Optimal Approximate Polytope Membership

Sunil Arya
Hong Kong University of Science and Technology

Guilherme da Fonseca
Université d’Auvergne and LIMOS

David Mount
University of Maryland, College Park

SODA 2017
Polytope Membership Queries

Given a polytope $P$ in $d$-dimensional space, preprocess $P$ to answer membership queries:

Given a point $q$, is $q \in P$?

- Assume that dimension $d$ is a constant and $P$ is given as intersection of $n$ halfspaces
- Dual of halfspace emptiness searching
- For $d \leq 3$
  Query time: $O(\log n)$  Storage: $O(n)$
- For $d \geq 4$
  Query time: $O(\log n)$  Storage: $O(n^{\lceil d/2 \rceil})$
Approximate Polytope Membership Queries

Approximate Version

- An approximation parameter $\varepsilon > 0$ is given (at preprocessing time)
- Assume the polytope has diameter 1
- If the query point’s distance from $P$:
  - $0$: answer must be inside
  - $\geq \varepsilon$: answer must be outside
  - $> 0$ and $< \varepsilon$: either answer is acceptable

Previous solutions were either:

- **Time-efficient**
  - Query time: $O(\log \frac{1}{\varepsilon})$
  - Storage: $O(1/\varepsilon^{d-1})$

- **Space-efficient**
  - Query time: $\tilde{O}(1/\varepsilon^{(d-1)/8})$
  - Storage: $O(1/\varepsilon^{(d-1)/2})$
Time Efficient Solution [BFP82]

- Create a grid with cells of diameter $\varepsilon$
- For each column, store the topmost and bottommost cells intersecting $P$
- Query processing:
  - Locate the column that contains $q$
  - Compare $q$ with the two extreme values

Time Efficient Solution [BFP82]

- $O(1/\varepsilon^{d-1})$ columns
- Query time: $O(\log \frac{1}{\varepsilon})$ ← optimal
- Storage: $O(1/\varepsilon^{d-1})$
Time Efficient Solution [BFP82]

- Create a grid with cells of diameter $\varepsilon$
- For each column, store the topmost and bottommost cells intersecting $P$

Query processing:
- Locate the column that contains $q$
- Compare $q$ with the two extreme values

Time Efficient Solution [BFP82]

- $O\left(\frac{1}{\varepsilon^{d-1}}\right)$ columns
- Query time: $O\left(\log \frac{1}{\varepsilon}\right)$ ← optimal
- Storage: $O\left(\frac{1}{\varepsilon^{d-1}}\right)$
Time Efficient Solution [BFP82]

- Create a grid with cells of diameter $\varepsilon$
- For each column, store the topmost and bottommost cells intersecting $P$
- Query processing:
  - Locate the column that contains $q$
  - Compare $q$ with the two extreme values

$O(1/\varepsilon^{d-1})$ columns

Query time: $O(\log \frac{1}{\varepsilon})$ ← optimal

Storage: $O(1/\varepsilon^{d-1})$
Space Efficient Solution [AFM11, AFM12]

Preprocess:
- Input $P$, $\varepsilon$
- $t = \tilde{O}(1/\varepsilon^{(d-1)/8})$
- $Q \leftarrow$ unit hypercube
- Split-Reduce($Q$)

Split-Reduce($Q$)
- Find an $\varepsilon$-approximation of $Q \cap P$
- If at most $t$ facets, then $Q$ stores them
- Otherwise, subdivide $Q$ and recurse

Query time: $\tilde{O}(1/\varepsilon^{(d-1)/8})$
Storage: $O(1/\varepsilon^{(d-1)/2})$ ← optimal
Space Efficient Solution [AFM11, AFM12]

Preprocess:
- Input $P$, $\varepsilon$
- $t = \widetilde{O}(1/\varepsilon^{(d-1)/8})$
- $Q \leftarrow$ unit hypercube
- Split-Reduce($Q$)

Split-Reduce($Q$)
- Find an $\varepsilon$-approximation of $Q \cap P$
- If at most $t$ facets, then $Q$ stores them
- Otherwise, subdivide $Q$ and recurse

- Query time: $\widetilde{O}(1/\varepsilon^{(d-1)/8})$
- Storage: $O(1/\varepsilon^{(d-1)/2}) \leftarrow$ optimal
Space Efficient Solution [AFM11, AFM12]

