sigma\Delta$
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** [continuous range $\mathbb{K}$ is not digital]

   $$\forall x \in X, \quad x = \sum x_n e_n, \quad x_n \in \mathbb{K}.$$ 

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

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

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

**Methods**

**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.

**Coarse quantization**, e.g., $\Sigma\Delta$. Use fewer bits to exploit redundancy.

**SRQP**

Norbert Wiener Center
Quantization

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

$$Q(u) = \arg \min \{|u - q| : q \in A^K_\delta\} = q_u$$
Replace \( x_n \leftrightarrow q_n = \arg\{\min |x_n - q| : q \in \mathcal{A}_K^\delta\} \). Then

\[
(PCM) \quad \tilde{x} = \frac{d}{N} \sum_{n=1}^{N} q_n e_n
\]
satisfies

\[
\|x - \tilde{x}\| \leq \frac{d}{N} \left\| \sum_{n=1}^{N} (x_n - q_n) e_n \right\| \leq \frac{d}{N} \frac{\delta}{2} \sum_{n=1}^{N} \|e_n\| = \frac{d}{2} \delta.
\]

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

\[
\text{MSE} = E\|x - \tilde{x}\|^2 \leq \frac{d}{12A} \delta^2 = \frac{(d\delta)^2}{12N}
\]
\[ \mathcal{A}_1^2 = \{-1, 1\} \text{ and } E_7 \]

Let \( x = \left( \frac{1}{3}, \frac{1}{2} \right) \), \( E_7 = \{(\cos(\frac{2n\pi}{7}), \sin(\frac{2n\pi}{7}))\}_{n=1}^7 \). Consider quantizers with \( \mathcal{A} = \{-1, 1\} \).
\( \mathcal{A}_1^2 = \{-1, 1\} \) and \( E_7 \)

Let \( x = \left( \frac{1}{3}, \frac{1}{2} \right) \), \( E_7 = \left\{ (\cos \left( \frac{2n\pi}{7} \right), \sin \left( \frac{2n\pi}{7} \right)) \right\}_{n=1}^7 \). Consider quantizers with \( \mathcal{A} = \{-1, 1\} \).
$A_1^2 = \{-1, 1\}$ and $E_7$

Let $x = \left(\frac{1}{3}, \frac{1}{2}\right)$, $E_7 = \{(\cos\left(\frac{2n\pi}{7}\right), \sin\left(\frac{2n\pi}{7}\right))\}_{n=1}^7$. Consider quantizers with $A = \{-1, 1\}$. 
$A_1^2 = \{-1, 1\}$ and $E_7$

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 $A = \{-1, 1\}$. 
Sigma-Delta quantization – background

- History from 1950s.
- Treatises of Candy, Temes (1992) and Norsworthy, Schreier, Temes (1997).
- PCM for finite frames and $\Sigma\Delta$ for $PW_\Omega$:
  Bölcsei, Daubechies, DeVore, Goyal, Gunturk, Kovačević, Thao, Vetterli.
- Combination of $\Sigma\Delta$ and finite frames:
  Powell, Yılmaz, and B.
- Subsequent work based on this $\Sigma\Delta$ finite frame theory:
  Bodman and Paulsen; Boufounos and Oppenheim; Jimenez and Yang Wang; Lammers, Powell, and Yılmaz.
- Genuinely apply it.
Let \( x \in \mathbb{C}^d, \{ e_n \}_{n=1}^N \) be a frame for \( \mathbb{C}^d \).

- **PCM:** \( \forall n = 1, \ldots, N, \quad q_n = Q_\delta(\langle x, e_n \rangle) \),

- **First Order Sigma-Delta:** Let \( \rho \) be a permutation of \( \{1, \ldots, N\} \). First Order Sigma-Delta quantization generates quantized sequence \( \{q_n\}_{n=1}^N \) by the iteration

\[
q_n = Q_\delta(u_{n-1} + \langle x, e_{\rho(n)} \rangle),
\]
\[
u_n = u_{n-1} + \langle x, e_{\rho(n)} \rangle - q_n,
\]

for \( n = 1, \ldots, N \), with an initial condition \( u_0 \).

