Coresets (for directional width)

Geometric Approximation Algorithms

Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

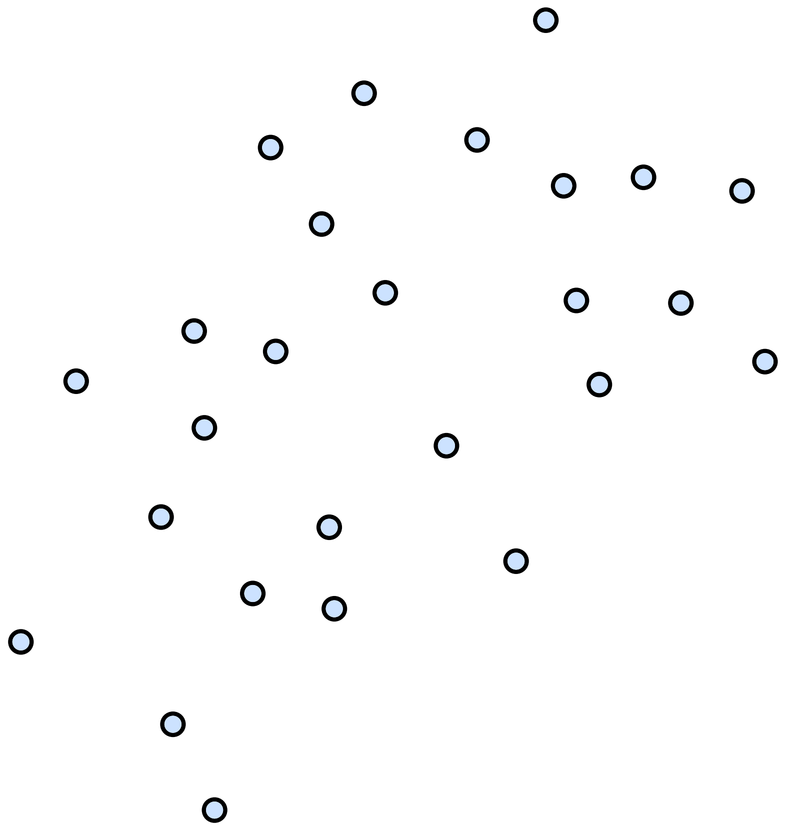
Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

a trivial (useless) example:

geometric problem: smallest axis-aligned bounding box of $P \subset \mathbb{R}^d$



Coreset

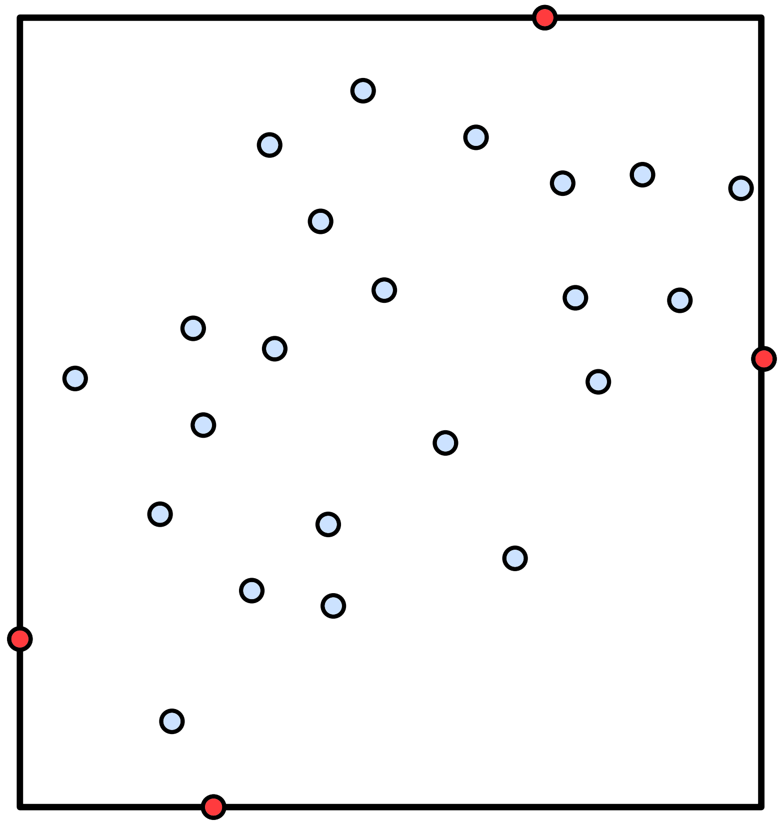
setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

a trivial (useless) example:

geometric problem: smallest axis-aligned bounding box of $P \subset \mathbb{R}^d$

S : points with minimum and maximum coordinate (for each coordinate)



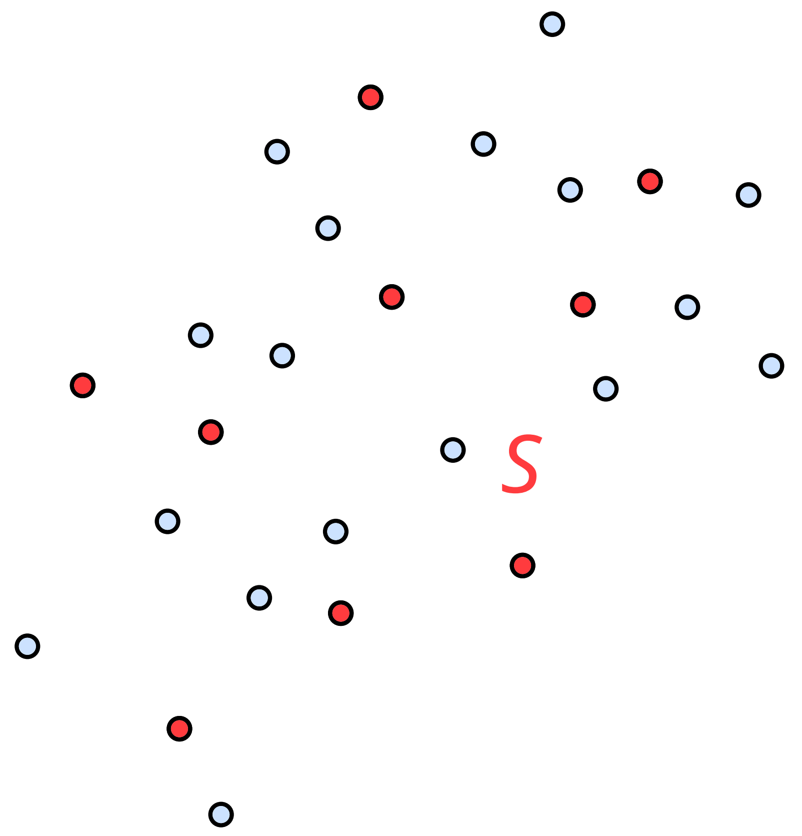
Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

example:

geometric problem: min-volume bounding box of $P \subset \mathbb{R}^d$



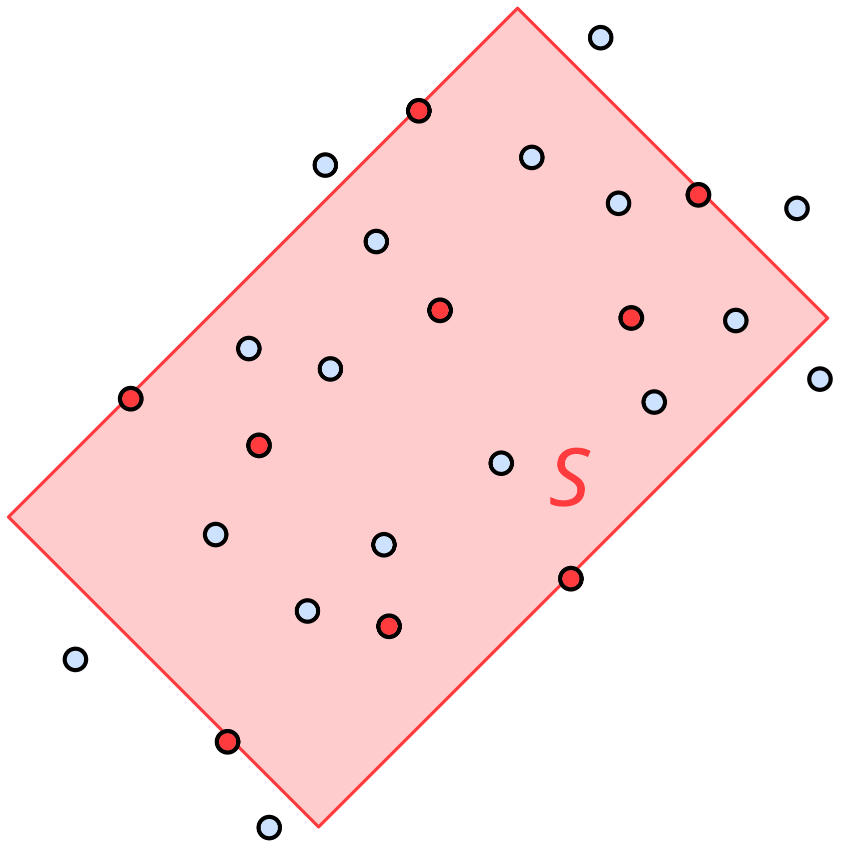
Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

example:

geometric problem: min-volume bounding box of $P \subset \mathbb{R}^d$



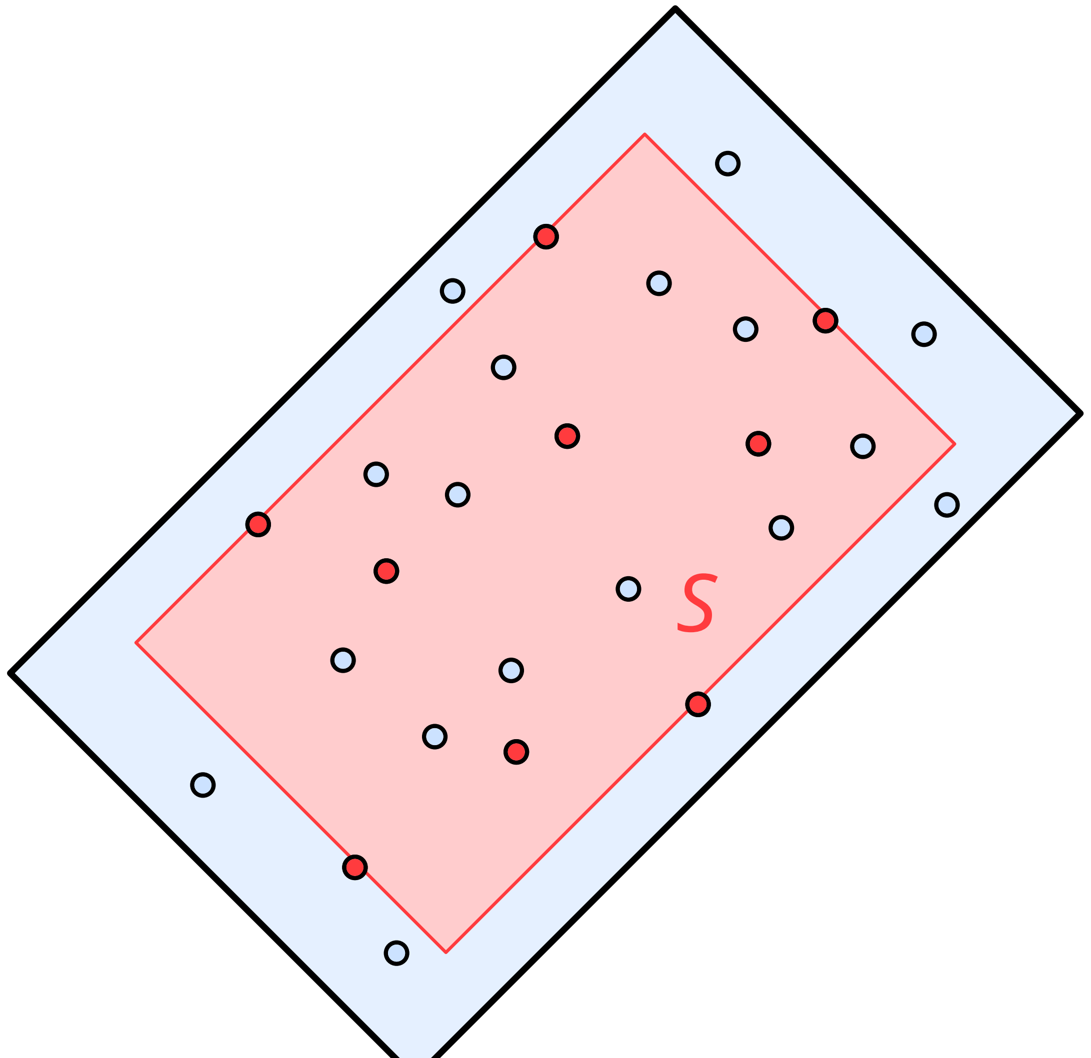
Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

example:

geometric problem: min-volume bounding box of $P \subset \mathbb{R}^d$



Coreset

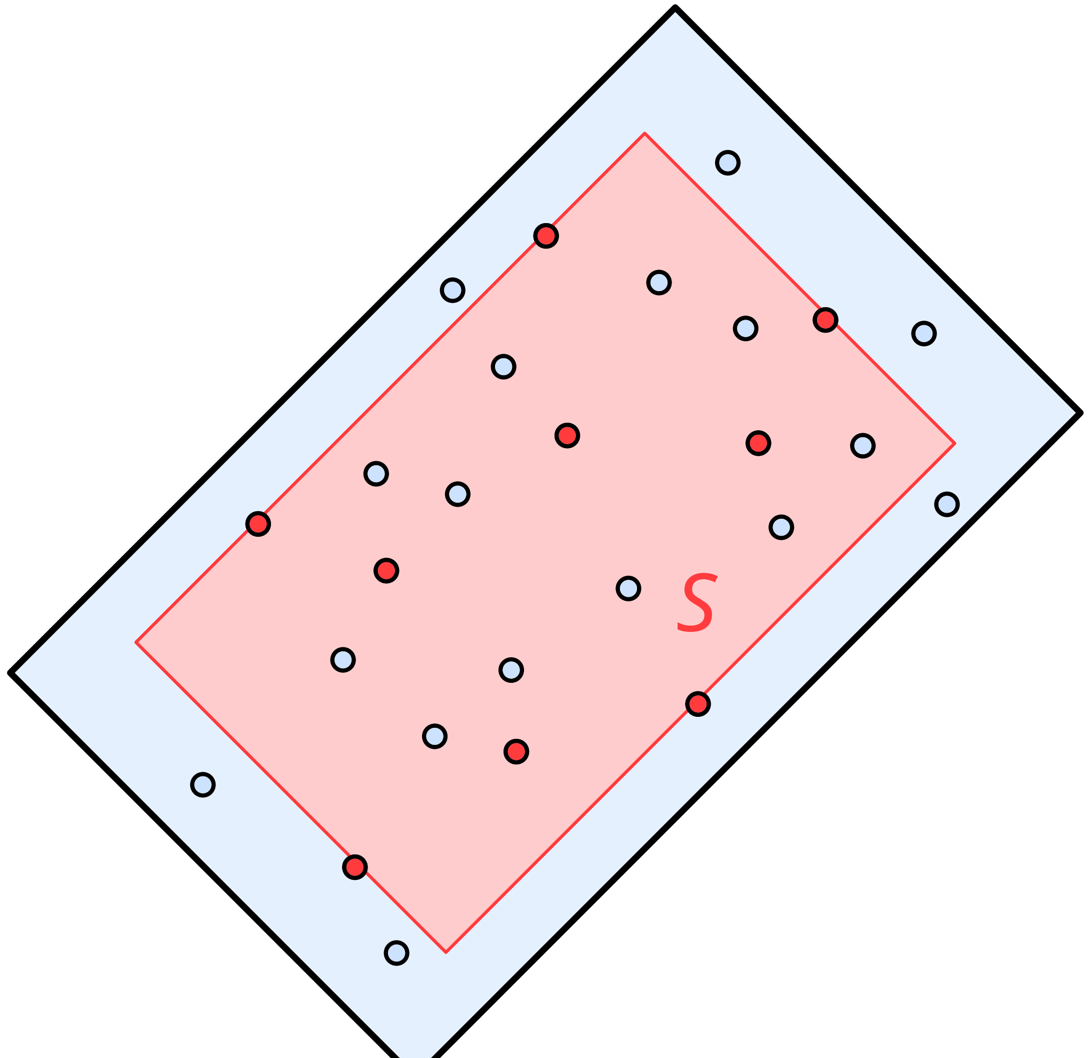
setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

example:

geometric problem: min-volume bounding box of $P \subset \mathbb{R}^d$

Why?



Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

example:

geometric problem: min-volume bounding box of $P \subset \mathbb{R}^d$

Why?

slow (exact) algorithm + coreset = fast approximation algorithm

Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

example:

geometric problem: min-volume bounding box of $P \subset \mathbb{R}^d$

Why?

slow (exact) algorithm + coreset = fast approximation algorithm

example (continued):

exact algorithm: min-volume bounding box of $P \subset \mathbb{R}^3$ in $O(n^3)$ time.

Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

example:

geometric problem: min-volume bounding box of $P \subset \mathbb{R}^d$

Why?

slow (exact) algorithm + coreset = fast approximation algorithm

example (continued):

exact algorithm: min-volume bounding box of $P \subset \mathbb{R}^3$ in $O(n^3)$ time.

coreset for min-volume in $\mathbb{R}^3$ with "error" ε:

size $O(1/\varepsilon^{3/2})$, construction time $O(n + 1/\varepsilon^{9/2})$

Coreset

setting: Geometric optimization problem on a point set P

coreset: Small subset $S \subset P$ that captures the structure of optimal solution

example:

geometric problem: min-volume bounding box of $P \subset \mathbb{R}^d$

Why?

slow (exact) algorithm + coreset = fast approximation algorithm

example (continued):

exact algorithm: min-volume bounding box of $P \subset \mathbb{R}^3$ in $O(n^3)$ time.

coreset for min-volume in $\mathbb{R}^3$ with "error" ε:

size $O(1/\varepsilon^{3/2})$, construction time $O(n + 1/\varepsilon^{9/2})$

combined: $(1 + \varepsilon)$-approximation in $O(n + 1/\varepsilon^{9/2})$ time

Overview

Coreset for directional width

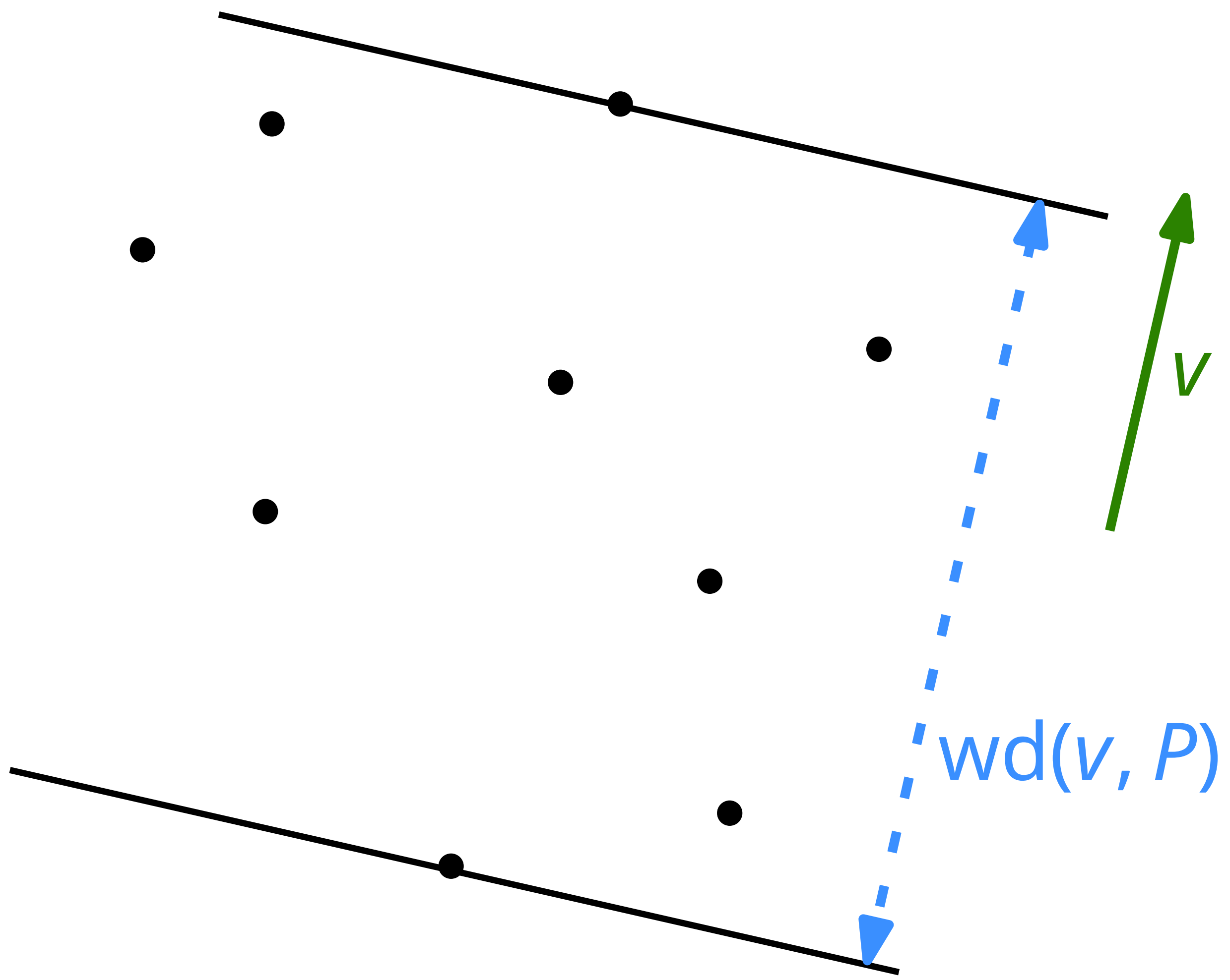
- definition
- applications
- construction algorithm

Extra ingredient: Minimum volume bounding box

Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

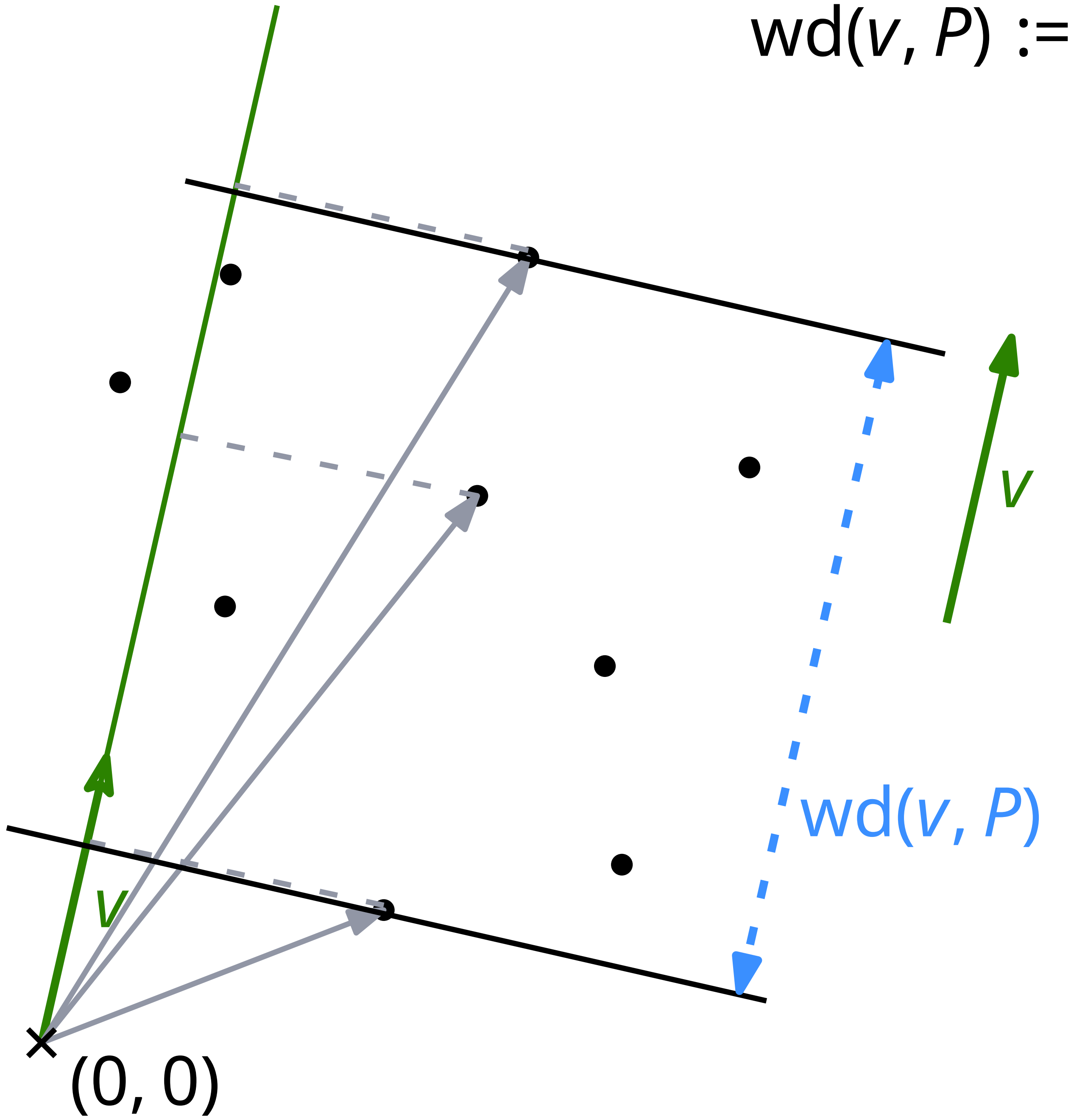
$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$



Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

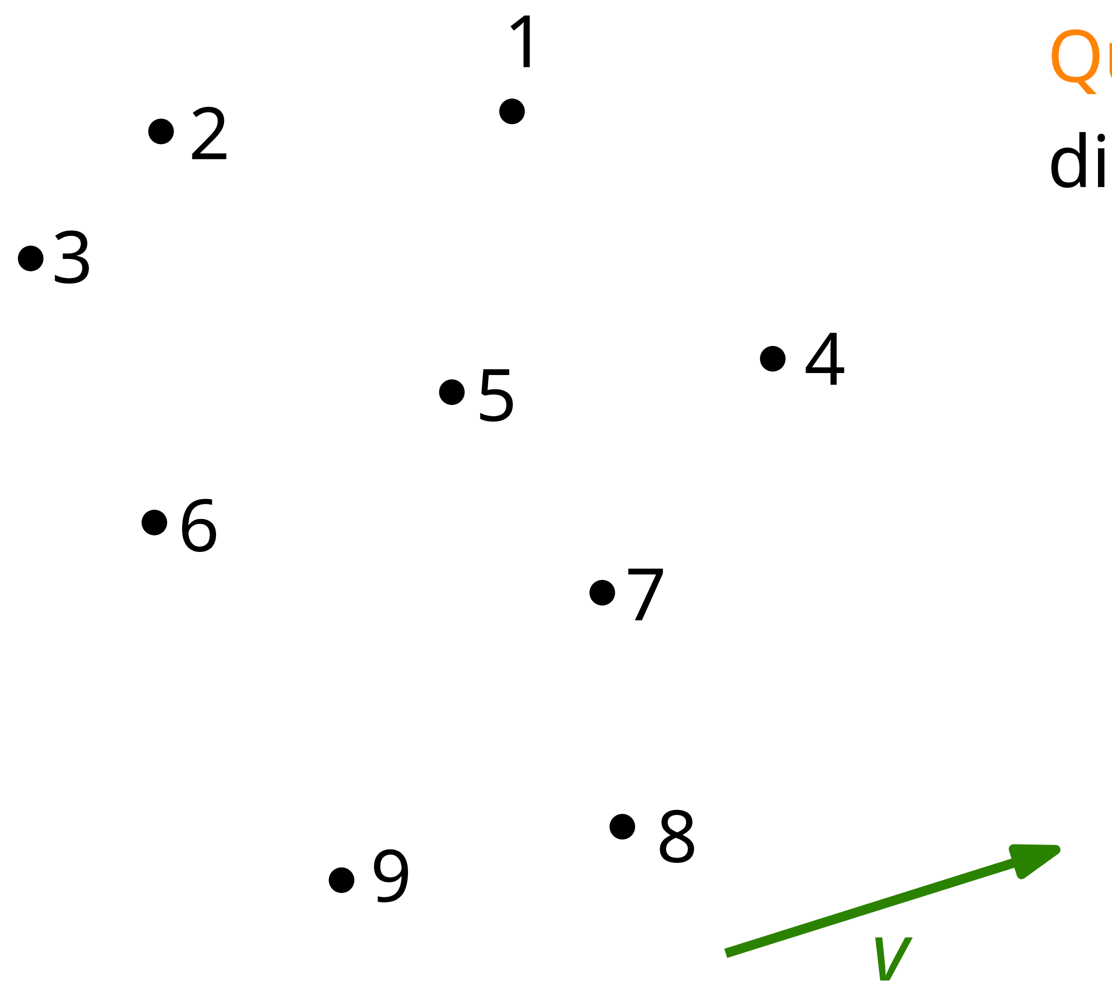
$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$



Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$

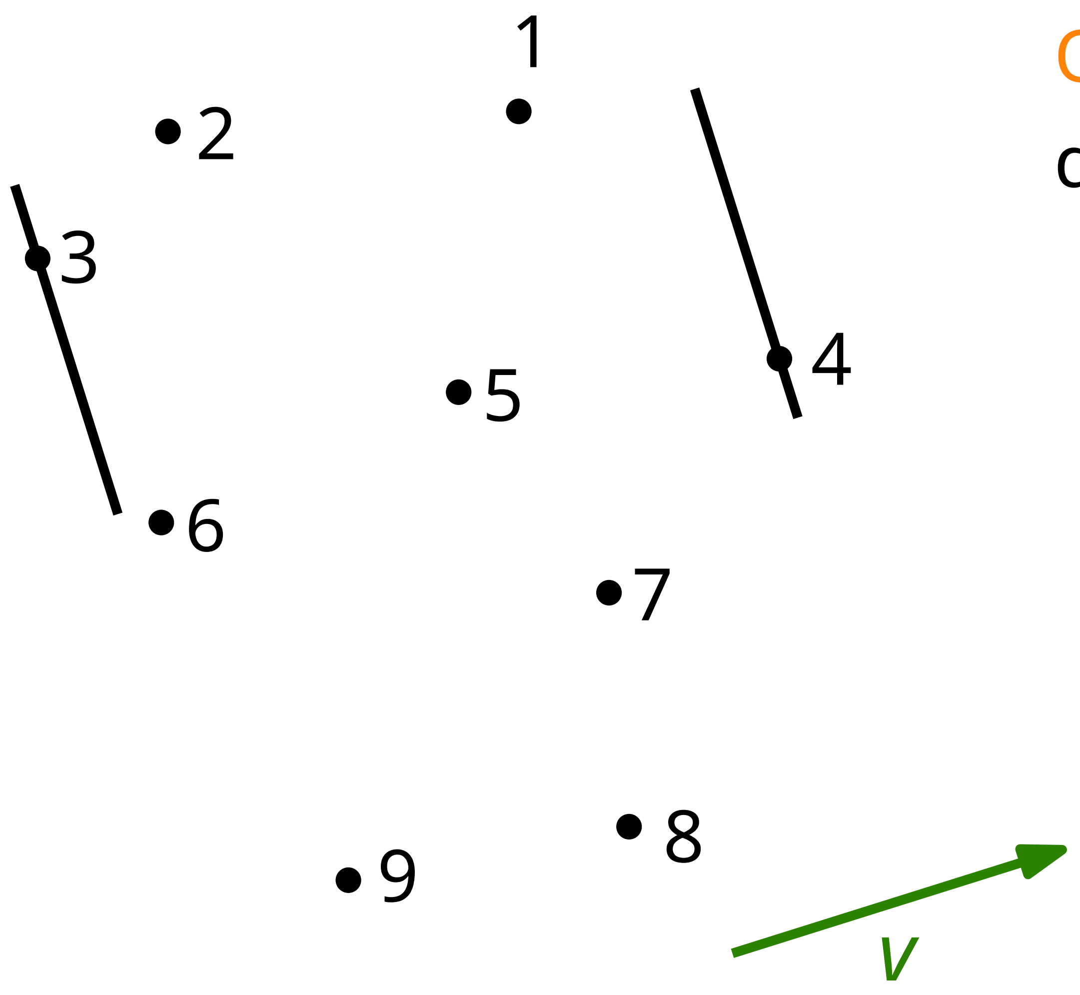


Question 1: Which two points give the width in direction v?

Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$



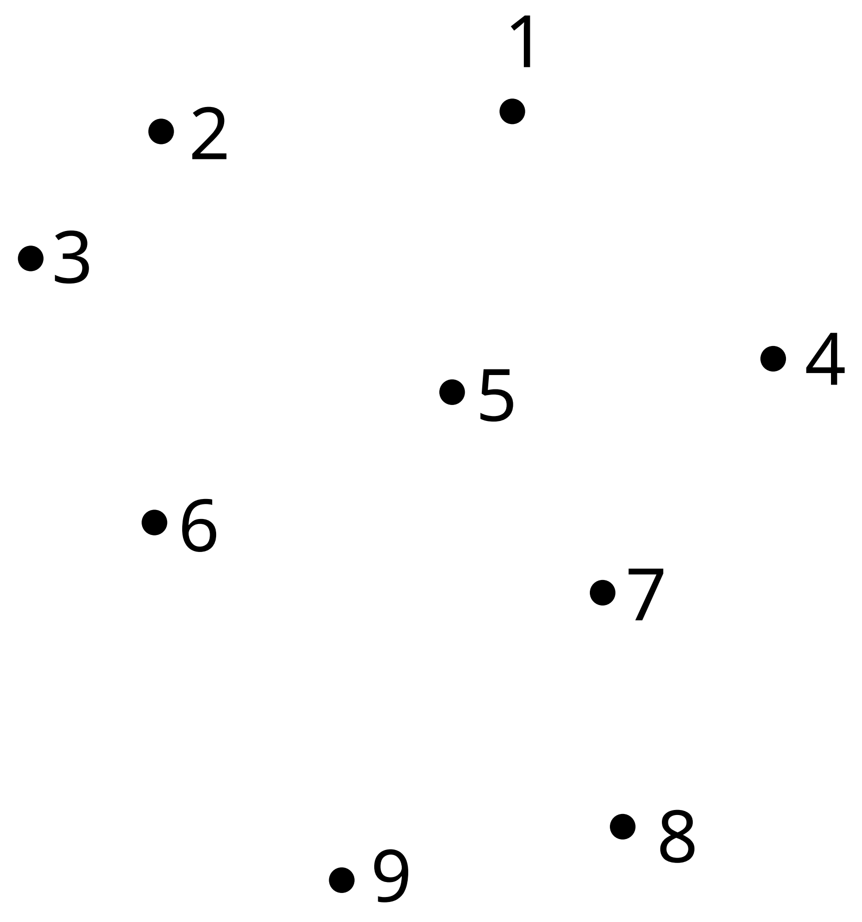
Question 1: Which two points give the width in direction v?

3, 4

Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$



Question 1: Which two points give the width in direction v?