Preprocess:
- Input $P$, $\varepsilon$
- $t = \tilde{O}(1/\varepsilon^{(d-1)/8})$
- $Q \leftarrow$ unit hypercube
- Split-Reduce($Q$)

Split-Reduce($Q$)
- Find an $\varepsilon$-approximation of $Q \cap P$
- If at most $t$ facets, then $Q$ stores them
- Otherwise, subdivide $Q$ and recurse

Query time: $\tilde{O}(1/\varepsilon^{(d-1)/8})$
Storage: $O(1/\varepsilon^{(d-1)/2})$ ← optimal

$t = 2$
Space Efficient Solution [AFM11, AFM12]

**Preprocess:**
- Input $P$, $\varepsilon$
- $t = \tilde{O}(1/\varepsilon^{(d-1)/8})$
- $Q \leftarrow$ unit hypercube
- Split-Reduce($Q$)

**Split-Reduce($Q$)**
- Find an $\varepsilon$-approximation of $Q \cap P$
- If at most $t$ facets, then $Q$ stores them
- Otherwise, subdivide $Q$ and recurse

- Query time: $\tilde{O}(1/\varepsilon^{(d-1)/8})$
- Storage: $O(1/\varepsilon^{(d-1)/2})$ $\leftarrow$ optimal
Space Efficient Solution [AFM11, AFM12]

Preprocess:
- Input $P, \varepsilon$
- $t = \tilde{O}(1/\varepsilon^{(d-1)/8})$
- $Q \leftarrow$ unit hypercube
- Split-Reduce($Q$)

Split-Reduce($Q$)
- Find an $\varepsilon$-approximation of $Q \cap P$
- If at most $t$ facets, then $Q$ stores them
- Otherwise, subdivide $Q$ and recurse

Query time: $\tilde{O}(1/\varepsilon^{(d-1)/8})$
Storage: $O(1/\varepsilon^{(d-1)/2}) \leftarrow$ optimal
New solution is **space-efficient** and **time-efficient**:

### Approximate Polytope Membership:

- Query time: $O(\log \frac{1}{\varepsilon})$ ← optimal
- Storage: $O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$ ← optimal

(Previous storage: $O\left(\frac{1}{\varepsilon^{d-1}}\right)$ [BFP82])

### Consequence:

### Approximate Nearest Neighbor Searching:

- Query time: $O(\log n)$
- Storage: $O\left(\frac{n}{\varepsilon^{d/2}}\right)$

(Previous storage: $O\left(\frac{n}{\varepsilon^{d-1}}\right)$ [Har01])
New solution is **space-efficient** and **time-efficient**:

### Approximate Polytope Membership:

- Query time: $O\left(\log \frac{1}{\varepsilon}\right)$ ← optimal
- Storage: $O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$ ← optimal

(Previous storage: $O\left(\frac{1}{\varepsilon^{d-1}}\right)$ [BFP82])

### Consequence:

### Approximate Nearest Neighbor Searching:

- Query time: $O\left(\log n\right)$
- Storage: $O\left(\frac{n}{\varepsilon^{d/2}}\right)$

(Previous storage: $O\left(\frac{n}{\varepsilon^{d-1}}\right)$ [Har01])
Techniques

- Previous solutions use grids and quadtrees
  - Similar width in all directions
- Our solution uses a hierarchy of Macbeath regions:
  - Adapt to the curvature of the body
  - Narrow in directions of high curvature
  - Wide in directions of low curvature
Macbeath Regions [Mac52]

Given a convex body $K$, $x \in K$, and $\lambda > 0$:

- $M^\lambda(x) = x + \lambda((K - x) \cap (x - K))$
- $M(x) = M^1(x)$: intersection of $K$ and $K$ reflected around $x$
- $M'(x) = M^{1/5}(x)$

Properties

- $M'(x) \cap M'(y) \neq \emptyset \Rightarrow M'(x) \subseteq M(y)$
- $y \in M'(x) \Rightarrow \delta(y) = \Theta(\delta(x))$, where $\delta(x)$: distance from $x$ to $\partial K$
Macbeath Regions [Mac52]

Given a convex body $K$, $x \in K$, and $\lambda > 0$:

- $M^\lambda(x) = x + \lambda((K - x) \cap (x - K))$
- $M(x) = M^1(x)$: intersection of $K$ and $K$ reflected around $x$
- $M'(x) = M^{1/5}(x)$

Properties

- $M'(x) \cap M'(y) \neq \emptyset \implies M'(x) \subseteq M(y)$
- $y \in M'(x) \implies \delta(y) = \Theta(\delta(x))$, where $\delta(x)$: distance from $x$ to $\partial K$
Macbeath Regions [Mac52]

Given a convex body $K$, $x \in K$, and $\lambda > 0$:

- $M^\lambda(x) = x + \lambda((K - x) \cap (x - K))$
- $M(x) = M^1(x)$: intersection of $K$ and $K$ reflected around $x$
- $M'(x) = M^{1/5}(x)$

Properties

- $M'(x) \cap M'(y) \neq \emptyset \Rightarrow M'(x) \subseteq M(y)$
- $y \in M'(x) \Rightarrow \delta(y) = \Theta(\delta(x))$, where $\delta(x)$: distance from $x$ to $\partial K$
Macbeath Regions [Mac52]

Given a convex body $K$, $x \in K$, and $\lambda > 0$:

- $M^\lambda(x) = x + \lambda((K - x) \cap (x - K))$
- $M(x) = M^1(x)$: intersection of $K$ and $K$ reflected around $x$
- $M'(x) = M^{1/5}(x)$

Properties

- $M'(x) \cap M'(y) \neq \emptyset \Rightarrow M'(x) \subseteq M(y)$
- $y \in M'(x) \Rightarrow \delta(y) = \Theta(\delta(x))$, where $\delta(x)$: distance from $x$ to $\partial K$
Macbeath Regions [Mac52]

Given a convex body \( K, x \in K \), and \( \lambda > 0 \):
- \( \mathcal{M}^\lambda(x) = x + \lambda((K - x) \cap (x - K)) \)
- \( \mathcal{M}(x) = \mathcal{M}^1(x) \): intersection of \( K \) and \( K \) reflected around \( x \)
- \( \mathcal{M}'(x) = \mathcal{M}^{1/5}(x) \)

Properties
- \( \mathcal{M}'(x) \cap \mathcal{M}'(y) \neq \emptyset \Rightarrow \mathcal{M}'(x) \subseteq \mathcal{M}(y) \)
- \( y \in \mathcal{M}'(x) \Rightarrow \delta(y) = \Theta(\delta(x)) \), where \( \delta(x) \): distance from \( x \) to \( \partial K \)
Macbeath Regions [Mac52]

Given a convex body $K$, $x \in K$, and $\lambda > 0$:

- $M^\lambda(x) = x + \lambda((K - x) \cap (x - K))$
- $M(x) = M^1(x)$: intersection of $K$ and $K$ reflected around $x$
- $M'(x) = M^{1/5}(x)$

Properties

- $M'(x) \cap M'(y) \neq \emptyset \Rightarrow M'(x) \subseteq M(y)$
- $y \in M'(x) \Rightarrow \delta(y) = \Theta(\delta(x))$, where $\delta(x)$: distance from $x$ to $\partial K$
Macbeath Regions [Mac52]

Given a convex body $K$, $x \in K$, and $\lambda > 0$:

- $M^\lambda(x) = x + \lambda((K - x) \cap (x - K))$
- $M(x) = M^1(x)$: intersection of $K$ and $K$ reflected around $x$
- $M'(x) = M^{1/5}(x)$

Properties

- $M'(x) \cap M'(y) \neq \emptyset \Rightarrow M'(x) \subseteq M(y)$
- $y \in M'(x) \Rightarrow \delta(y) = \Theta(\delta(x))$, where $\delta(x)$: distance from $x$ to $\partial K$
Macbeath Regions [Mac52]

Given a convex body $K$, $x \in K$, and $\lambda > 0$:

- $M^\lambda(x) = x + \lambda((K - x) \cap (x - K))$
- $M(x) = M^1(x)$: intersection of $K$ and $K$ reflected around $x$
- $M'(x) = M^{1/5}(x)$

Properties

- $M'(x) \cap M'(y) \neq \emptyset \Rightarrow M'(x) \subseteq M(y)$
- $y \in M'(x) \Rightarrow \delta(y) = \Theta(\delta(x))$, where $\delta(x)$: distance from $x$ to $\partial K$
Macbeath Regions [Mac52]