In either case, the quantized estimate is

\[
\tilde{x} = \frac{d}{N} \sum_{n=1}^N q_n e_n = \frac{d}{N}L^* q
\]
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 $A_K^\delta$
Quantizer function $Q(u) = \text{arg}\{\min |u - q| : q \in A_K^\delta\}$

Define the first-order $\Sigma \Delta$ quantizer with ordering $p$ and with the quantizer alphabet $A_K^\delta$ by means of the following recursion.

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

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

\[
|x_n| \leq (K - 1/2)\delta, \quad n = 1, \ldots, N,
\]

then the state sequence \( \{u_n\}_{n=0}^{N} \) generated by the first-order \( \Sigma\Delta \) quantizer with alphabet \( A^\delta_K \) satisfies \( |u_n| \leq \delta/2, n = 1, \ldots, N \).

- The first-order \( \Sigma\Delta \) scheme is equivalent to
  \[
u_n = \sum_{j=1}^{n} x_{p(j)} - \sum_{j=1}^{n} q_j, \quad n = 1, \ldots, N.
\]

- Stability results lead to tiling problems for higher order schemes.
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\}. The variation $\sigma(F, p)$ is

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

Error estimate

**Theorem**

Let $F = \{e_n\}_{n=1}^N$ be an A-FUNT 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 $\Sigma\Delta$ quantizer with ordering $p$ and with the quantizer alphabet $A_K^\delta$ satisfies

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

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

**Definition**

\( H = \mathbb{C}^d \). 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 \( n \)th row is

\[
e_n^N = \sqrt{\frac{2}{d}} \left( \cos(\frac{2\pi n}{N}), \sin(\frac{2\pi n}{N}), \ldots, \cos(\frac{2\pi (d/2)n}{N}), \sin(\frac{2\pi (d/2)n}{N}) \right).
\]

- Harmonic frames are FUNTFs.
- 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

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

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

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

The first order $\Sigma \Delta$ scheme achieves the asymptotically optimal $\text{MSE}_{\text{PCM}}$ for harmonic frames.

The digital encoding

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

in PCM format leaves open the possibility that decoding (consistent nonlinear reconstruction, with additional numerical complexity this entails) could lead to

$$\text{"MSE}^\text{opt}_{\text{PCM}} \ll O\left(\frac{1}{N}\right).$$

Goyal, Vetterli, Thao (1998) proved

$$\text{"MSE}^\text{opt}_{\text{PCM}} \sim \frac{\tilde{C}_d}{N^2} \delta^2.$$
A comparison of Σ-Δ and PCM
Comparison of 1-bit PCM and 1-bit $\Sigma\Delta$

Let $x \in \mathbb{C}^d$, $\|x\| \leq 1$.

**Definition**

- $q_{PCM}(x)$ is the sequence to which $x$ is mapped by PCM.
- $q_{\Sigma\Delta}(x)$ is the sequence to which $x$ is mapped by $\Sigma\Delta$.

$$\text{err}_{PCM}(x) = \|x - \frac{d}{N}L^* q_{PCM}(x)\|$$

$$\text{err}_{\Sigma\Delta}(x) = \|x - \frac{d}{N}L^* q_{\Sigma\Delta}(x)\|$$

Fickus question: We shall analyze to what extent $\text{err}_{\Sigma\Delta}(x) < \text{err}_{PCM}(x)$ beyond our results with Powell and Yilmaz.
Let $x \in \mathbb{C}^d$.
Let $F = \{e_n\}_{n=1}^N$ be a FUNTF for $\mathbb{C}^d$ with the analysis matrix $L$.