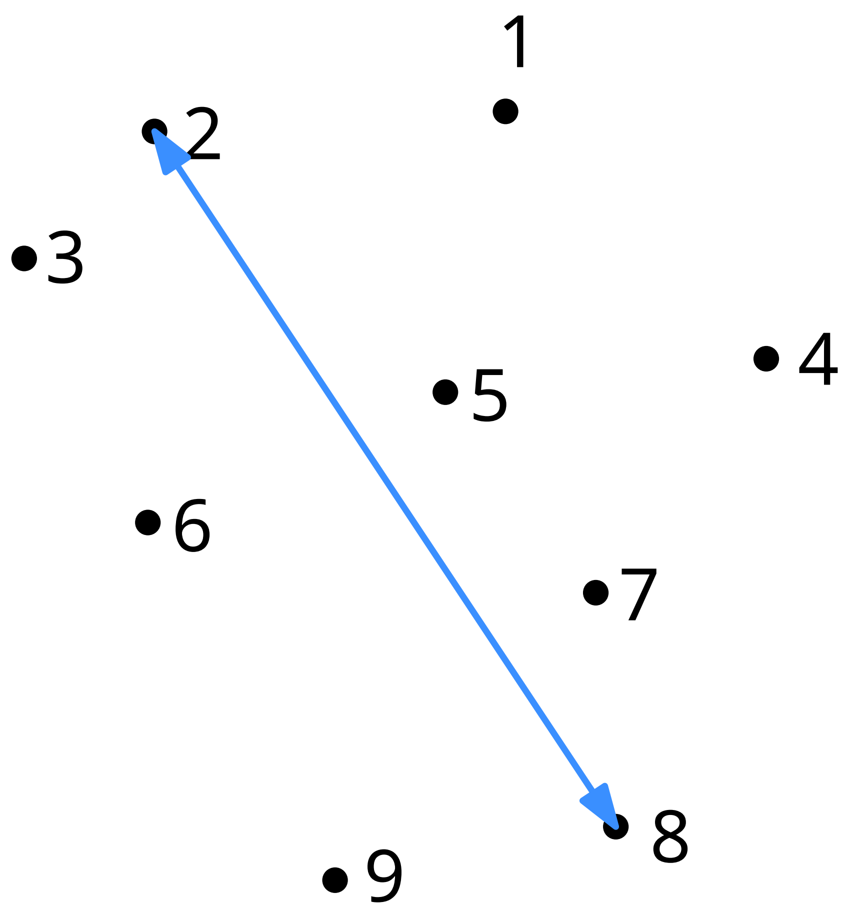
3, 4

Question 2: Which v with $\|v\| = 1$ has the largest directional width?

Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$



Question 1: Which two points give the width in direction v?

3, 4

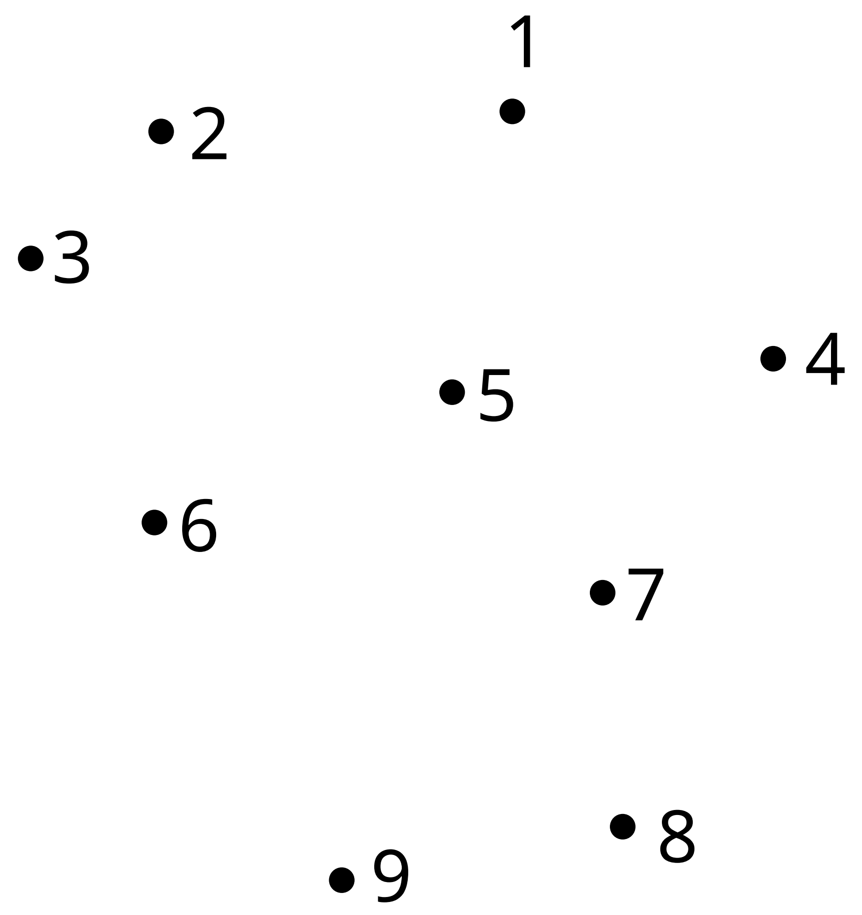
Question 2: Which v with $\|v\| = 1$ has the largest directional width?

The unit vector in direction from 2 to 8 (or vice versa)

Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$



Question 1: Which two points give the width in direction v?

3, 4

Question 2: Which v with $\|v\| = 1$ has the largest directional width?

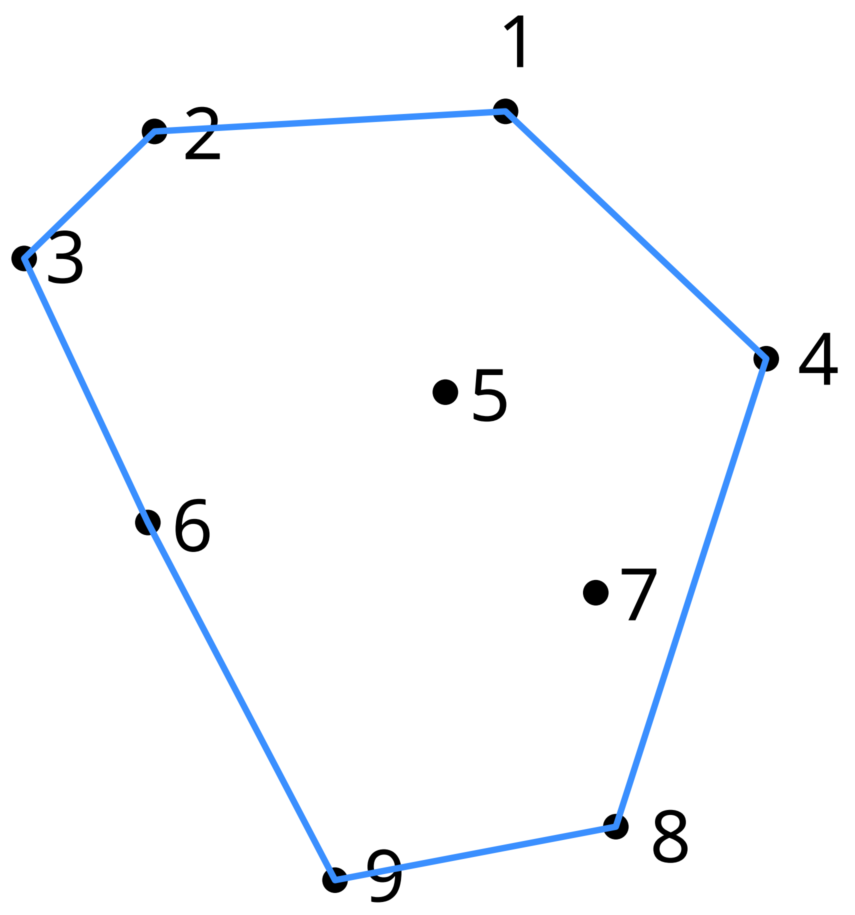
The unit vector in direction from 2 to 8 (or vice versa)

Question 3: Which are the only points of P relevant for computing directional width?

Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$



Question 1: Which two points give the width in direction v?

3, 4

Question 2: Which v with $\|v\| = 1$ has the largest directional width?

The unit vector in direction from 2 to 8 (or vice versa)

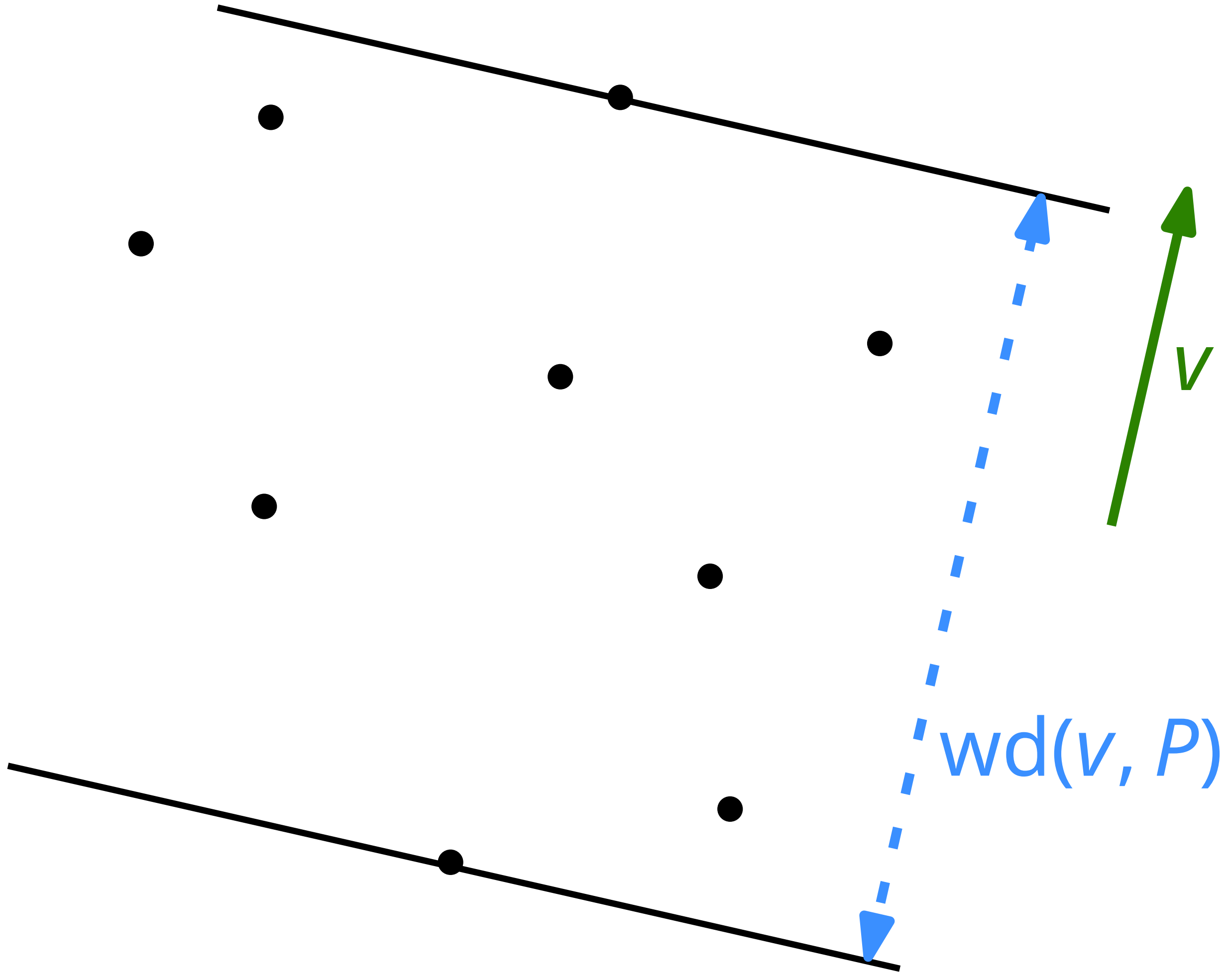
Question 3: Which are the only points of P relevant for computing directional width?

Only points on convex hull
(i.e. 1, 2, 3, 6, 9, 8, 4)

Directional Width

Definition: The directional width of $P \subset \mathbb{R}^d$ with respect to a vector $v \in \mathbb{R}^d \setminus \{0\}$ is

$$\mathrm{wd}(v, P) := \max_{p \in P} \langle v, p \rangle - \min_{p \in P} \langle v, p \rangle$$



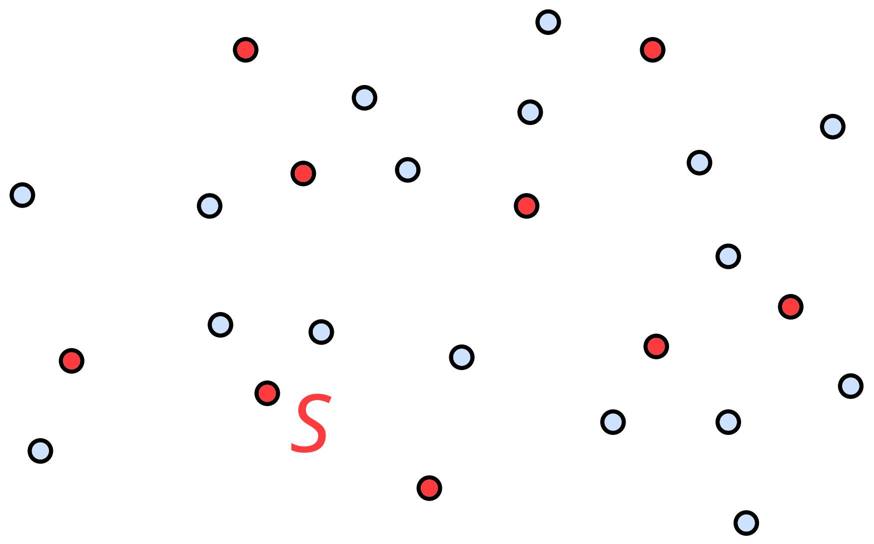
Properties:

- translation invariant
- scales linearly
- $\mathrm{wd}(v, P) = \mathrm{wd}(v, \mathrm{conv}(P))$
- monotone: if $Q \subset P$, then $\mathrm{wd}(v, Q) \leq \mathrm{wd}(v, P)$

Coreset for directional width

Definition: The set $S \subset P$ is an ε-coreset for directional width if for every unit vector v

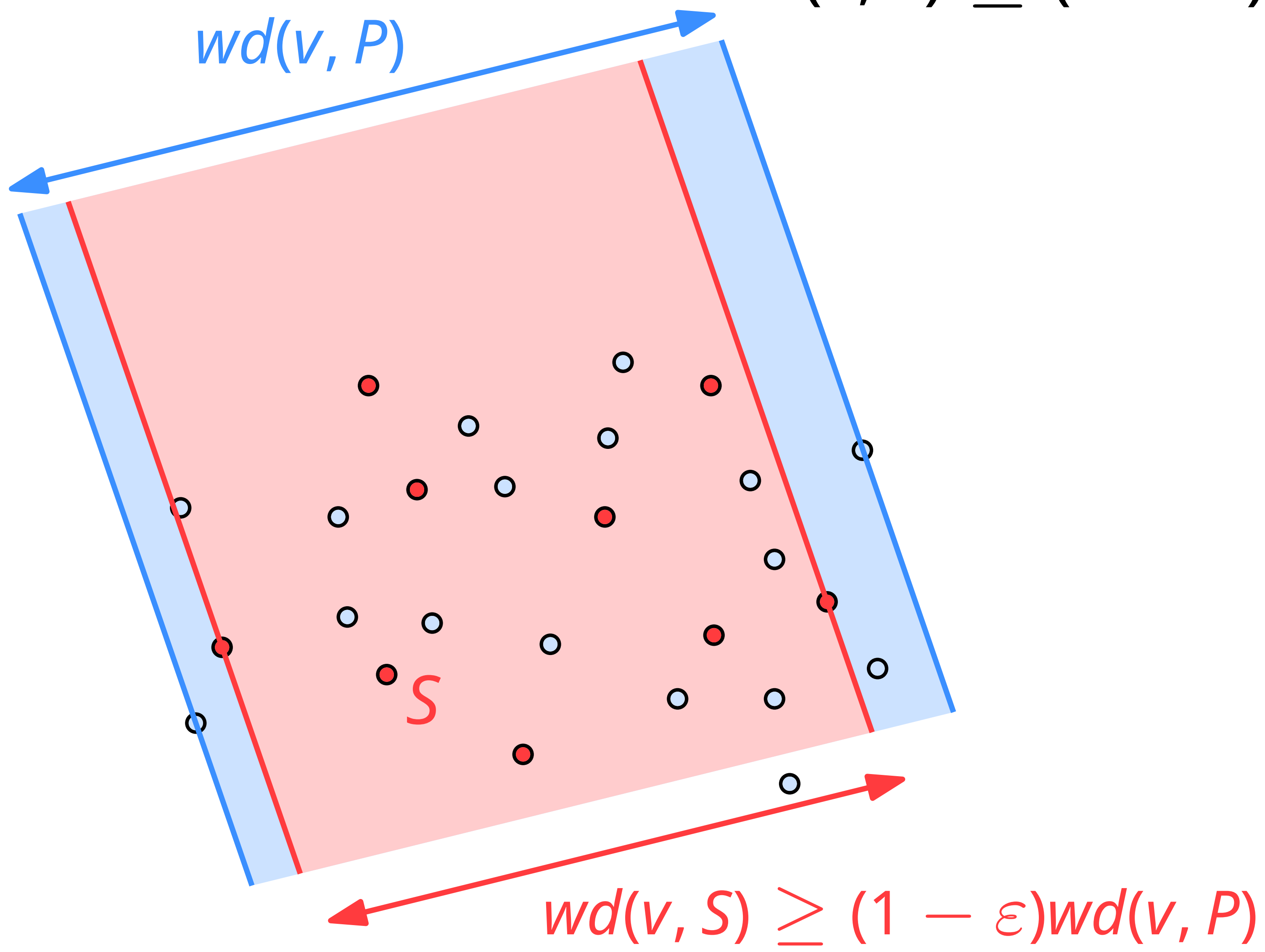
$$\text{wd}(v, S) \geq (1 - \varepsilon)\text{wd}(v, P)$$



