Approximation by Sampling:
- Running time too slow?
- Maybe your data size is too large!
- Idea:
  - Extract a small subset, $P' \subseteq P$
  - Run solve problem exactly on $P'$
  - Prove that the answer on $P'$ is close to optimal on $P$.

How to compute $P'$?
- Depends on your problem
- Random sampling is most common, but not necessarily best

Example: Minimum Enclosing Ball (MEB)
- Given a set $P = \{p_1, \ldots, p_n\}$ in $\mathbb{R}^d$
  compute the Euclidean ball of min. radius enclosing $P$. 
Problem with random sampling:
- \( \text{MEB}(P) \) depends on points near periphery
- Random sample extracts many irrelevant points.
- Smarter: Use a sampling method that gives priority to peripheral points.

**Coreset:** Let \( P \) be the input set. \( f^*(P) \rightarrow \mathbb{R} \) is our objective function (e.g., \( f^*(P) = \) radius of MEB).

Given \( \varepsilon > 0 \), an \( \varepsilon \)-coreset is a subset \( Q \subseteq P \) s.t.

\[
1 - 3\varepsilon \leq \frac{f^*(Q)}{f^*(P)} \leq 1 + 3\varepsilon
\]

The opt. soln. for \( Q \) is close to opt. for \( P \).
Questions:
- For what optimization problems do (small) coresets exist?
- (As a function of $n + \varepsilon$) how small is the coreset?
- How fast can we compute a coreset?

Coreset for Directional Width:
(also called $\varepsilon$-kernel)

- Given a point set $P \subseteq \mathbb{R}^d$
- Given a unit vector $\hat{u}$

- Directional width of $P$ in direction $\hat{u}$ is:

$$W_u(P) = \max_{p \in P} (\bar{p} \cdot \hat{u}) - \min_{p \in P} (\bar{p} \cdot \hat{u})$$
Given $\varepsilon > 0$, an $\varepsilon$-coreset for direc. width (also called $\varepsilon$-kernel) is a subset $R \subseteq P$ s.t.

$$\forall \text{ unit vect. } \hat{u}:$$

$$(1 - \varepsilon) \, W_u(P) \leq W_u(R) \leq W_u(P)$$

Trivially true since $R \subseteq P$

Getting this is the objective

Aside: When computing approx. lower bounds we sometimes write:

$$(1 - \varepsilon) \cdot \text{exact} \leq \text{approx}$$

and other times:

$$\frac{\text{exact}}{1 + \varepsilon} \leq \text{approx}$$

Does the form matter?

Not really. If $0 < \varepsilon < 1$, then

$$1 - \varepsilon \leq \frac{1}{1 + \varepsilon} \leq 1 - \frac{\varepsilon}{2}$$

- Only constant factors are affected
Useful Facts:

**Chain Property**: If $X$ is an $\varepsilon$-kernel for $Y$ and $Y$ is an $\varepsilon'$-kernel for $Z$ then $X$ is an $(\varepsilon + \varepsilon')$-kernel for $Z$

**Union Property**: If $X$ is an $\varepsilon$-kernel for $P$, $X'$ is an $\varepsilon'$-kernel for $P'$ then $X \cup X'$ is an $\varepsilon$-kernel for $P \cup P'$

**Canonical Position**: We like fat things...

**Fat**: Given $0 \leq \alpha \leq 1$, a convex body $K$ is $\alpha$-fat if $K$ can be sandwiched between two concentric balls of radii $\lambda_- \leq \lambda_+$ where $\alpha = \lambda_- / \lambda_+$

**Canonical Position**: Convex body $K$ is in $\alpha$-canonical form if it is sandwiched between balls of radius $\lambda_- = \frac{1}{2} \alpha + \lambda_+ = \frac{1}{2}$ centered at the origin.
Why \( \frac{1}{2} \) \( \Rightarrow \) \( K \)'s diameter \( \leq 1 \)

A point set \( P \) is \( \alpha \)-fat \( \alpha \)-canonical form \( \) if \( \text{conv}(P) \) is.

We can convert any pt set into canonical form.