**Definition**

- $q_{PCM}(x, F, b)$ is the quantized sequence given by $b$-bit PCM,
- $q_{\Sigma\Delta}(x, F, b)$ is the quantized sequence given by $b$-bit Sigma-Delta.

$$
err_{PCM}(x, F, b) = \|x - \frac{d}{N} L^* q_{PCM}(x)\|,
$$
$$
err_{\Sigma\Delta}(x, F, b) = \|x - \frac{d}{N} L^* q_{\Sigma\Delta}(x)\|.
$$
Comparison of 1-bit PCM and 1-bit $\Sigma\Delta$

Definition

A function $e : [a, b] \to \mathbb{C}^d$ is of bounded variation (BV) if there is a $K > 0$ such that for every $a \leq t_1 < t_2 < \cdots < t_N \leq b$, 

$$
\sum_{n=1}^{N-1} \| e(t_n) - e(t_{n+1}) \| \leq K.
$$

The smallest such $K$ is denoted by $|e|_{BV}$, and defines a seminorm for the space of BV functions.
Theorem 1

Let $x \in \mathbb{C}^d$ satisfy $0 < \|x\| \leq 1$, and let $F = \{e_n\}_{n=1}^N$ be a FUNTF for $\mathbb{C}^d$. Then, the 1-bit PCM error satisfies

$$\text{err}_{PCM}(x, F, 1) \geq \alpha_F + 1 - \|x\|$$

where

$$\alpha_F := \inf_{\|x\|=1} \frac{d}{N} \sum_{n=1}^N (|\text{Re}(\langle x, e_n \rangle)| + |\text{Im}(\langle x, e_n \rangle)|) - 1 \geq 0.$$
Comparison of 1-bit PCM and 1-bit Sigma-Delta

Theorem 2

Let \( \{ F_N = \{ e_n^N \}_{n=1}^N \} \) be a family of FUNTFs for \( \mathbb{C}^d \). Then,

\[
\forall \varepsilon > 0, \exists N_0 > 0, \text{ such that } \forall N \geq N_0 \text{ and } \forall 0 < \| x \| \leq 1 - \varepsilon
\]

\[
er_{\Sigma\Delta}(x, F_N, 1) \leq er_{\text{PCM}}(x, F_N, 1).
\]

Numerical experiments suggest that, we can choose \( N \) significantly smaller than \( (M/\varepsilon)^{2d} \).
Comparison of 1-bit PCM and 1-bit Sigma-Delta

If \( \{\alpha_{F_N}\} \) is bounded below by a positive number, then we can improve Theorem 2:

**Theorem 3**

Let \( \{F_N = \{e^N_n\}_{n=1}^N\} \) be a family of FUNTFs for \( \mathbb{C}^d \) such that

\[
\exists a > 0, \forall N, \alpha_{F_N} \geq a.
\]

Then,

\[
\exists N_0 > 0 \ such \ that \ \forall N \geq N_0 \ and \ \forall 0 < \|x\| \leq 1 \\
err_{\Sigma\Delta}(x, F_N, 1) \leq err_{PCM}(x, F_N, 1).
\]
Comparison of 1-bit PCM and 1-bit Sigma-Delta

Below is a family \( \{ F_N \} \) of FUNTFs where \( \{ \alpha_{F_N} \} \) is bounded below by a positive constant. Harmonic frames are examples to such families.

**Theorem 4**

Let \( e : [0, 1] \to \{ x \in \mathbb{C}^d : \| x \| = 1 \} \) be continuous function of bounded variation such that \( F_N = \{ e(n/N) \}_{n=1}^N \) is a FUNTF for \( \mathbb{C}^d \) for every \( N \). Then,

\[
\exists N_0 > 0 \text{ such that } \forall N \geq N_0 \text{ and } \forall 0 < \| x \| \leq 1
\]

\[
\text{err}_{\Sigma \Delta}(x, F_N, 1) \leq \text{err}_{PCM}(x, F_N, 1).
\]

One can show that \( \alpha := \lim_{N \to \infty} \alpha_{F_N} \) is positive, and that

\[
\alpha + 1 = d \inf_{\| x \| = 1} \int_0^1 (| \text{Re}(\langle x, e(t) \rangle) | + | \text{Im}(\langle x, e(t) \rangle) |) \, dt.
\]
Comparison of 1-bit PCM and 1-bit $\Sigma\Delta$