Coreset for directional width

Definition: The set $S \subset P$ is an ε-coreset for directional width if for every unit vector v

$$\mathrm{wd}(v, S) \geq (1 - \varepsilon)\mathrm{wd}(v, P)$$



Coreset for directional width

Definition: The set $S \subset P$ is an ε-coreset for directional width if for every unit vector v

$$\text{wd}(v, S) \geq (1 - \varepsilon)\text{wd}(v, P)$$

- captures the (outer) geometry of P, and (hopefully) is much smaller than P

Coreset for directional width

Definition: The set $S \subset P$ is an ε-coreset for directional width if for every unit vector v

$$\mathrm{wd}(v, S) \geq (1 - \varepsilon)\mathrm{wd}(v, P)$$

- captures the (outer) geometry of P, and (hopefully) is much smaller than P
- sketch property: If S is ε-coreset of S' and S' is ε'-coreset of P, then S is $(\varepsilon + \varepsilon')$-coreset of P (Exercise)

Coreset for directional width

Definition: The set $S \subset P$ is an ε-coreset for directional width if for every unit vector v

$$\mathrm{wd}(v, S) \geq (1 - \varepsilon)\mathrm{wd}(v, P)$$

- captures the (outer) geometry of P, and (hopefully) is much smaller than P
- sketch property: If S is ε-coreset of S' and S' is ε'-coreset of P, then S is $(\varepsilon + \varepsilon')$-coreset of P (Exercise)
- merge property: If S is ε-coreset of P and S' is ε'-coreset of P', then $S \cup S'$ is $(\varepsilon + \varepsilon')$-coreset of $P \cup P'$ (Exercise)

Coreset for directional width

Definition: The set $S \subset P$ is an ε-coreset for directional width if for every unit vector v

$$\mathrm{wd}(v, S) \geq (1 - \varepsilon)\mathrm{wd}(v, P)$$

- captures the (outer) geometry of P, and (hopefully) is much smaller than P
- sketch property: If S is ε-coreset of S' and S' is ε'-coreset of P, then S is $(\varepsilon + \varepsilon')$-coreset of P (Exercise)
- merge property: If S is ε-coreset of P and S' is ε'-coreset of P', then $S \cup S'$ is $(\varepsilon + \varepsilon')$-coreset of $P \cup P'$ (Exercise)
- general concept: other objective functions f instead of directional width

Overview

Coreset for directional width

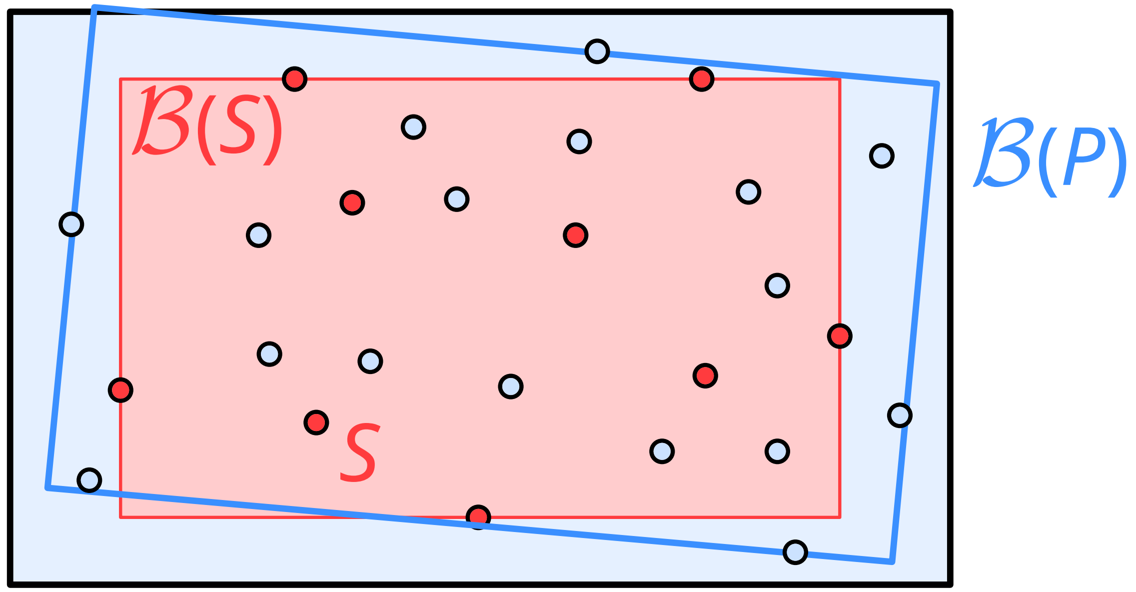
- definition
- applications ←
- construction algorithm

Extra ingredient: Minimum volume bounding box

Use case: min volume bounding box

Given $\varepsilon > 0$, $P \subset \mathbb{R}^d$, let S be a δ-coreset of P for directional width ($\delta = \varepsilon/(8d)$). Let $\mathcal{B}(P)$ (resp. $\mathcal{B}(S)$) be the min volume bounding box of P (resp. S), and let B be $\mathcal{B}(S)$ scaled by $(1 + 3\delta)$ around the center of $\mathcal{B}(S)$.

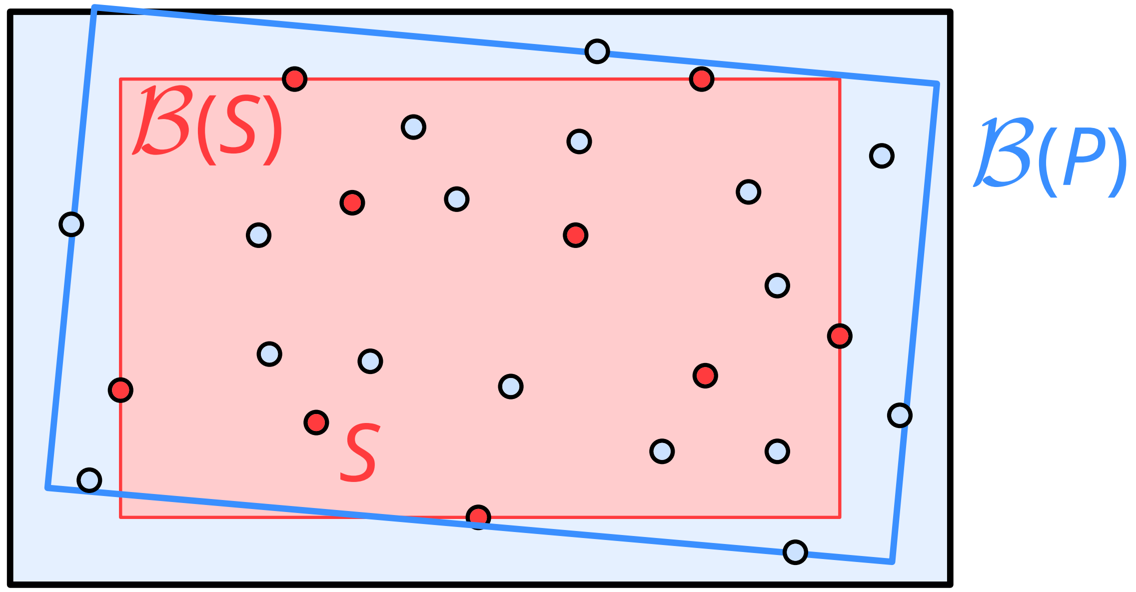
Then $P \subset B$, and $\mathrm{Vol}(B) \leq (1 + \varepsilon)\mathrm{Vol}(\mathcal{B}(P))$.



Use case: min volume bounding box

Given $\varepsilon > 0$, $P \subset \mathbb{R}^d$, let S be a δ-coreset of P for directional width ($\delta = \varepsilon/(8d)$). Let $\mathcal{B}(P)$ (resp. $\mathcal{B}(S)$) be the min volume bounding box of P (resp. S), and let B be $\mathcal{B}(S)$ scaled by $(1 + 3\delta)$ around the center of $\mathcal{B}(S)$.

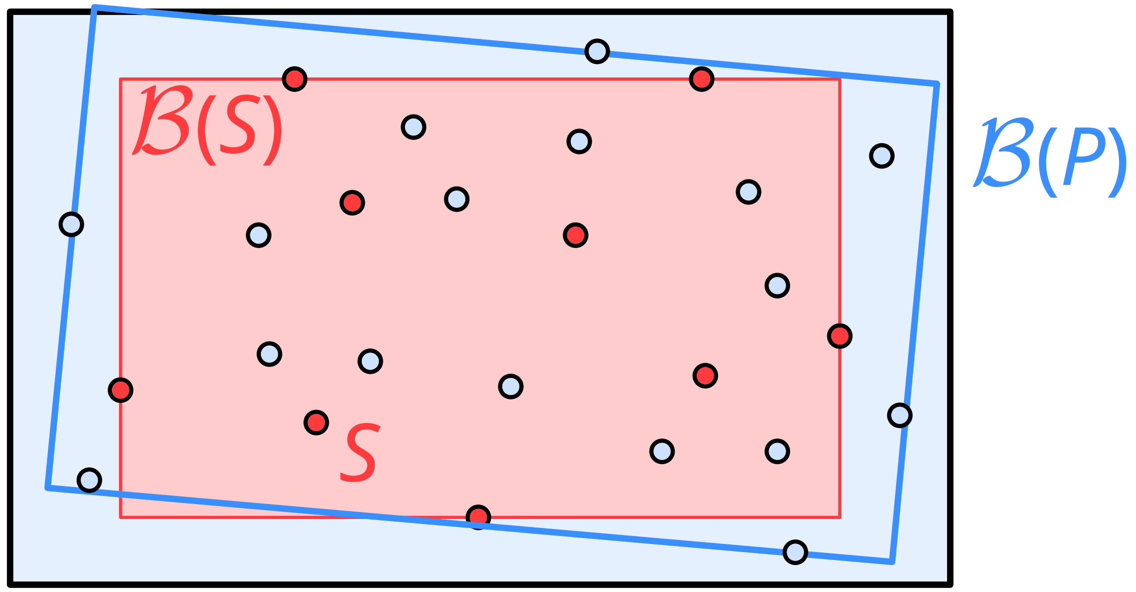
Then $P \subset B$, and $\mathrm{Vol}(B) \leq (1 + \varepsilon)\mathrm{Vol}(\mathcal{B}(P))$.



Proof.

Use case: min volume bounding box

Given $\varepsilon > 0$, $P \subset \mathbb{R}^d$, let S be a δ-coreset of P for directional width ($\delta = \varepsilon/(8d)$). Let $\mathcal{B}(P)$ (resp. $\mathcal{B}(S)$) be the min volume bounding box of P (resp. S), and let B be $\mathcal{B}(S)$ scaled by $(1 + 3\delta)$ around the center of $\mathcal{B}(S)$.



Then $P \subset B$, and $\text{Vol}(B) \leq (1 + \varepsilon)\text{Vol}(\mathcal{B}(P))$.

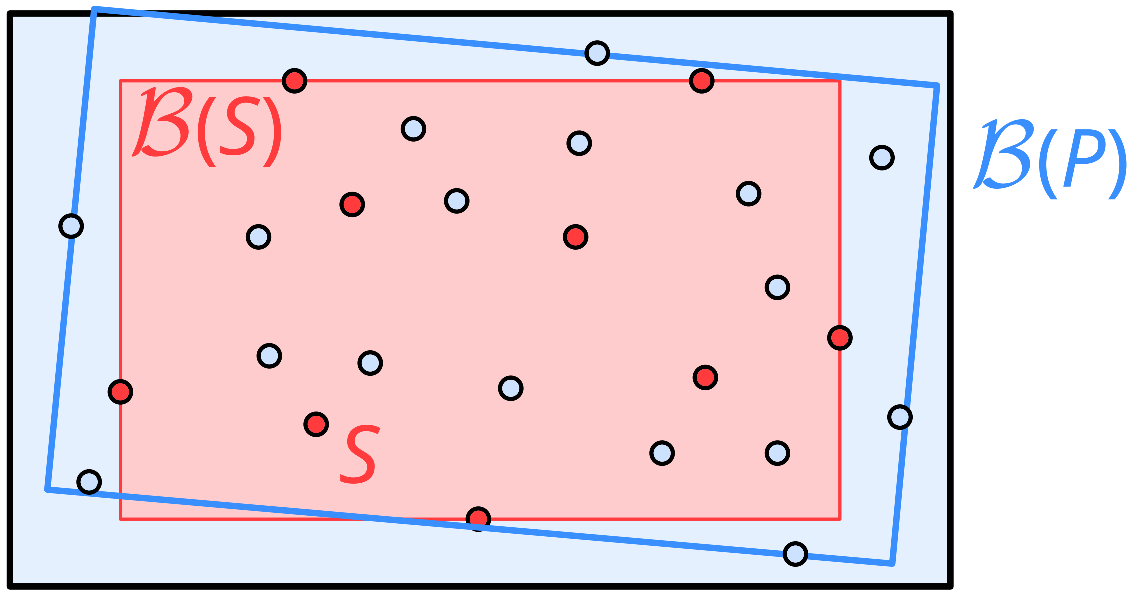
Proof.

Volume claim: $(1 + 3\delta)^d < (1 + \varepsilon)$

Use case: min volume bounding box

Given $\varepsilon > 0$, $P \subset \mathbb{R}^d$, let S be a δ-coreset of P for directional width ($\delta = \varepsilon/(8d)$). Let $\mathcal{B}(P)$ (resp. $\mathcal{B}(S)$) be the min volume bounding box of P (resp. S), and let B be $\mathcal{B}(S)$ scaled by $(1 + 3\delta)$ around the center of $\mathcal{B}(S)$.

Then $P \subset B$, and $\text{Vol}(B) \leq (1 + \varepsilon)\text{Vol}(\mathcal{B}(P))$.



Proof.

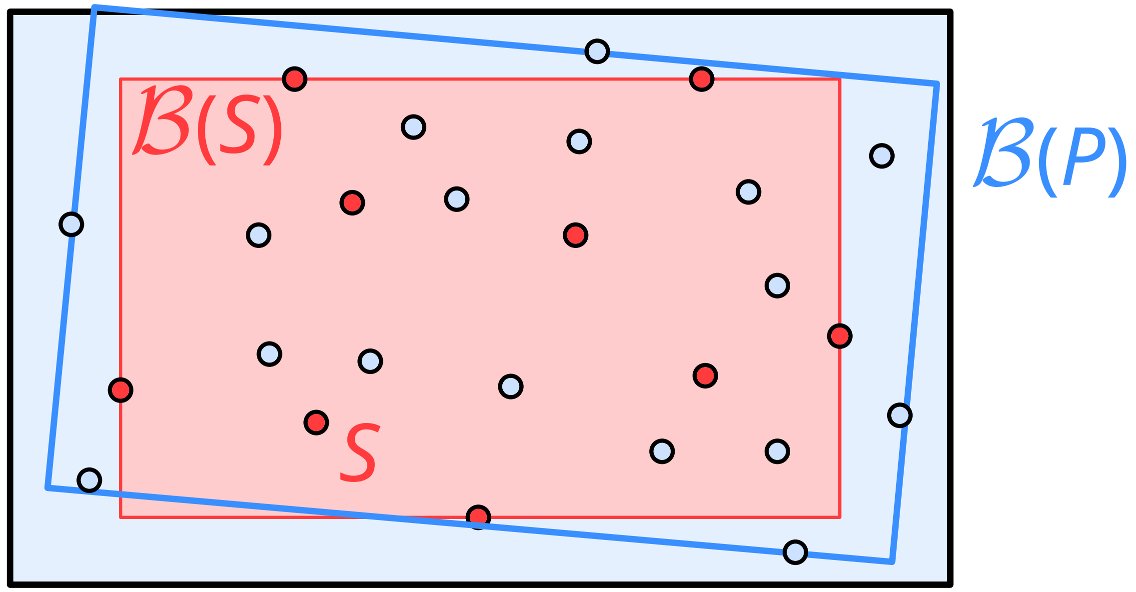
Volume claim: $(1 + 3\delta)^d < (1 + \varepsilon)$

$$(1 + 3\delta)^d = \left(1 + \frac{3\varepsilon/8}{d}\right)^d \leq \exp(3\varepsilon/8)$$

Use case: min volume bounding box

Given $\varepsilon > 0$, $P \subset \mathbb{R}^d$, let S be a δ-coreset of P for directional width ($\delta = \varepsilon/(8d)$). Let $\mathcal{B}(P)$ (resp. $\mathcal{B}(S)$) be the min volume bounding box of P (resp. S), and let B be $\mathcal{B}(S)$ scaled by $(1 + 3\delta)$ around the center of $\mathcal{B}(S)$.

Then $P \subset B$, and $\mathrm{Vol}(B) \leq (1 + \varepsilon)\mathrm{Vol}(\mathcal{B}(P))$.



Proof.