**Affine Transformation:** Is a linear transformation (scaling + rotation + shearing) + translation

**Lemma:** Given any \( n \)-element pt. set \( P \subseteq \mathbb{R}^d \), there exists an affine transformation \( T \) that maps \( P \) into \( (\frac{1}{d}) \)-canonical form.

- \( R \subseteq P \) is an \( \varepsilon \)-kernel for \( P \) iff \( T(R) \) is an \( \varepsilon \)-kernel for \( T(P) \)
- \( T \) can be computed in \( O(n) \) time
Proof makes use of important fact: (1948)

**John’s Theorem:** Given any convex body \( K \subseteq \mathbb{R}^d \), let \( E \) be the maximum volume ellipsoid contained in \( K \), then

\[
E \subseteq K \subseteq d \cdot E
\]

where \( d \cdot E \) is a factor-\( d \) scaling \( E \) about its center.

**Equiv:** Given pt. set \( P \), let \( E \) be the minimum volume ellipsoid containing \( P \), then

\[
\frac{1}{d} E \subseteq \text{conv}(P) \subseteq E
\]

- The ellipsoid is called the **John Ellipsoid** or **Löwner-John Ellipsoid**.
- Can compute it in \( \mathcal{O}(n) \) time (incremental).
Fact: Given any ellipsoid $E$, there exists an affine transformation that maps $E$ to a unit ball, centered at origin.

Proof (of canonical form lemma):
1. Compute $P$'s outer John ellipsoid $E$
2. Find affine transformation mapping $E$ to unit ball centered at origin
3. Scale by $\frac{1}{2}$ → output resulting transformation $T$

Why are directional width approximations preserved?
- Affine transformations preserve ratios of parallel lengths (details omitted)
Quick + Dirty Kernel: Simple but not optimal size $- \mathcal{O}(1/\varepsilon^d)$

Given $P \subseteq \mathbb{R}^d + \varepsilon > 0$

1. Map $P$ to $\frac{1}{d}$-canonical position
   
   Note: $\forall u, \frac{1}{d} \leq \|u(P)\|_2 \leq 1$
   
   $\Rightarrow$ absolute error of $\varepsilon/d \Rightarrow$ rel. error $\leq \varepsilon$

2. Let $\varepsilon' = \varepsilon/d$

3. Let $H = [-\frac{1}{2}, +\frac{1}{2}]^d$ be unit hypercube containing $P$

4. Subdivide $H$ into square grid of diameter $\leq \varepsilon'/(2d)$ (equiv., side length $= \varepsilon'/2\sqrt{d}$)

   Note: No. of grid cells is $(\frac{1}{\varepsilon'/2\sqrt{d}})^d = \mathcal{O}(1/\varepsilon^d)$

5. $R \leftarrow$ take one pt of $P$ from each occupied cell

   Note: $|R| = \mathcal{O}(1/\varepsilon^d)$, Computable in $O(n)$ time
Running time: $O(n + \frac{1}{\varepsilon^d})$
- Canonical position - $O(n)$
- Place pts in grid cells - $O(n)$
  [integer division + hashing]
- Output $R$ - $O(1/\varepsilon^d)$

Correctness:
- Given any direction $\hat{u}$, let $p_1, p_2 \in P$ be pts that define $W_u(L)$
- Let $q_1, q_2 \in R$ be corresponding representatives from $p_1$ and $p_2$'s cells

- Since cell diameter $\leq \varepsilon'/2$, it follows that
  $W_u(L) \leq \varepsilon'/2 + W_u(R) + \varepsilon'/2$
  $= \varepsilon' + W_u(R) = \varepsilon'/d + W_u(R)$
- By canonical form, $W_u(L) \geq \frac{1}{d}$
  $W_u(L) \leq \varepsilon \cdot W_u(L) + W_u(R)$
  $\Rightarrow (1 - \varepsilon) W_u(L) \leq W_u(R) \leq W_u(L)$
  $\Rightarrow R$ is an $\varepsilon$-kernel

\[ \square \]
Small Improvement: $O(1/\varepsilon^d) \rightarrow O(1/\varepsilon^{d-1})$

- Quick + dirty's grid includes many internal points $\rightarrow$ wasteful
- Rather than take:
  - one representative per cell, instead
  - two per column - toppmost + bottommost