**Theorem**

Let $e : [0, 1] \to \{ x \in \mathbb{C}^d : \|x\| = 1 \}$ be continuous function of bounded variation such that $F_N = (e(n/N))_{n=1}^N$ is a FUNTF for $\mathbb{C}^d$ for every $N$. Then,

$$\exists N_0 > 0 \text{ such that } \forall N \geq N_0 \text{ and } \forall 0 < \|x\| \leq 1$$

$$\text{err}_{\Sigma\Delta}(x) \leq \text{err}_{PCM}(x).$$

Moreover, a lower bound for $N_0$ is $d(1 + |e|_{BV})/(\sqrt{d} - 1)$. 
Comparison of 1-bit PCM and 1-bit $\Sigma\Delta$

Example (Roots of unity frames for $\mathbb{R}^2$)

$$e_n^N = (\cos(2\pi n/N), \sin(2\pi n/N)).$$

Here, $e(t) = (\cos(2\pi t), \sin(2\pi t))$, 
$M = |e|_{BV} = 2\pi, \lim_{N \to \infty} \alpha_{F_N} = 2/\pi$.

Example (Real Harmonic Frames for $\mathbb{R}^{2k}$)

$$e_n^N = \frac{1}{\sqrt{k}}(\cos(2\pi n/N), \sin(2\pi n/N), \ldots, \cos(2\pi kn/N), \sin(2\pi kn/N)).$$

In this case, $e(t) = \frac{1}{\sqrt{k}}(\cos(2\pi t), \sin(2\pi t), \ldots, \cos(2\pi kt), \sin(2\pi kt))$, 
$M = |e|_{BV} = 2\pi \sqrt{\frac{1}{d} \sum_{k=1}^{d} k^2}$. 
Comparison of 2-bit PCM and 1-bit $\Sigma\Delta$

Red: $\text{err}_{\text{PCM}}(x) < \text{err}_{\Sigma\Delta}(x)$, Green: $\text{err}_{\text{PCM}}(x) = \text{err}_{\Sigma\Delta}(x)$
Comparison of 2-bit PCM and 1-bit $\Sigma\Delta$

81st Roots of 1 frame, 2bit PCM vs 1bit $\Sigma\Delta$

Red: $\text{err}_{PCM}(x) < \text{err}_{\Sigma\Delta}(x)$, Green: $\text{err}_{PCM}(x) = \text{err}_{\Sigma\Delta}(x)$
Comparison of 2-bit PCM and 1-bit $\Sigma\Delta$

101st Roots of 1 frame, 2bit PCM vs 1bit $\Sigma\Delta$

Red: $err_{PCM}(x) < err_{\Sigma\Delta}(x)$, Green: $err_{PCM}(x) = err_{\Sigma\Delta}(x)$
Comparison of 3-bit PCM and 1-bit $\Sigma\Delta$

Red: $\text{err}_{PCM}(x) < \text{err}_{\Sigma\Delta}(x)$, Green: $\text{err}_{PCM}(x) = \text{err}_{\Sigma\Delta}(x)$
Comparison of 3-bit PCM and 1-bit $\Sigma\Delta$

101st Roots of 1 frame, 3bit PCM vs 1bit $\Sigma\Delta$

Red: $\text{err}_{PCM}(x) < \text{err}_{\Sigma\Delta}(x)$, Green: $\text{err}_{PCM}(x) = \text{err}_{\Sigma\Delta}(x)$
Comparison of 3-bit PCM and 1-bit $\Sigma\Delta$

201st Roots of 1 frame, 3bit PCM vs 1bit $\Sigma\Delta$

Red: $\text{err}_{PCM}(x) < \text{err}_{\Sigma\Delta}(x)$, Green: $\text{err}_{PCM}(x) = \text{err}_{\Sigma\Delta}(x)$
Comparison of 3-bit PCM and 2-bit $\Sigma\Delta$