Volume claim: $(1 + 3\delta)^d < (1 + \varepsilon)$

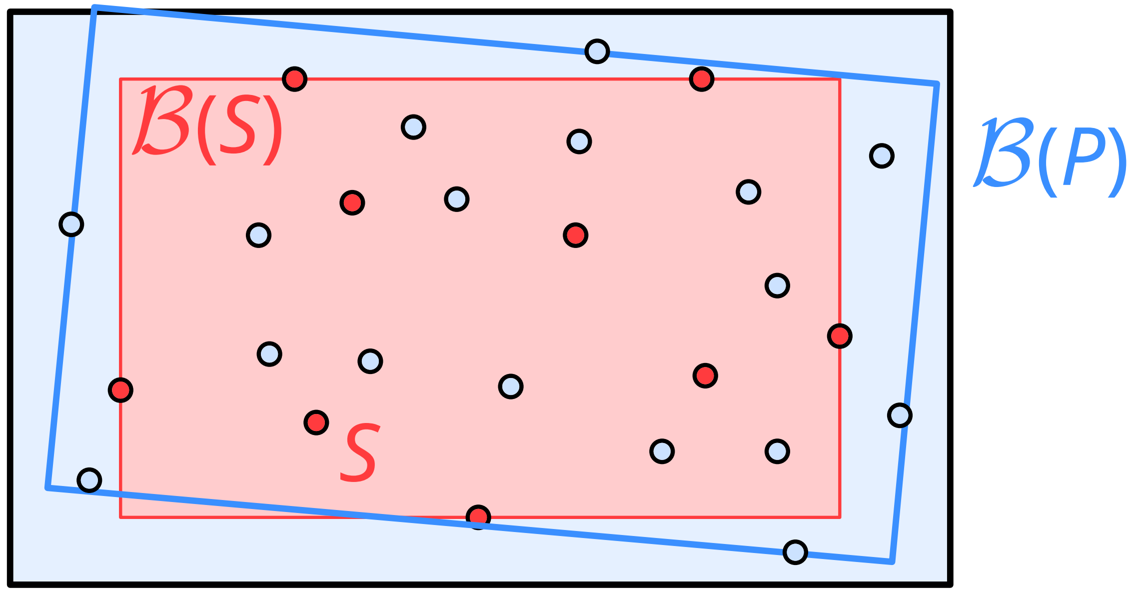
$$(1 + 3\delta)^d = \left(1 + \frac{3\varepsilon/8}{d}\right)^d \leq \exp(3\varepsilon/8)$$

$\leq 1 + 6\varepsilon/8 < 1 + \varepsilon$ $\qquad$ ($\exp(x) \leq 1 + 2x$ for $0 \leq x \leq 1$)

Use case: min volume bounding box

Given $\varepsilon > 0$, $P \subset \mathbb{R}^d$, let S be a δ-coreset of P for directional width ($\delta = \varepsilon/(8d)$). Let $\mathcal{B}(P)$ (resp. $\mathcal{B}(S)$) be the min volume bounding box of P (resp. S), and let B be $\mathcal{B}(S)$ scaled by $(1 + 3\delta)$ around the center of $\mathcal{B}(S)$.

Then $P \subset B$, and $\mathrm{Vol}(B) \leq (1 + \varepsilon)\mathrm{Vol}(\mathcal{B}(P))$.



Proof.

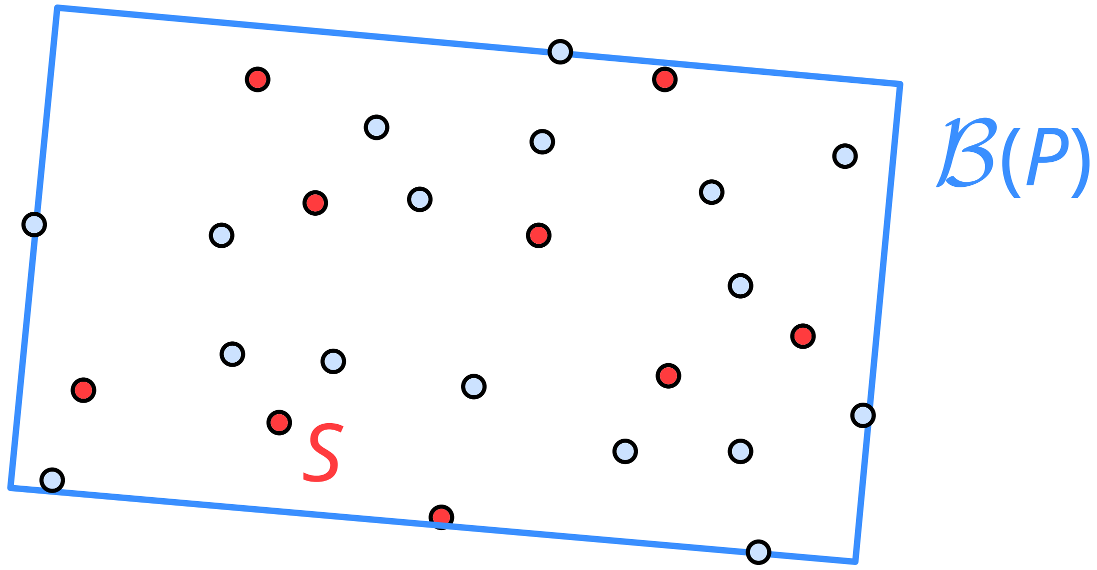
Volume claim: $(1 + 3\delta)^d < (1 + \varepsilon)$

$$(1 + 3\delta)^d = \left(1 + \frac{3\varepsilon/8}{d}\right)^d \leq \exp(3\varepsilon/8)$$

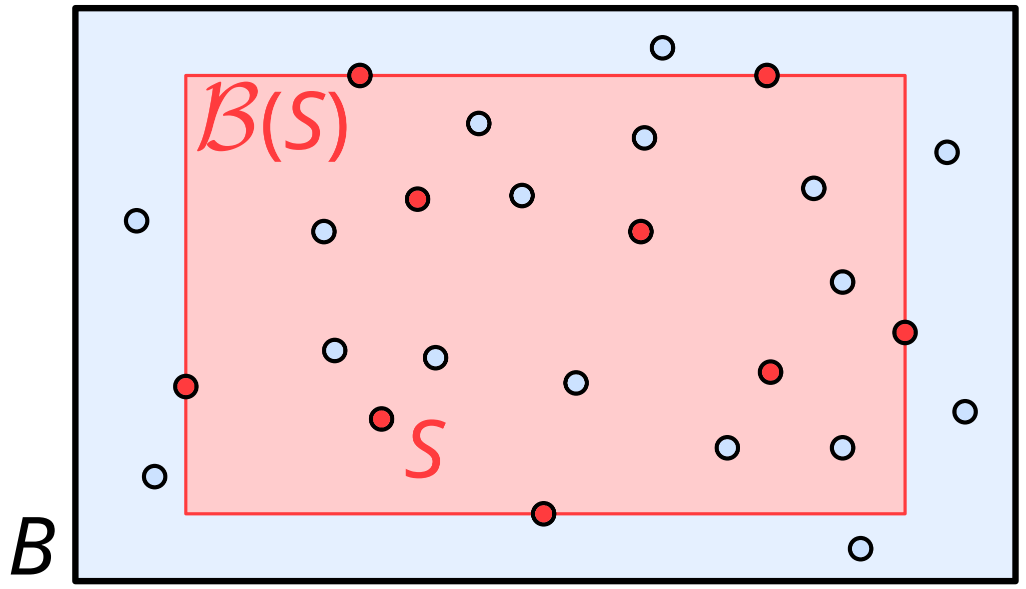
$\leq 1 + 6\varepsilon/8 < 1 + \varepsilon$ $\qquad$ ($\exp(x) \leq 1 + 2x$ for $0 \leq x \leq 1$)