- How many? Top grid has $O(1/\varepsilon^{d-1})$ cells
  $|R| = 2 \cdot O(1/\varepsilon^{d-1}) = O(1/\varepsilon^{d-1})$

- Correctness? Let $p$ be extreme pt in direction $\hat{u}$ + let $q \in R$ be toppmost (or bottommost) pt in column

  Directional distance betw. $q, p$ is $\leq \varepsilon/2$

  ... remaining details omitted
Big Improvement: ε-kernel of size $O\left(\frac{1}{\varepsilon^2} \right)$ [optimal in the worst case]

Construction based on idea discovered (independently) by Dudley + Bronstein + Ivanov (~1974)

1. Map $\mathcal{P}$ to $\frac{1}{d}$-canonical position
   
   Note: $\forall u, \frac{1}{d} \leq \mathcal{W}_u(\mathcal{P}) \leq 1$  
   $\Rightarrow$ absolute error of $\varepsilon/d \Rightarrow$ rel. error $\leq \varepsilon$

2. Let $\varepsilon' = \varepsilon/d$

3. Let $S$ = sphere of radius $2$ centered at origin, let $\delta = \sqrt{\varepsilon/4d}$

4. Let $Q$ be a set of points on $S$ s.t. any point of $S$ is within distance $\delta$ of some pt of $Q$. ($Q$ is “$\delta$-dense”)

5. For each $q \in Q$, let $\text{nn}(q) \in \mathcal{P}$ be its closest pt.  
Return: $R = \bigcup_{q \in Q} \text{nn}(q)$
Size: $|R| \leq |Q|$
- Claim that $|Q| = O\left(\frac{1}{\sqrt{\varepsilon}} \right)^{d-1} = O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$
- **Intuition:** Each $q \in Q$ covers a spherical cap of radius $\delta \approx \sqrt{\varepsilon}$
  - Such a cap has surface area $\approx \delta^{d-1} \approx \sqrt{\varepsilon}^{d-1} \approx \varepsilon^{(d-1)/2}$
  - $S$ has constant radius $\Rightarrow$ constant area
  - No. caps needed to cover $S$
    $\approx \text{const} / \varepsilon^{(d-1)/2} = O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$
  $\Rightarrow |R| = O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$

Running Time:
- **Canonical position:** $O(n)$
- Computing $\delta$-dense $Q$
  $O(|Q|) = O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$
  How? Enclose $S$ in a hypercube
  Cover hypercube with grid $\sim \delta$
  Project onto $S$
- Compute $\text{nn}(q)$
  - Sufficient to use approx nn
    $- O\left(\text{poly}\left(\frac{1}{\varepsilon}\right) \cdot \log n\right)$

![Diagram showing the spherical cap and hypercube relationship](image)
Correctness: (Complex—see latex notes)
We’ll consider a simpler question:

Why \( \sqrt{E} \) spacing?

Simple case: Suppose \( \text{conv}(P) \) is a unit ball

To generate a large error select \( u \) to hit \( \text{conv}(P) \) midway between pts of \( Q \).

By Pythagorean Thm:

\[
\text{error} \leq 1 - \sqrt{1 - \left(\frac{x}{2}\right)^2}
\]

We want

\[
\text{error} \leq \varepsilon
\]

That is, want \( x \) s.t.
Solving for $x$, we have:

\[
1 - \frac{1 - (x/2)^2}{(1 - \varepsilon)^2} \leq \varepsilon
\]

\[
\Leftrightarrow 1 - \frac{(x/2)^2}{1 - \varepsilon} \geq (1 - \varepsilon)^2 = 1 - 2\varepsilon + \varepsilon^2
\]

if $\varepsilon \leq 1$, then $\varepsilon^2 \leq \varepsilon \Rightarrow 1 - 2\varepsilon + \varepsilon^2 \leq 1 - 3$

\[
\Leftrightarrow 1 - \frac{(x/2)^2}{1 - \varepsilon} \geq 1 - 3
\]

\[
\Leftrightarrow \frac{x}{2} \leq \sqrt{\varepsilon}
\]

\[
x \leq 2\sqrt{\varepsilon}
\]

- This explains why $\text{spacing} \sim \sqrt{\varepsilon}$ is the right thing to do.

- Notice this is tight up to constant factors.