81st Roots of 1 frame, 3bit PCM vs 2bit $\Sigma\Delta$

Red: $\text{err}_{PCM}(x) < \text{err}_{\Sigma\Delta}(x)$, Green: $\text{err}_{PCM}(x) = \text{err}_{\Sigma\Delta}(x)$
Comparison of 3-bit PCM and 2-bit \( \Sigma \Delta \)

Red: \( \text{err}_{\text{PCM}}(x) < \text{err}_{\Sigma \Delta}(x) \), Green: \( \text{err}_{\text{PCM}}(x) = \text{err}_{\Sigma \Delta}(x) \)
Comparison of 3-bit PCM and 2-bit $\Sigma\Delta$

Red: $\text{err}_{PCM}(x) < \text{err}_{\Sigma\Delta}(x)$, Green: $\text{err}_{PCM}(x) = \text{err}_{\Sigma\Delta}(x)$
Let $K \in \mathbb{N}$ and $\delta > 0$. The *midrise* quantization alphabet is

$$A^K_\delta = \left\{ \left( m + \frac{1}{2} \right) \delta + in\delta : m = -K, \ldots, K - 1, \ n = -K, \ldots, K \right\}$$

**Figure:** $A^K_\delta$ for $K = 3\delta$. 

Norbert Wiener Center
for Harmonic Analysis and Applications
For $K > 0$ (we consider only $K = 1$) and $b \geq 1$, an integer representing the number of bits, let $\delta = 2K/(2^b - 1)$.

$$A^K_\delta = \{(-K + m\delta) + i(-K + n\delta) : m, n = 0, \ldots, 2^b - 1\}.$$ 

The associated \textit{scalar uniform quantizer} is

$$Q_\delta(u + iv) = \delta \left( \frac{1}{2} + \left\lfloor \frac{u}{\delta} \right\rfloor + i \left( \frac{1}{2} + \left\lfloor \frac{v}{\delta} \right\rfloor \right) \right).$$

In particular, for 1-bit case, $Q(u + iv) = \text{sign}(u) + i\text{sign}(v)$
Complex $\Sigma\Delta$

The **scalar uniform quantizer** associated to $A^\delta_K$ is

$$Q_\delta(a + ib) = \delta \left( \frac{1}{2} + \left\lfloor \frac{a}{\delta} \right\rfloor + i \left\lfloor \frac{b}{\delta} \right\rfloor \right),$$

where $\lfloor x \rfloor$ is the largest integer smaller than $x$.
For any $z = a + ib$ with $|a| \leq K$ and $|b| \leq K$, $Q$ satisfies

$$|z - Q_\delta(z)| \leq \min_{\zeta \in A^\delta_K} |z - \zeta|.$$

Let $\{x_n\}_{n=1}^N \subseteq \mathbb{C}$ and let $p$ be a permutation of $\{1, \ldots, N\}$. Analogous to the real case, the first order $\Sigma\Delta$ quantization is defined by the iteration

$$u_n = u_{n-1} + x_{p(n)} - q_n,$$
$$q_n = Q_\delta(u_{n-1} + x_{p(n)}).$$
The following theorem is analogous to BPY

**Theorem**

Let \( F = \{e_n\}_{n=1}^{N} \) be a finite unit norm frame for \( \mathbb{C}^d \), let \( p \) be a permutation of \( \{1, \ldots, N\} \), let \( |u_0| \leq \delta/2 \), and let \( x \in \mathbb{C}^d \) satisfy \( \|x\| \leq (K - 1/2)\delta \). The \( \Sigma\Delta \) approximation error \( \|x - \tilde{x}\| \) satisfies

\[
\|x - \tilde{x}\| \leq \sqrt{2}\|S^{-1}\|_{op} \left( \sigma(F, p)\frac{\delta}{2} + |u_N| + |u_0| \right),
\]

where \( S^{-1} \) is the inverse frame operator. In particular, if \( F \) is a FUNTF, then