Still need: B contains P

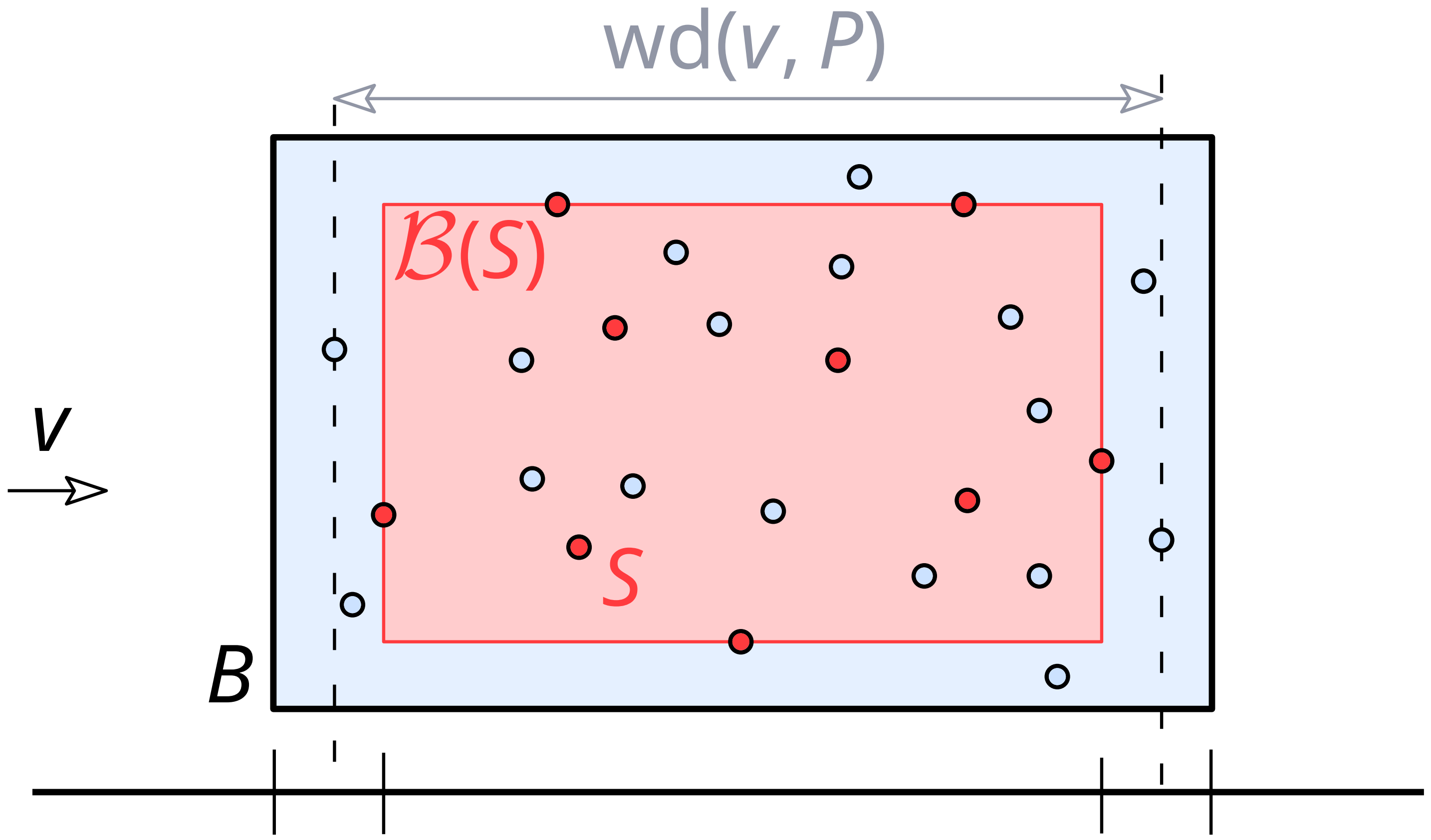
Proof: $P \subset B$



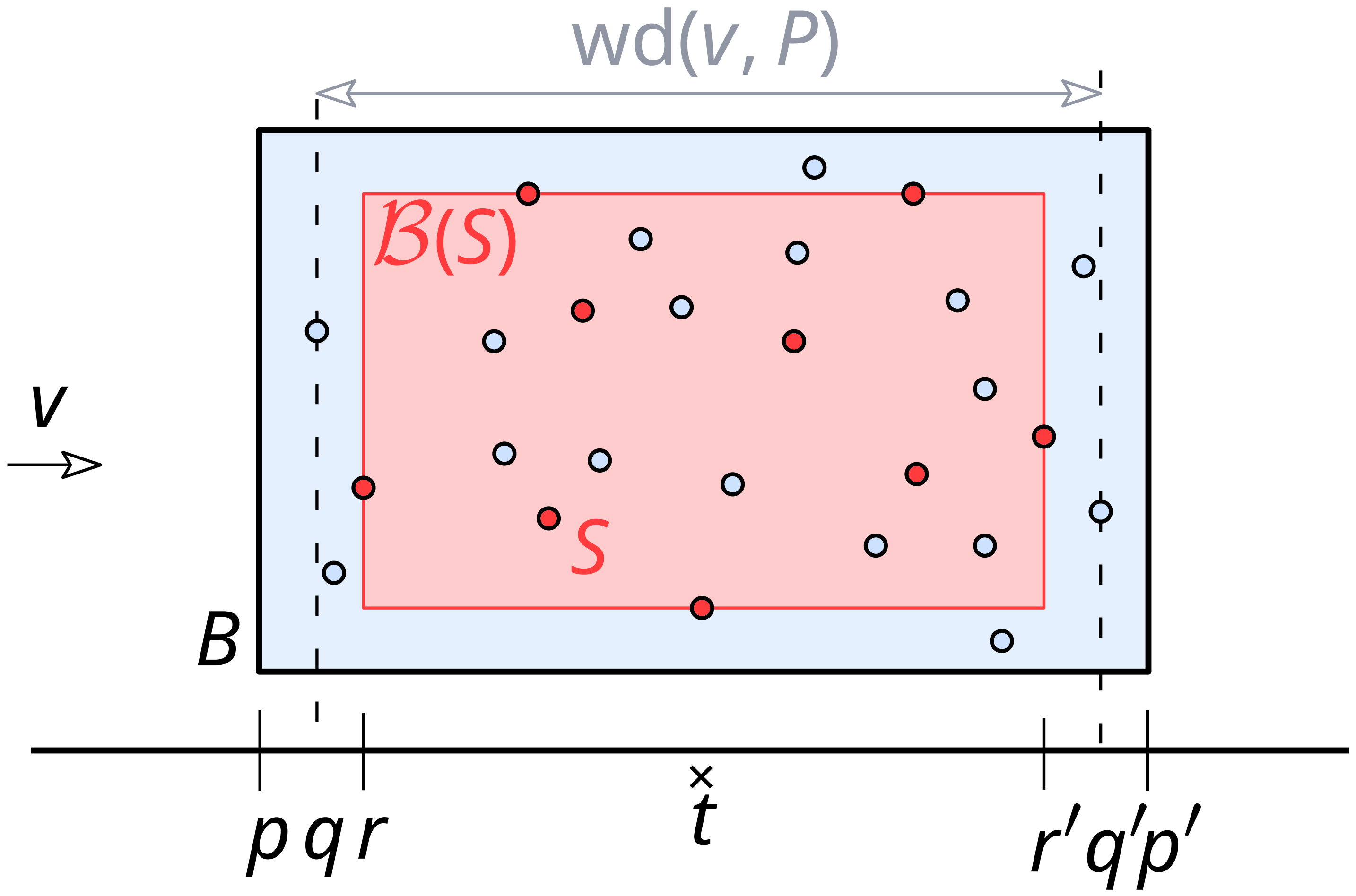
Proof: $P \subset B$



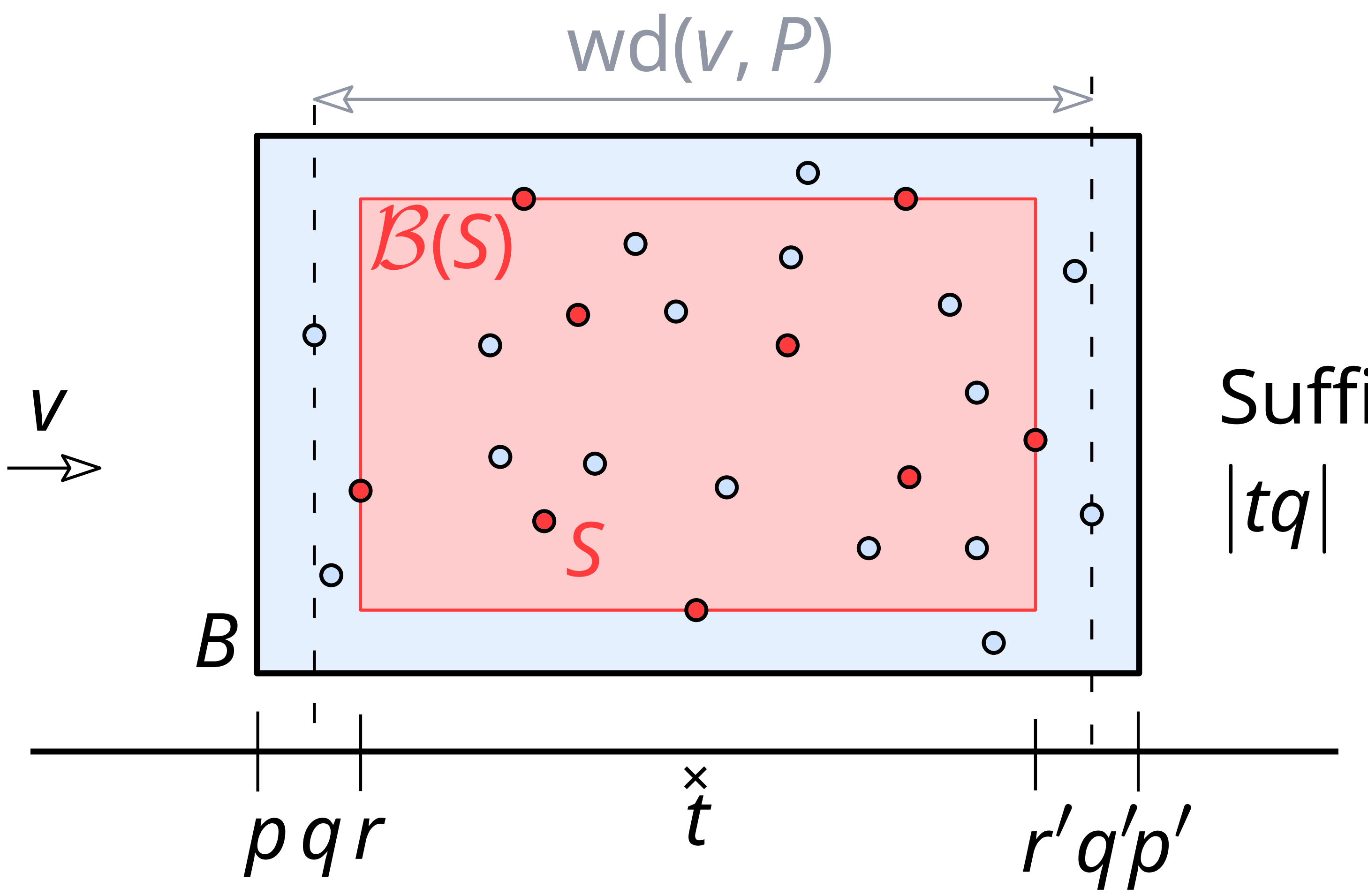
Proof: $P \subset B$



Proof: $P \subset B$

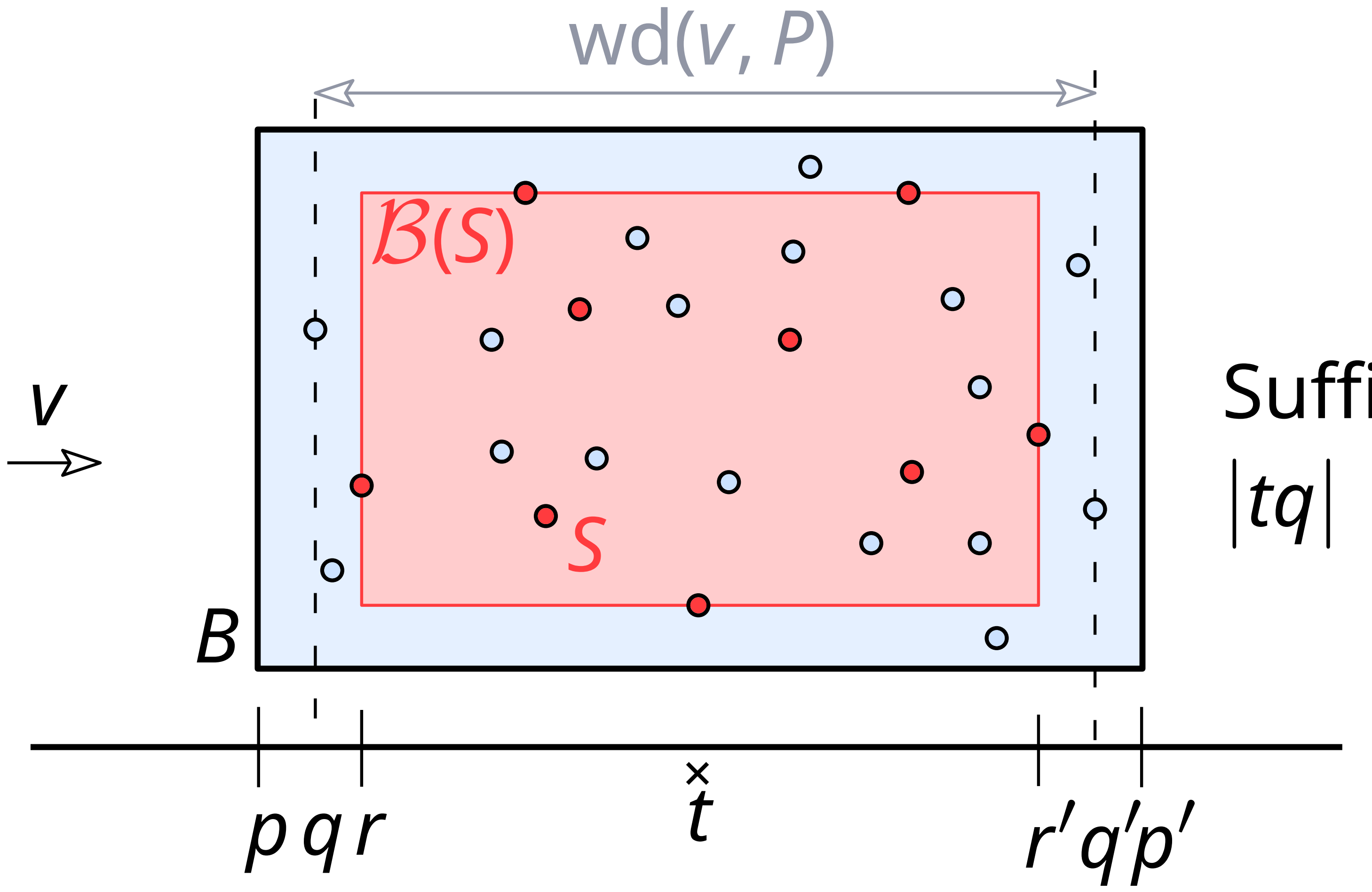


Proof: $P \subset B$



Sufficient to prove for each “axis” v of B:
$|tq| \leq |tp|$ (and $|tq'| \leq |tp'|$)

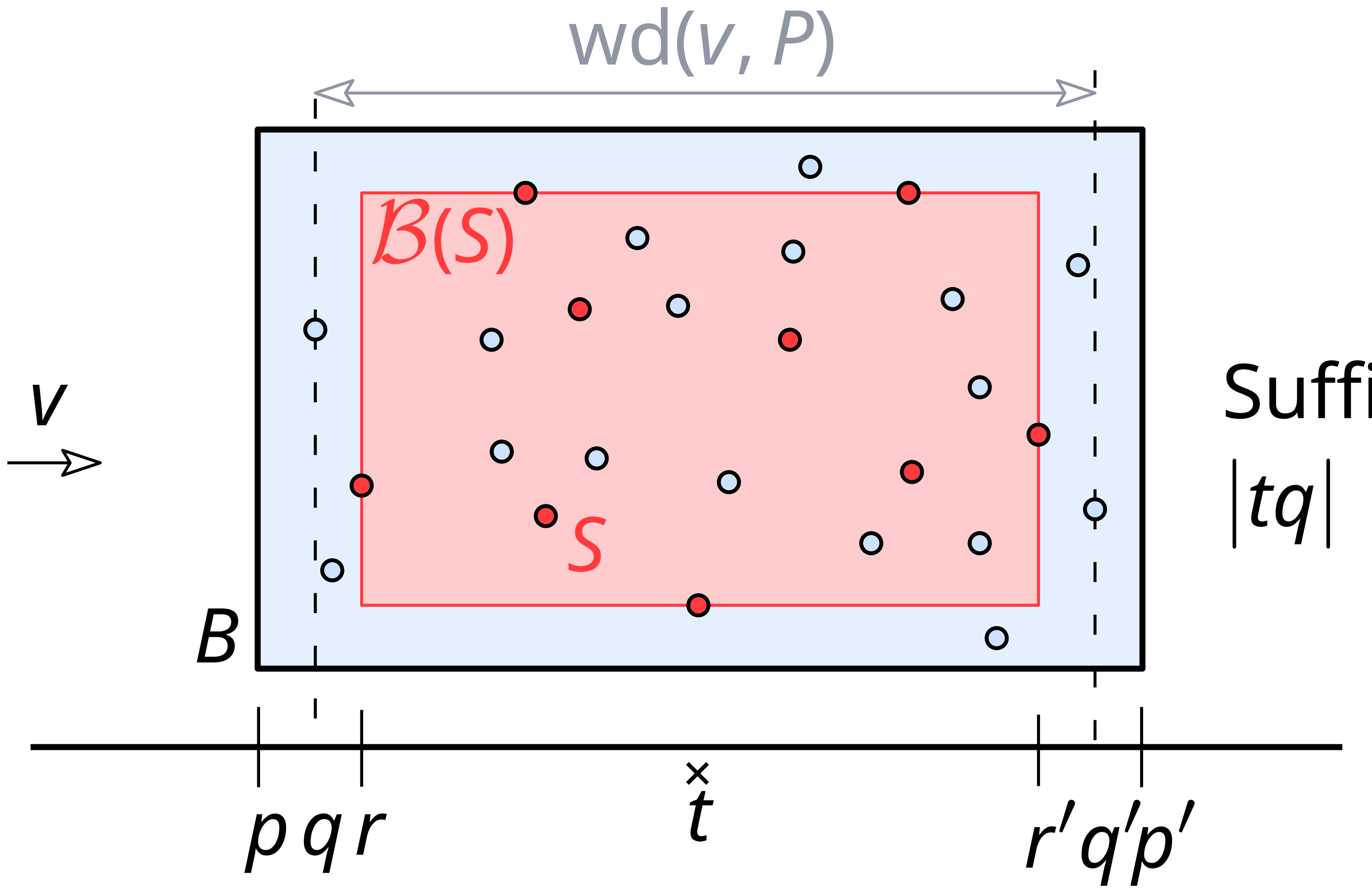
Proof: $P \subset B$



Sufficient to prove for each "axis" v of B:
$|tq| \leq |tp|$ (and $|tq'| \leq |tp'|$)

$(1-\delta)|qq'| \leq |rr'| = 2|tr|$ as S is δ-coreset.

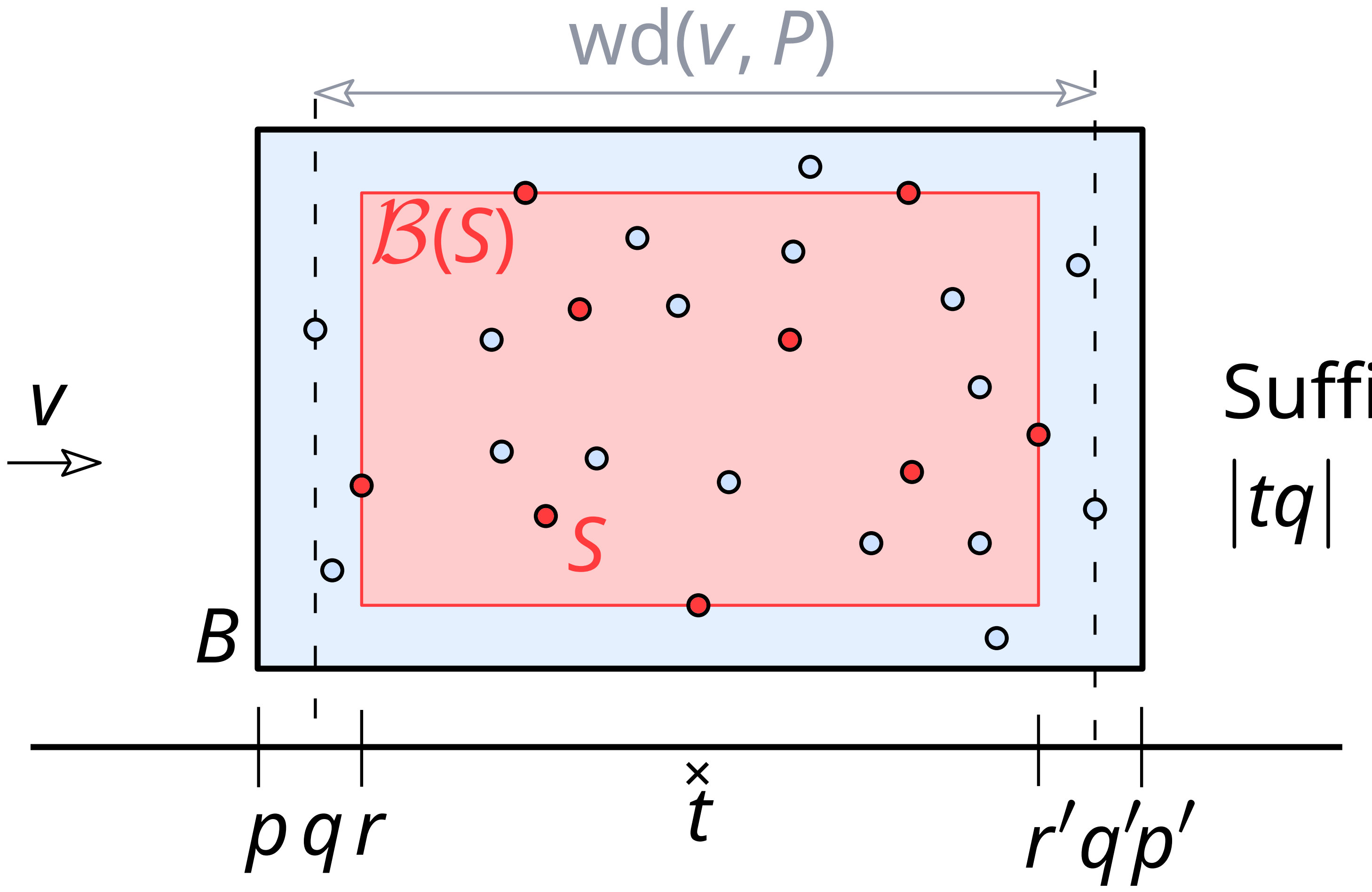
Proof: $P \subset B$



Sufficient to prove for each "axis" v of B:
$|tq| \leq |tp|$ (and $|tq'| \leq |tp'|$)

$(1-\delta)|qq'| \leq |rr'| = 2|tr|$ as S is δ-coreset. $\Rightarrow |qq'| - |rr'| \leq \delta|qq'|$

Proof: $P \subset B$

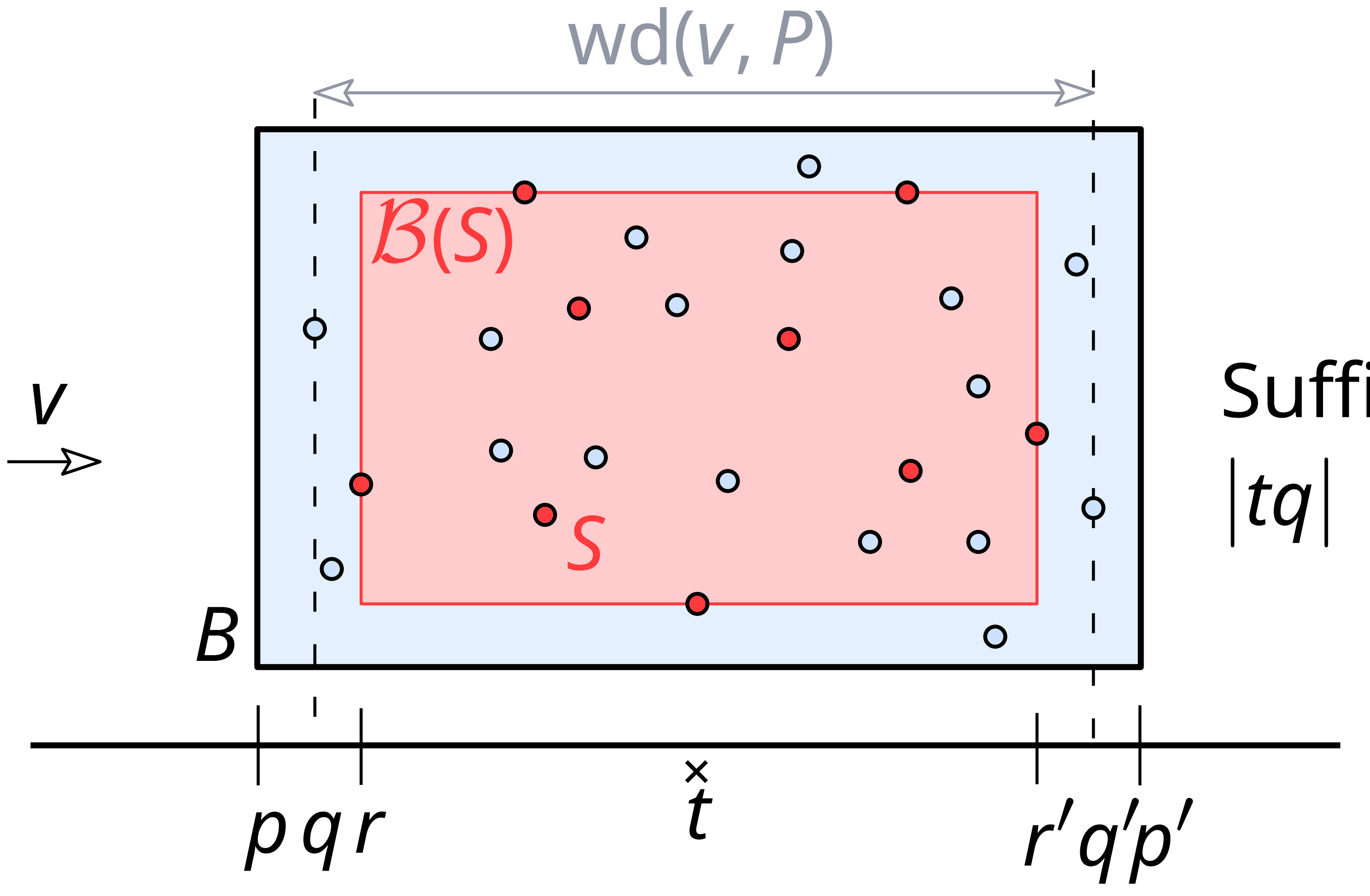


Sufficient to prove for each "axis" v of B:
$|tq| \leq |tp|$ (and $|tq'| \leq |tp'|$)

$(1-\delta)|qq'| \leq |rr'| = 2|tr|$ as S is δ-coreset. $\Rightarrow |qq'| - |rr'| \leq \delta|qq'|$

$$|tq| \leq |tr| + \delta|qq'| \leq \left(1 + \frac{2\delta}{1-\delta}\right)|tr| \leq (1+3\delta)|tr| = |tp|$$

Proof: $P \subset B$



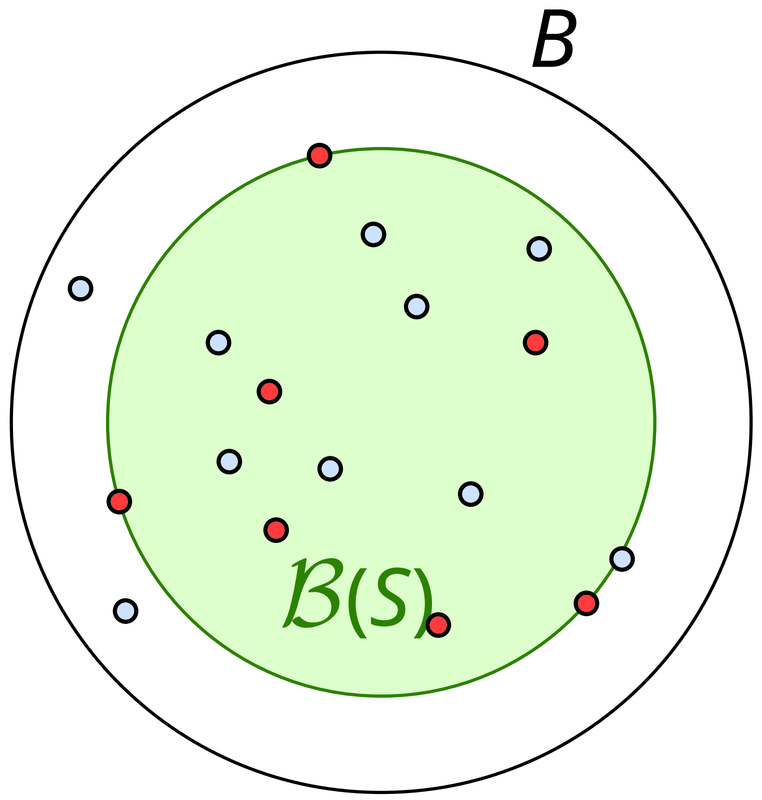
Sufficient to prove for each "axis" v of B:
$|tq| \leq |tp|$ (and $|tq'| \leq |tp'|$)

$(1-\delta)|qq'| \leq |rr'| = 2|tr|$ as S is δ-coreset. $\Rightarrow |qq'| - |rr'| \leq \delta|qq'|$

$$|tq| \leq |tr| + \delta|qq'| \leq \left(1 + \frac{2\delta}{1-\delta}\right)|tr| \leq (1+3\delta)|tr| = |tp| \qquad \square$$

Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1+\varepsilon)$ around its center, B, contains P.