Given a convex body $K$, $x \in K$, and $\lambda > 0$:

- $M^\lambda(x) = x + \lambda((K - x) \cap (x - K))$
- $M(x) = M^1(x)$: intersection of $K$ and $K$ reflected around $x$
- $M'(x) = M^{1/5}(x)$

Properties

- $M'(x) \cap M'(y) \neq \emptyset \Rightarrow M'(x) \subseteq M(y)$
- $y \in M'(x) \Rightarrow \delta(y) = \Theta(\delta(x))$, where $\delta(x)$: distance from $x$ to $\partial K$
Macbeath Ellipsoids

John Ellipsoid [Joh48]
For every centrally symmetric convex body $K$ in $\mathbb{R}^d$, there exist ellipsoids $E_1, E_2$ such that $E_1 \subseteq K \subseteq E_2$ and $E_2$ is a $\sqrt{d}$-scaling of $E_1$.

Macbeath Ellipsoid
- $E(x)$: enclosed John ellipsoid of $M'(x)$
- $M^\lambda(x) \subseteq E(x) \subseteq M'(x)$ for $\lambda = 1/(5\sqrt{d})$
Macbeath Ellipsoids

John Ellipsoid [Joh48]
For every centrally symmetric convex body $K$ in $\mathbb{R}^d$, there exist ellipsoids $E_1, E_2$ such that $E_1 \subseteq K \subseteq E_2$ and $E_2$ is a $\sqrt{d}$-scaling of $E_1$.

Macbeath Ellipsoid
- $E(x)$: enclosed John ellipsoid of $M'(x)$
- $M^\lambda(x) \subseteq E(x) \subseteq M'(x)$ for $\lambda = 1/(5\sqrt{d})$
Macbeath Ellipsoids

John Ellipsoid [Joh48]

For every centrally symmetric convex body $K$ in $\mathbb{R}^d$, there exist ellipsoids $E_1, E_2$ such that $E_1 \subseteq K \subseteq E_2$ and $E_2$ is a $\sqrt{d}$-scaling of $E_1$

Macbeath Ellipsoid

- $E(x)$: enclosed John ellipsoid of $M'(x)$
- $M^\lambda(x) \subseteq E(x) \subseteq M'(x)$ for $\lambda = 1/(5\sqrt{d})$
Covering with Macbeath Ellipsoids

Covering (see [Bar07])

Given:
- $K$: convex body
- $\delta$: small positive parameter

There exist ellipsoids $E(x_1), \ldots, E(x_k)$
- $\delta(x_1) = \cdots = \delta(x_k) = \delta$
- Cover: Every ray from the origin intersects some ellipsoid

$k = O\left(1/\delta^{(d-1)/2}\right)$ [AFM16]
Covering with Macbeath Ellipsoids

Given:
- $K$: convex body
- $\delta$: small positive parameter

There exist ellipsoids $E(x_1), \ldots, E(x_k)$
- $\delta(x_1) = \cdots = \delta(x_k) = \delta$

Cover: Every ray from the origin intersects some ellipsoid
- $k = O(1/\delta^{(d-1)/2})$ [AFM16]
Covering (see [Bar07])

Given:
- \( K \): convex body
- \( \delta \): small positive parameter

There exist ellipsoids \( E(x_1), \ldots, E(x_k) \)
- \( \delta(x_1) = \cdots = \delta(x_k) = \delta \)
- **Cover**: Every ray from the origin intersects some ellipsoid
- \( k = O(1/\delta^{(d-1)/2}) \) [AFM16]
Hierarchy of Macbeath Ellipsoids

**Hierarchy**

**Given:**
- $K$: convex body
- $\varepsilon$: small positive parameter

**Hierarchy:**
- Each level $i$ a $\delta_i$-covering
- $\ell = \Theta(\log \frac{1}{\varepsilon})$ levels
- $\delta_0 = \Theta(1)$, $\delta_\ell = \Theta(\varepsilon)$
- $\delta_{i+1} = \delta_i/2$
- $E, E'$ are parent/child if
  - Levels are consecutive
  - Same ray from the origin intersects $E$ and $E'$
- Each node has $O(1)$ children
Hierarchy of Macbeath Ellipsoids