\[
\|x - \tilde{x}\| \leq \sqrt{2}d\frac{d}{N} \left( \sigma(F, p)\frac{\delta}{2} + |u_N| + |u_0| \right),
\]
Let \( \{F_N\} \) be a family of FUNTFs, and \( \rho_N \) be a permutation of \( \{1, \ldots, N\} \). Then the frame variation \( \sigma(F_N, \rho_N) \) is a function of \( N \). If \( \sigma(F_N, \rho_N) \) is bounded, then

\[
\|x - \tilde{x}\| = \mathcal{O}(N^{-1}) \text{ as } N \to \infty.
\]

Wang gives an upper bound for the frame variation of frames for \( \mathbb{R}^d \), using the results from the Travelling Salesman Problem.

**Theorem YW**

Let \( S = \{v_j\}_{j=1}^N \subseteq [-\frac{1}{2}, \frac{1}{2}]^d \) with \( d \geq 3 \). There exists a permutation \( \rho \) of \( \{1, \ldots, N\} \) such that

\[
\sum_{j=1}^{N-1} \|v_{\rho(j)} - v_{\rho(j+1)}\| \leq 2\sqrt{d + 3N^{1 - \frac{1}{d}}} - 2\sqrt{d} + 3.
\]
Theorem

Let \( F = \{ e_n \}_{n=1}^{N} \) be a FUNTF for \( \mathbb{R}^d \), \( |u_0| \leq \delta / 2 \), and let \( x \in \mathbb{R}^d \) satisfy \( \|x\| \leq (K - 1/2)\delta \). Then, there exists a permutation \( p \) of \( \{1, 2, \ldots, N\} \) such that the approximation error \( \|x - \tilde{x}\| \) satisfies

\[
\|x - \tilde{x}\| \leq \sqrt{2} \delta d \left( (1 - \sqrt{d + 3}) N^{-1} + \sqrt{d + 3} N^{-\frac{1}{d}} \right)
\]

This theorem guarantees that

\[
\|x - \tilde{x}\| \leq O(N^{-\frac{1}{d}}) \text{ as } N \to \infty
\]

for FUNTFs for \( \mathbb{R}^d \).
Preprocessing for clutter mitigation

- Massive sensor data set $\rightarrow$ dimension reduction $\rightarrow$ sparse representation
- False targets caused by clutter inhibit data triage, waste vital resources, and degrade sparse representation algorithms
- View clutter mitigation as preprocessing step for ATR/ATE
- For active sensors, choose waveform to reduce clutter effects by limiting side lobe magnitude
  - improves concise data representation
  - supports dimensionality reduction processing
Sparse coefficient sets for stable representation

• Opportunistic sensing systems can utilize large networks of diverse sensors
  – sensor quality may vary, e.g., low cost wireless sensors
  – massive amount of noisy sensor data
• Signal representations using sparse coefficient sets
  – compensate for hardware errors
  – ensure numerical stability
  – frame setting → frame dimension reduction
Frame variation and $\Sigma\Delta$

- $F = \{e_j\}_{j=1}^N$ a FUNTF for $\mathbb{C}^d$
- $x \in \mathbb{C}^d$, $p$ a permutation of $\{1, \ldots, N\}$, $x_p(n) = \langle x, e_p(n) \rangle$,
  \[ x = \frac{d}{N} \sum_{n=1}^N x_p(n) e_p(n) \quad \text{and} \quad \tilde{x} \equiv \frac{d}{N} \sum_{n=1}^N q_n e_p(n) \]

- Frame variation,
  \[ \sigma(F, p) = \sum_{n=1}^{N-1} \| e_p(n) - e_p(n+1) \| \]

- Transport $\Sigma\Delta$ FUNTF setting to coefficient sparse representation point of view.
Summary

Given a signal $x$ and a tolerance $r > 0$

- Define frames using Frame Potential Energy and SQP (or other optimization)
- Analyze Frame Variation in terms of our permutation algorithm
- Compute $\tilde{x}$ having separated coefficients taken from a fixed small and sparse alphabet
- Ensure that $\|x - \tilde{x}\| < r$.