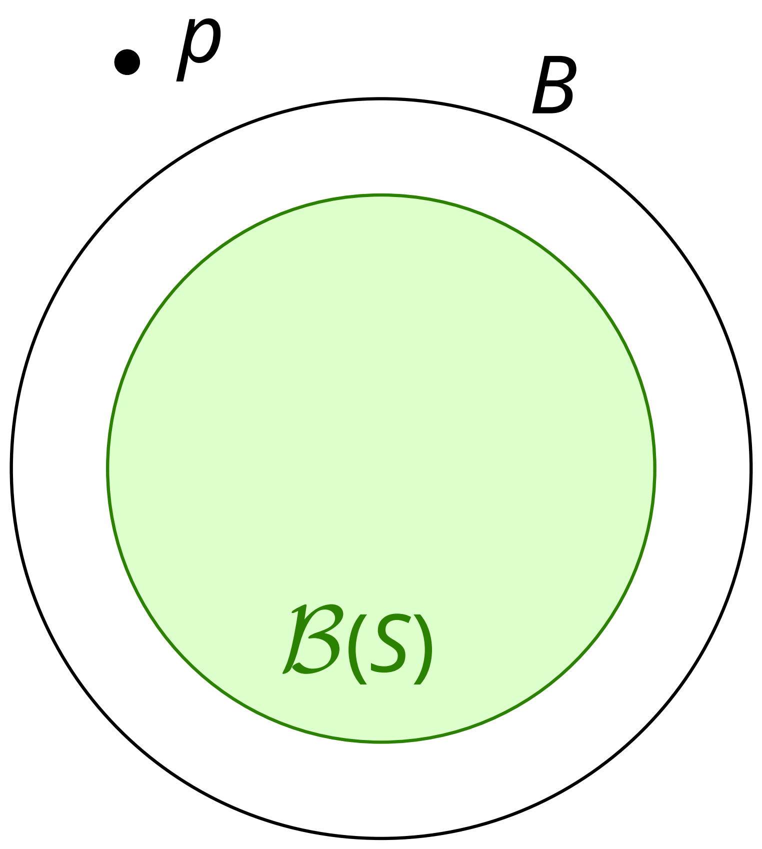


Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.

Proof by contradiction:

Suppose $p \in P$ is outside B

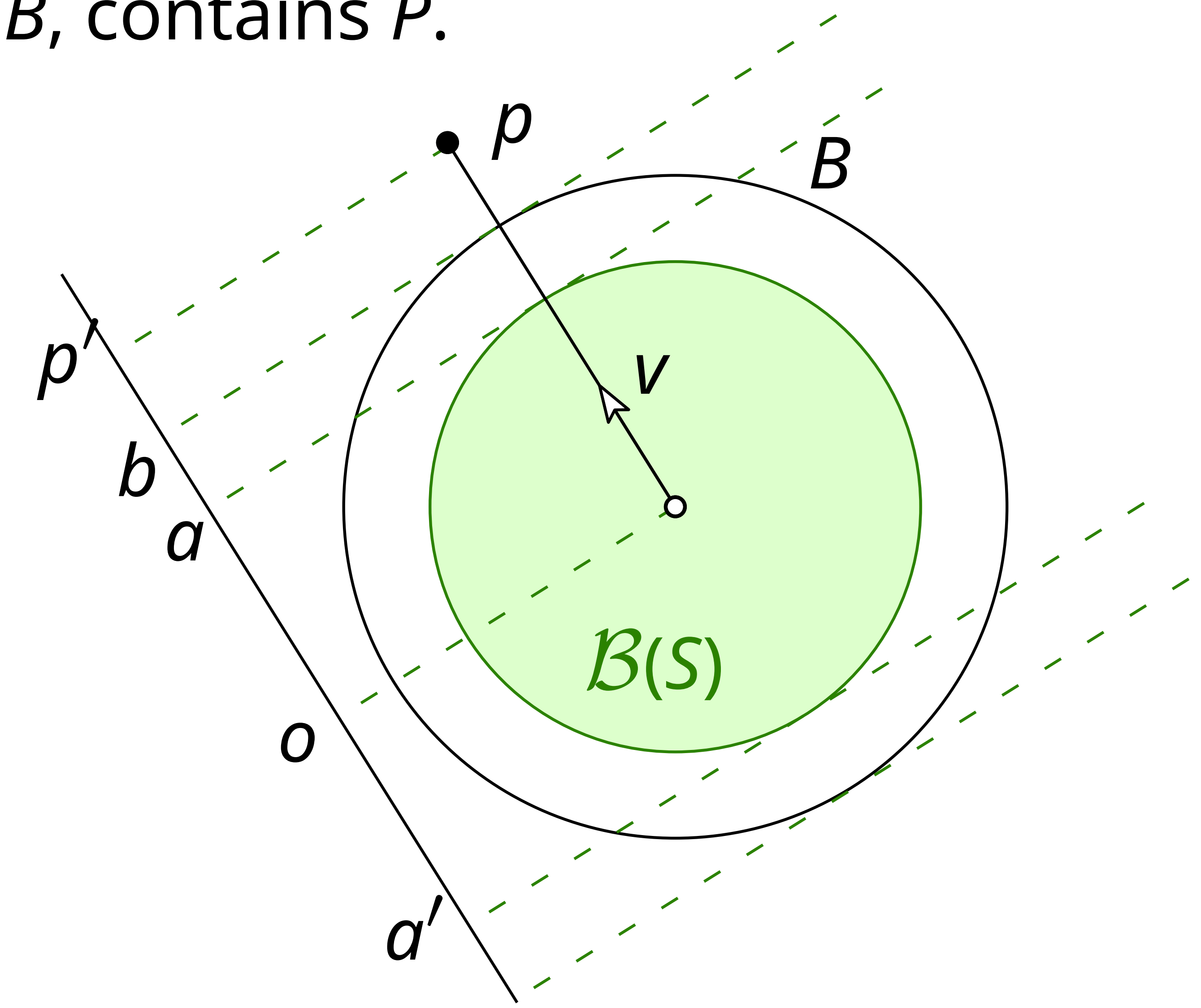


Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.

Proof by contradiction:

Suppose $p \in P$ is outside B



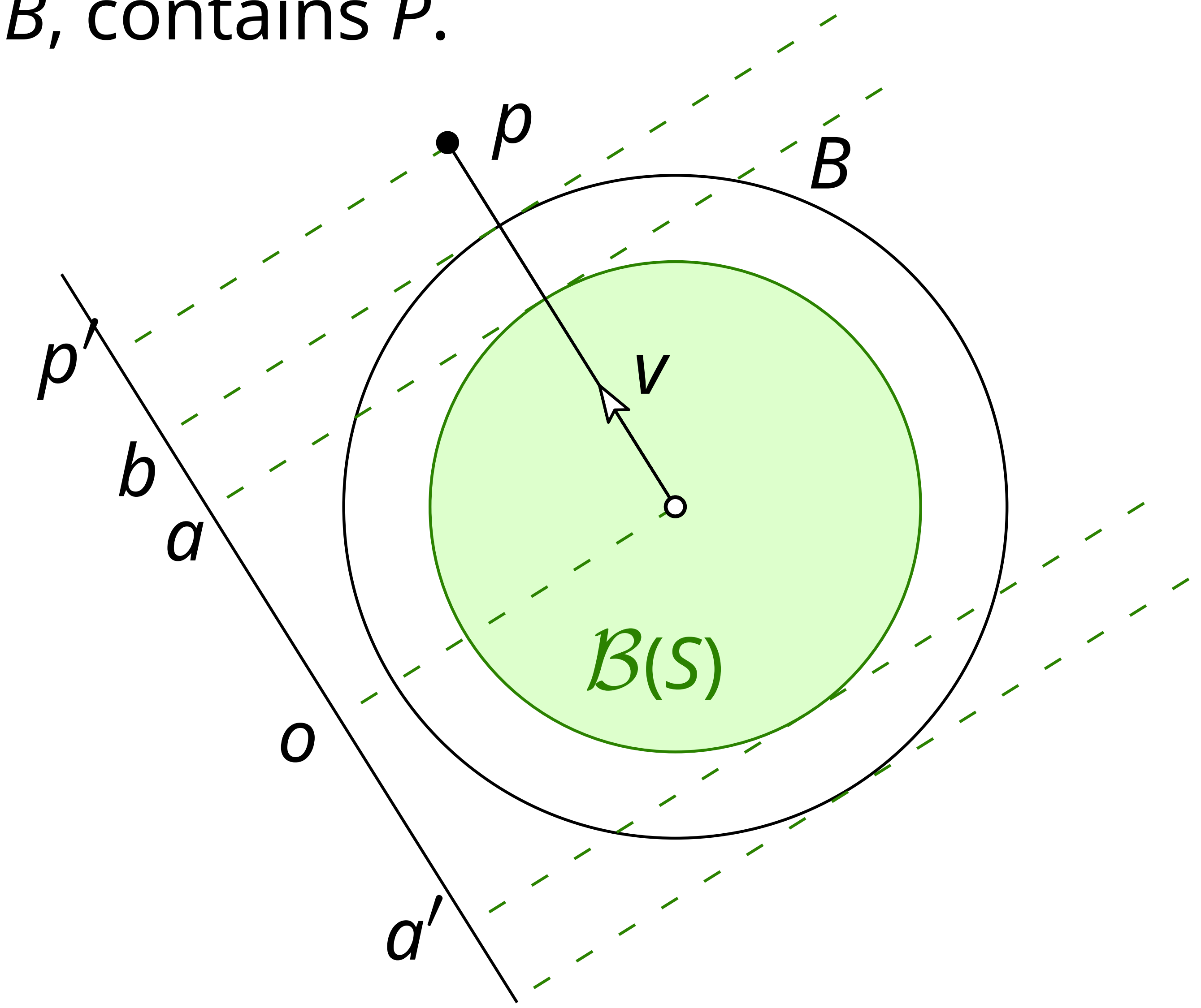
Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.

Proof by contradiction:

Suppose $p \in P$ is outside B

$\Rightarrow |op'| > |ob| = (1 + \varepsilon)|oa|$



Use Case: minimum enclosing ball

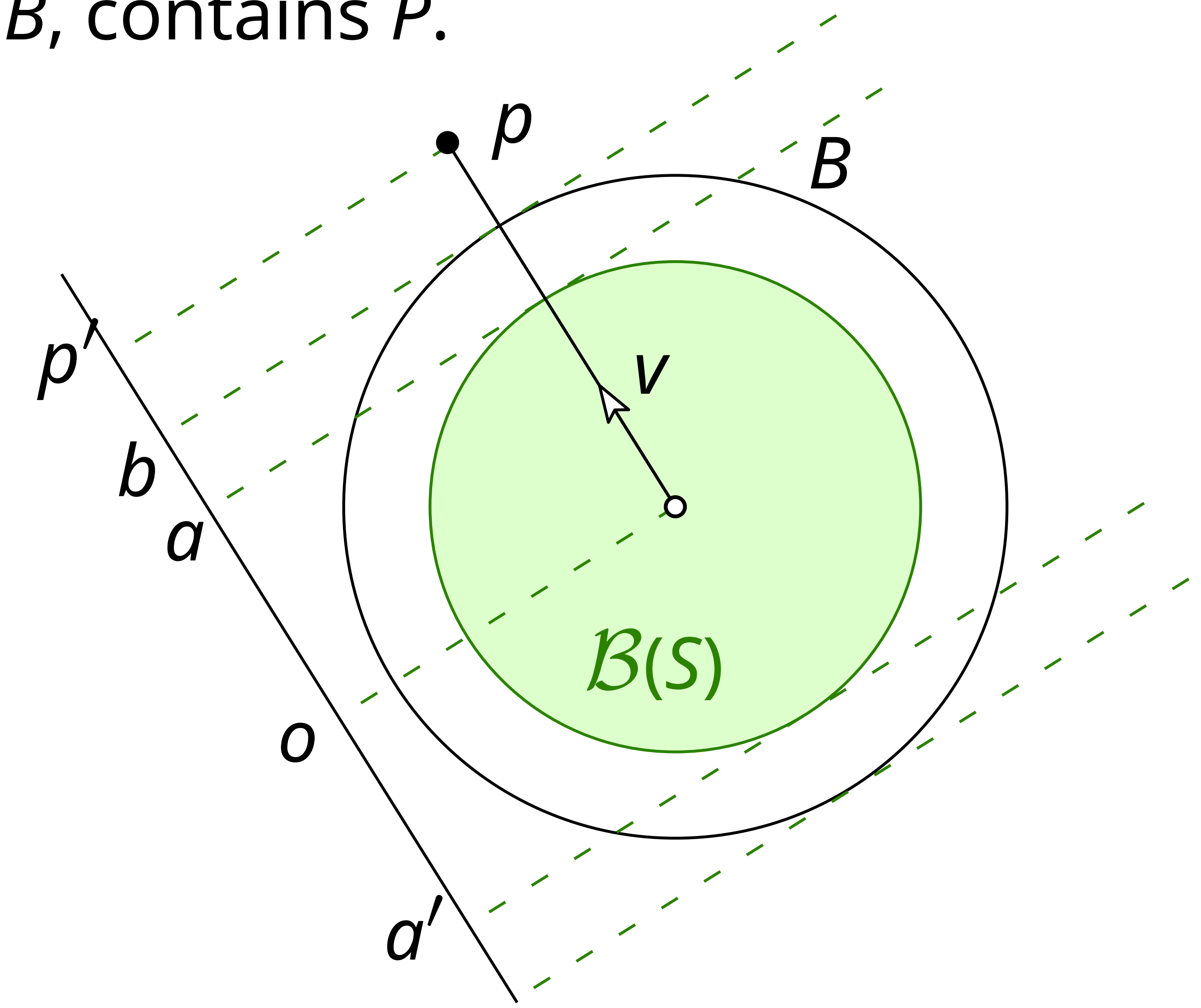
If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.