Hierarchy

Given:
- \( K \): convex body
- \( \varepsilon \): small positive parameter

Hierarchy:
- Each level \( i \) a \( \delta_i \)-covering
- \( \ell = \Theta(\log \frac{1}{\varepsilon}) \) levels
- \( \delta_0 = \Theta(1), \delta_\ell = \Theta(\varepsilon) \)
- \( \delta_{i+1} = \delta_i / 2 \)
- \( E, E' \) are parent/child if
  - Levels are consecutive
  - Same ray from the origin intersects \( E \) and \( E' \)
- Each node has \( O(1) \) children
Hierarchy of Macbeath Ellipsoids

**Hierarchy**

**Given:**
- \( K \): convex body
- \( \varepsilon \): small positive parameter

**Hierarchy:**
- Each level \( i \) a \( \delta_i \)-covering
- \( \ell = \Theta(\log \frac{1}{\varepsilon}) \) levels
- \( \delta_0 = \Theta(1) \), \( \delta_\ell = \Theta(\varepsilon) \)
- \( \delta_{i+1} = \delta_i / 2 \)
- \( E, E' \) are parent/child if
  - Levels are consecutive
  - Same ray from the origin intersects \( E \) and \( E' \)
- Each node has \( O(1) \) children
Ray Shooting from the Origin

generalizes polytope membership

Preprocess:
- \( K \): convex body
- \( \varepsilon \): small positive parameter

Query:
- \( Oq \): ray from the origin towards \( q \)

Query algorithm:
- Find an ellipsoid intersecting \( Oq \) at level 0
- Repeat among children at next level
- Stop at leaf node
- Leaf ellipsoid \( \varepsilon \)-approximates boundary
Ray Shooting from the Origin

**Ray Shooting from the Origin**
(generalizes polytope membership)

Preprocess:
- \( K \): convex body
- \( \varepsilon \): small positive parameter

Query:
- \( Oq \): ray from the origin towards \( q \)

Query algorithm:
- Find an ellipsoid intersecting \( Oq \) at level 0
- Repeat among children at next level
- Stop at leaf node
- Leaf ellipsoid \( \varepsilon \)-approximates boundary
Ray Shooting from the Origin

(generalizes polytope membership)

Preprocess:
- $K$: convex body
- $\varepsilon$: small positive parameter

Query:
- $Oq$: ray from the origin towards $q$

Query algorithm:
- Find an ellipsoid intersecting $Oq$
  at level 0
- Repeat among children at next level
- Stop at leaf node
- Leaf ellipsoid $\varepsilon$-approximates boundary
Analysis

- Out-degree: $O(1)$
- Query time per level: $O(1)$
- Number of levels: $O(\log \frac{1}{\varepsilon})$
- Query time: $O(\log \frac{1}{\varepsilon})$

- Storage for bottom level:
  $O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$

- Geometric progression of storage per level

- Total storage: $O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$
Approximate Nearest Neighbor

Preprocess $n$ points such that, given a query point $q$, we can find a point within at most $1 + \varepsilon$ times the distance to $q$’s nearest neighbor.

- For $\log \frac{1}{\varepsilon} \leq m \leq \frac{1}{\varepsilon^{d/2}}$
  
  Query time: $O(\log n + \frac{1}{m \varepsilon^{d/2}})$
  
  Storage: $O(nm)$

- If $m = \frac{1}{\varepsilon^{d/2}}$
  
  Query time: $O(\log n)$
  
  Storage: $O(n/\varepsilon^{d/2})$
What else is in the paper?

- Proofs
- Witness (important to find the approximate nearest neighbor)
- Reduction from ANN to approximate ray shooting

Full Paper
arxiv.org/abs/1612.01696
Conclusions and Open Problems

Our approximate polytope membership data structure is optimal

- Query time: $O(\log \frac{1}{\varepsilon})$
- Storage: $O\left(\frac{1}{\varepsilon^{(d-1)/2}}\right)$

Still, several open problems remain

- Further improvements to approximate nearest neighbor searching?
- Generalization to $k$-nearest neighbors?
- Other applications of the hierarchy?

Recent applications of the hierarchy

- Near-optimal $\varepsilon$-kernel computation
- Approximate diameter
- Approximate bichromatic closest pair
Bibliography


Thank you!

Sculpture by José Mérino