**Conclusion:** $\tilde{x}$ is a stable sparse coefficient approximant of $x$
Dimension reduction
Given data space $X$ of $N$ vectors in $\mathbb{R}^D$. ($N$ is the number of pixels in the hypercube, $D$ is the number of spectral bands.)

Two Steps:

1. Construction of an $N \times N$ symmetric, positive semi–definite kernel, $K$, from these $N$ data points in $\mathbb{R}^D$.

2. Diagonalization of $K$, and then choosing $d \leq D$ significant orthogonal eigenmaps of $K$. 
Different classes of interest may not be orthogonal to each other; however, they may be captured by different frame elements. It is plausible that classes may correspond to elements in a frame but not elements in a basis.

A *frame* generalizes the concept of an orthonormal basis. Frame elements are non–orthogonal.
Dimension reduction paradigm

- Given data space $X$ of $N$ vectors $x_m \in \mathbb{R}^D$, and let

$$K : X \times X \rightarrow \mathbb{R}$$

be a symmetric ($K(x, y) = K(y, x)$), positive semi–definite kernel.

- We map $X$ to a low dimensional space via the following mapping:

$$X \longrightarrow K \longrightarrow \mathbb{R}^d(K), \quad d < D$$

$$x_m \mapsto y_m = (y[m, n_1], y[m, n_2], \ldots, y[m, n_d]) \in \mathbb{R}^d(K),$$

where $y[\cdot, n] \in \mathbb{R}^N$ is an eigenvector of $K$. 

Norbert Wiener Center
for Harmonic Analysis and Applications

John J. Benedetto

Frame potential classification algorithm for retinal data
Consider the data points $X$ as the nodes of a graph.

Define a metric $\rho : X \times X \rightarrow \mathbb{R}^+$, e.g., $\rho(x_m, x_n) = \|x_m - x_n\|$ is the Euclidean distance.

Choose $q \in \mathbb{N}$.

For each $x_i$ choose the $q$ nodes $x_n$ closest to $x_i$ in the metric $\rho$, and place an edge between $x_i$ and each of these nodes.

This defines $N'(x_i)$, viz.,

$N'(x_i) = \{x \in X : \exists \text{ an edge between } x \text{ and } x_i \}.$

To define the weights on the edges, we compute:

$$W = \arg\min_{\tilde{W}} \left| x_i - \sum_{j \in N'(x_i)} \tilde{W}(x_i, x_j)x_j \right|^2.$$

Set $K = (I - W)(I - W^T)$ and diagonalize $K$.

$K$ is symmetric and positive semi–definite.
Consider the data points $X$ as the nodes of a graph.

Define a metric $\rho : X \times X \rightarrow \mathbb{R}^+$, e.g., $\rho(x_m, x_n) = \|x_m - x_n\|$ is the Euclidean distance.

Choose $q \in \mathbb{N}$.

For each $x_i$ choose the $q$ nodes $x_n$ closest to $x_i$ in the metric $\rho$, and place an edge between $x_i$ and each of these nodes.

This defines $N'(x_i)$, viz.,

$$N'(x_i) = \{ x \in X : \exists \text{ an edge between } x \text{ and } x_i \}. $$

To define the weights on the edges, we compute:

$$W_{ij} = \begin{cases} 
\exp(-\|x_i - x_j\|^2/\sigma) & \text{if } x_j \in N'(x_i) \text{ or } x_i \in N'(x_j) \\
0 & \text{otherwise}
\end{cases}$$

Set $K = D - W$, where $D_{ii} = \sum_j W_{ij}$ and $D_{ij} = 0$ for $i \neq j$;

Diagonalize $K$.

$K$ is symmetric and positive semi–definite.
Frames
Frame potential classification algorithm
Optimization problem: maximal separation

Goal: Construct a FUNTF $\{\psi_k\}_{k=1}^s$ such that each $\psi_k$ is associated to only one classifiable material.

For $\{\theta_k\}_{k=1}^s \in S^{d-1} \times \ldots \times S^{d-1}$ and $n = 1, \ldots, s$, set

$$p(\theta_n) = \sum_{m=1}^N |\langle y_m, \theta_n \rangle|$$

and consider the maximal separation

$$\sup \min\{|p(\theta_k) - p(\theta_n)| : k \neq n\}.$$
Combine maximal separation with frame potential to construct a pseudo-FUNTF $\Psi = \{\psi_k\}_{k=1}^S$ by solving the minimization problem:

$$\sup \left\{ \min \{ |p(\theta_k) - p(\theta_n)| : k \neq n \} : \{\theta_j\} \in \{ \arg \min_\Phi TFP(\Phi) \} \right\}, \quad (1)$$

where $\Phi = \{\phi_k\}_{k=1}^S$.

- (1) is solved using a new, fast gradient descent method for products of spheres.
- Nate Strawn created the method and developed new geometric ideas for such computation.
Combine frame potential with \("l^1\)-energy" to construct a FUNTF \(\Psi = \{\psi_k\}_{k=1}^s\) by solving a minimization problem of the following type:

\[
\min \{ TFP(\Theta) + P(Y, \Theta) : \Theta \in S^{d-1} \times \cdots \times S^{d-1} \},
\]

where

\[
P(Y, \Theta) = \sum_{n=1}^N \sum_{k=1}^s |\langle y_n, \theta_k \rangle| = \sum_{k=1}^s p(\theta_k).
\]

Remark. a. Minimization of \(P\) is convex optimization of \(l^1\)-energy of \(Y\) for a given frame.
b. By Candes and Tao (2005), under suitable conditions, this can yield a frame \(\Psi\) with a sparse set of coefficients \(\{\langle y_n, \psi_k \rangle\}\). We do not proceed this way to obtain sparsity.
Given $\psi = \{\psi_n\}_{n=1}^{s}$ and $m \in \{1, \ldots, N\}$. Consider the set of frame decompositions

$$y_m = \sum_{n=1}^{s} c_{m,n}^{\alpha} \psi_n, \quad \text{indexed by } \alpha \in \mathbb{R}.$$

- If $\psi$ is a FUNTF then $\alpha = 0$ designates the canonical dual, i.e.,

$$c_{m,n}^{0} = \frac{d}{s} \langle y_m, \psi_n \rangle.$$
For each $m \in \{1, \ldots, N\}$ choose an $\ell^1$ sparse decomposition

$$y_m = \sum_{n=1}^{s} c_{m,n}^{\alpha(m)} \Psi_n$$

defined by the inequality,

$$\forall \alpha, \quad \sum_{n=1}^{s} |c_{m,n}^{\alpha(m)}| \leq \sum_{n=1}^{s} |c_{m,n}^{\alpha}|.$$ 

There is $\ell^0$ theory.
Choose \( n \in \{1, \ldots, s\} \). Take a slice, \( P_n \), of the data cube at \( n \). \( P_n \) contains \( N \) points \( m \).

The image with \( N \) pixels \( m \), associated to the frame element \( \Psi_n \), is defined by \( \{c_{m,n}^{\alpha(m)} \mid m = 1, \ldots, N\} \).
Hyperspectral image processing
Urban data set classes

Figure: HYDICE Copperas Cove, TX — http://www.tec.army.mil/Hypercube/
NGA gave us 23 classes, associated with the different colors in the previous figure.

In fact, if the 23 classes were to correspond roughly to orthogonal subspaces, then one cannot achieve effective dimension reduction less than dimension $d = 23$.

However, we could have a frame with 23 elements in a space of reduced dimension $d < 23$. 
Spectral signatures of selected classes
Frame coefficients

(a) Original

(b) Road coefficients

(c) Tree coefficients

(d) White house coefficients
Frame coefficients

(a) Original
(b) Road coefficients
(c) Tree coefficients
(d) Dirt/grass coefficients
Overview of Classification Results