Proof by contradiction:

Suppose $p \in P$ is outside B

$\Rightarrow |op'| > |ob| = (1 + \varepsilon)|oa|$

$\text{wd}(v, P) - \text{wd}(v, S) \geq |p'a|$



Use Case: minimum enclosing ball

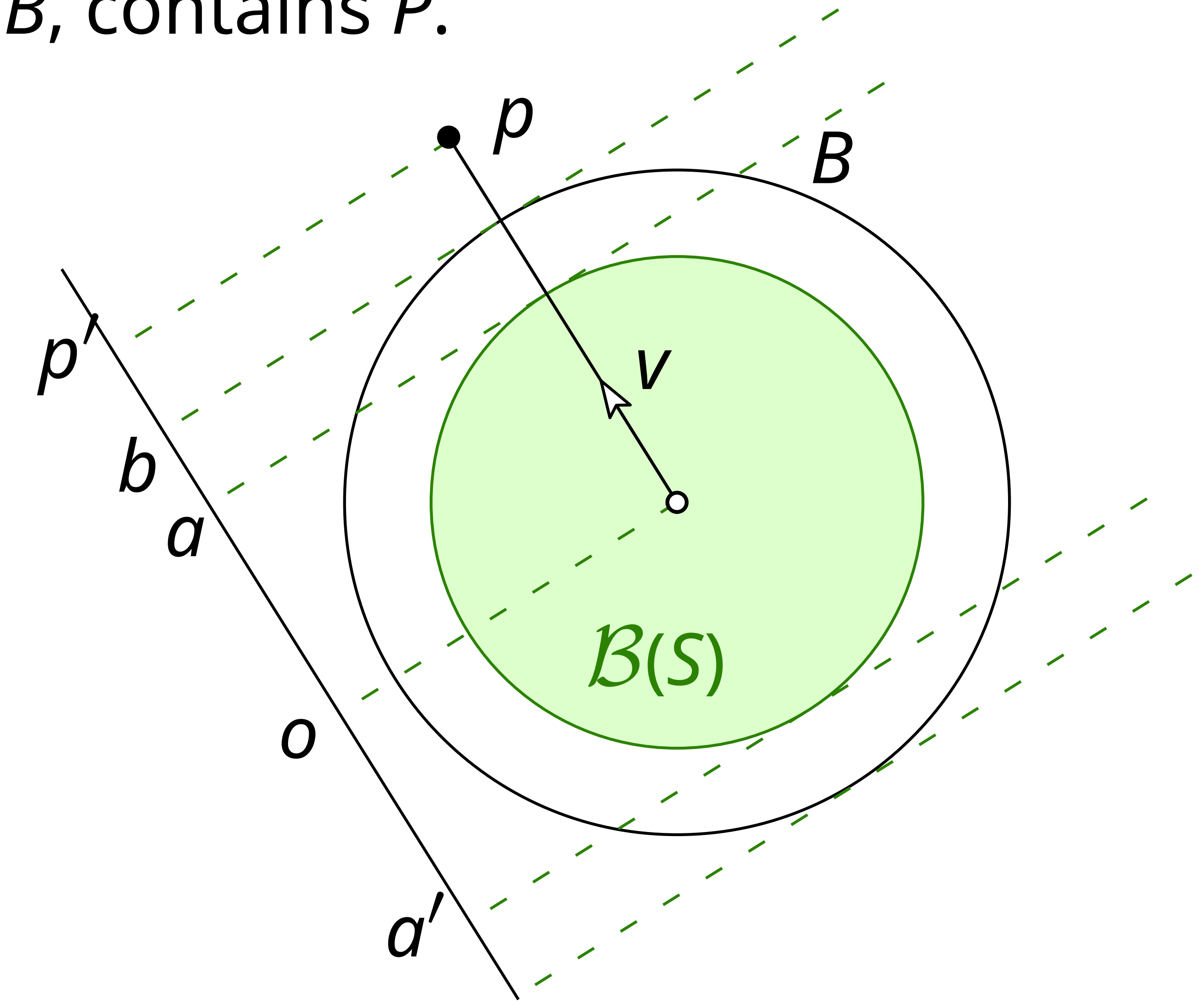
If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1+\varepsilon)$ around its center, B, contains P.

Proof by contradiction:

Suppose $p \in P$ is outside B

$\Rightarrow |op'| > |ob| = (1+\varepsilon)|oa|$

$$\begin{aligned} \mathrm{wd}(v, P) - \mathrm{wd}(v, S) &\geq |p'a| \\ &\geq \varepsilon\frac{|aa'|}{2} \geq \frac{\varepsilon}{2}\mathrm{wd}(v, S) \end{aligned}$$



Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1+\varepsilon)$ around its center, B, contains P.

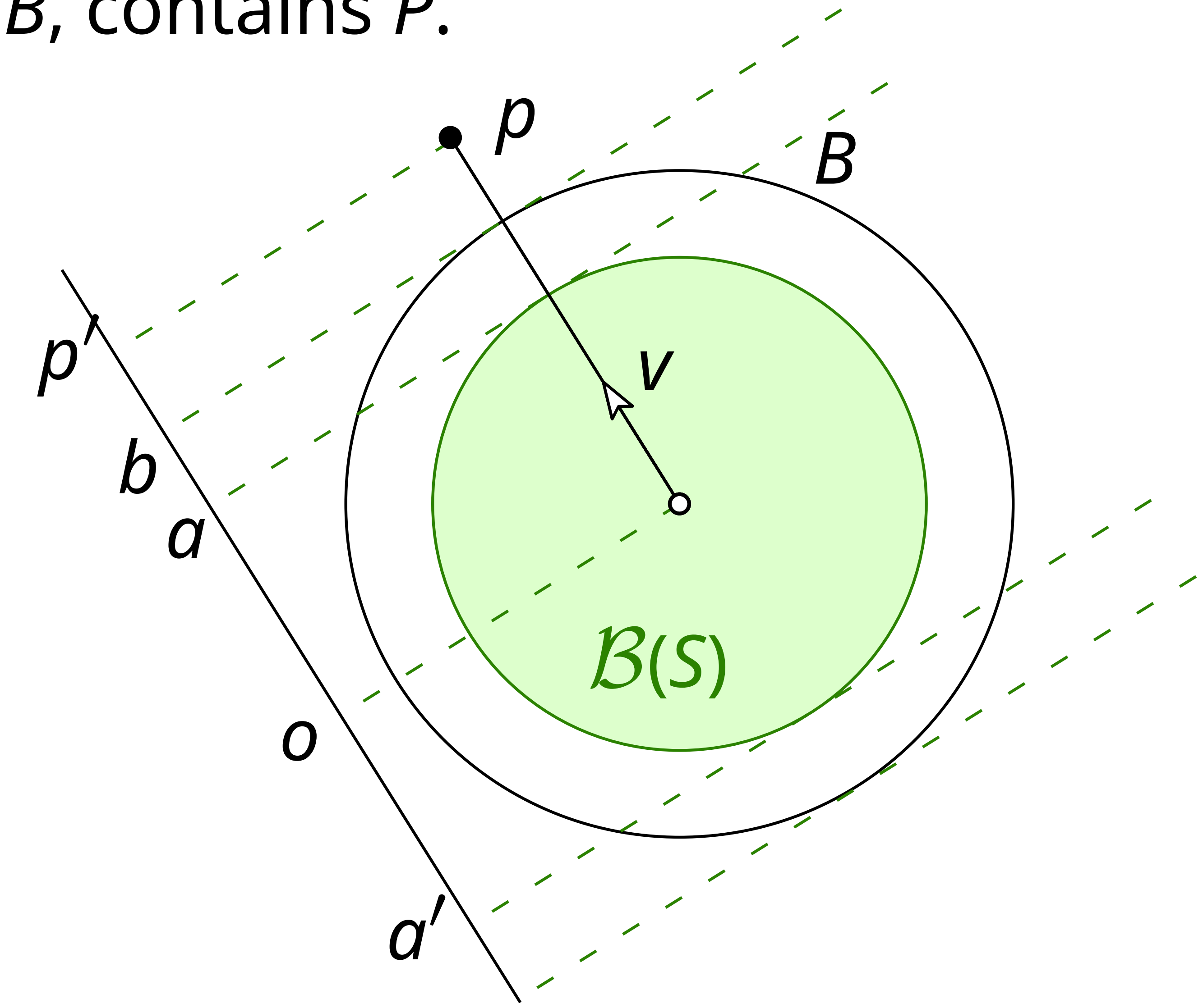
Proof by contradiction:

Suppose $p \in P$ is outside B

$\Rightarrow |op'| > |ob| = (1+\varepsilon)|oa|$

$$\mathrm{wd}(v,P) - \mathrm{wd}(v,S) \geq |p'a| \geq \varepsilon\frac{|aa'|}{2} \geq \frac{\varepsilon}{2}\mathrm{wd}(v,S)$$

$$\mathrm{wd}(v,S) \leq \frac{1}{1+\varepsilon/2}\mathrm{wd}(v,P)$$



Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1+\varepsilon)$ around its center, B, contains P.

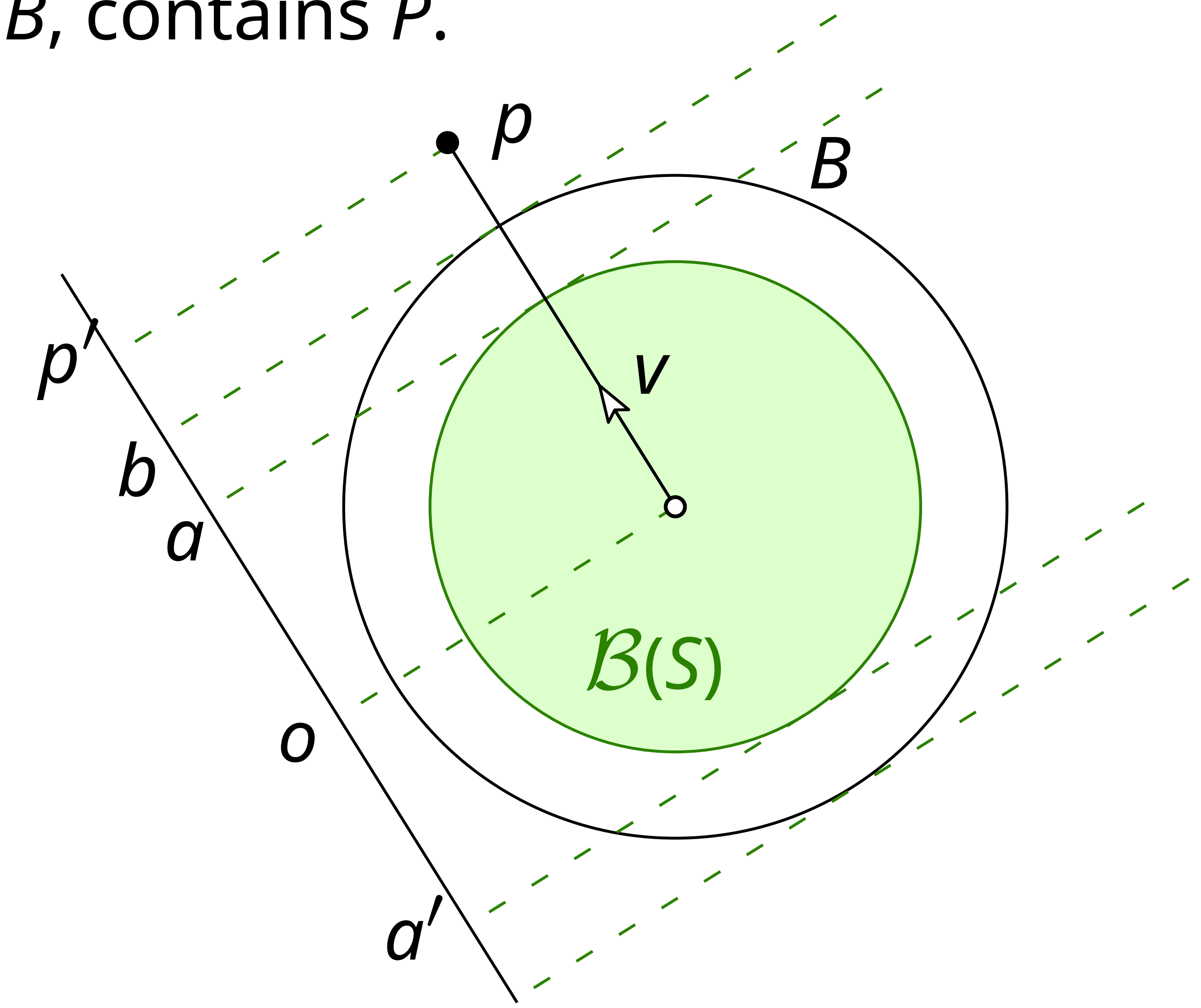
Proof by contradiction:

Suppose $p \in P$ is outside B

$\Rightarrow |op'| > |ob| = (1+\varepsilon)|oa|$

$$\mathrm{wd}(v,P) - \mathrm{wd}(v,S) \geq |p'a| \geq \varepsilon\frac{|aa'|}{2} \geq \frac{\varepsilon}{2}\mathrm{wd}(v,S)$$

$$\mathrm{wd}(v,S) \leq \frac{1}{1+\varepsilon/2}\mathrm{wd}(v,P) = \left(1 - \frac{\varepsilon/2}{1+\varepsilon/2}\right) \underset{\varepsilon<2}{<} (1-\frac{\varepsilon}{4})\mathrm{wd}(v,P)$$



Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1+\varepsilon)$ around its center, B, contains P.

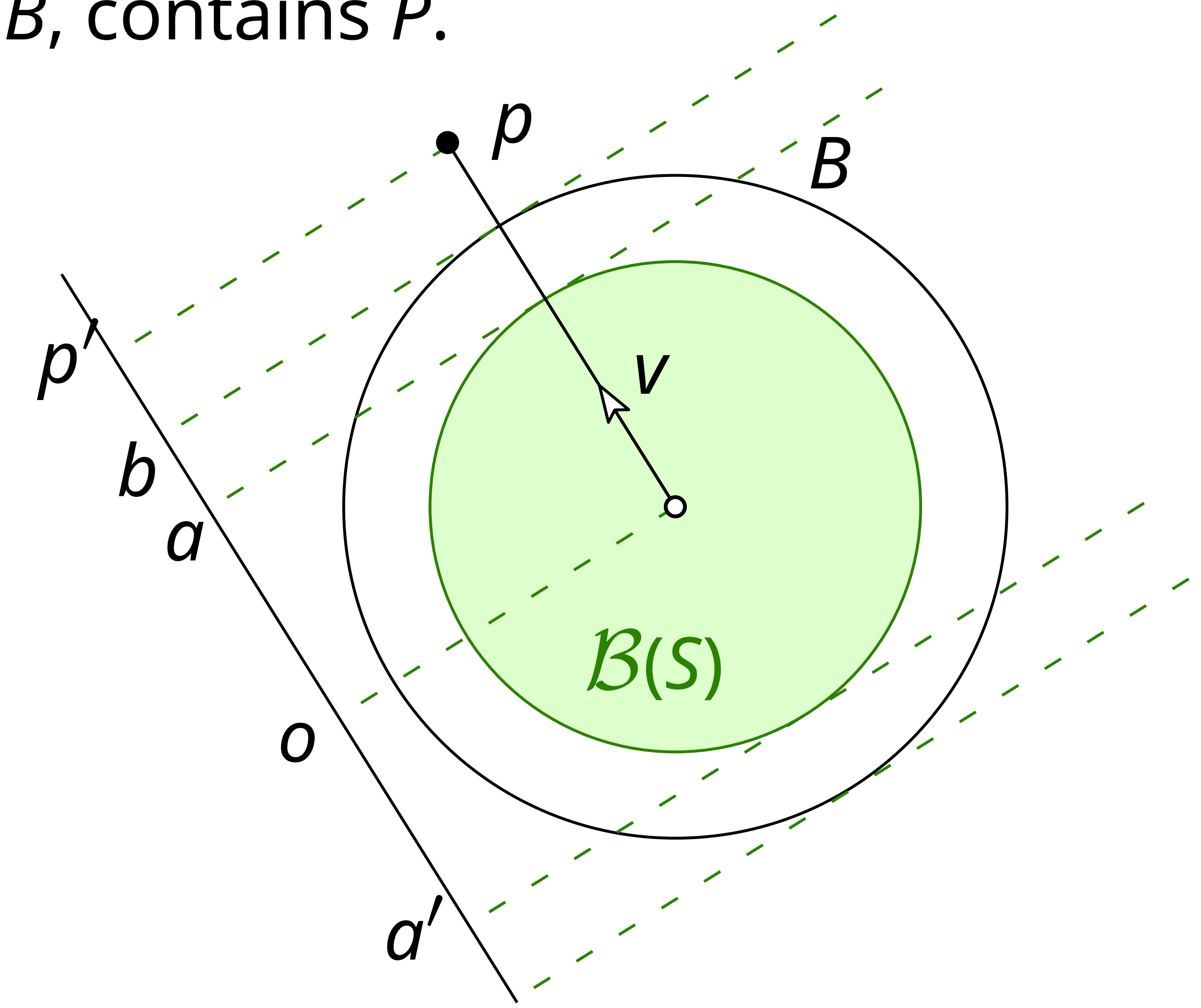
Proof by contradiction:

Suppose $p \in P$ is outside B

$\Rightarrow |op'| > |ob| = (1+\varepsilon)|oa|$

$$\mathrm{wd}(v,P) - \mathrm{wd}(v,S) \geq |p'a| \geq \varepsilon\frac{|aa'|}{2} \geq \frac{\varepsilon}{2}\mathrm{wd}(v,S)$$

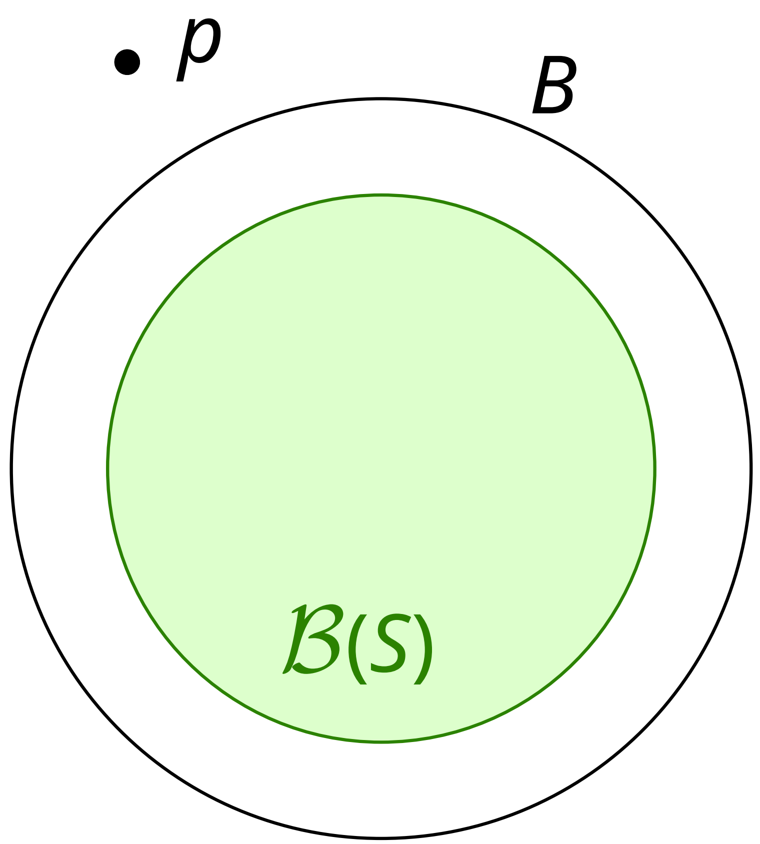
$$\mathrm{wd}(v,S) \leq \frac{1}{1+\varepsilon/2}\mathrm{wd}(v,P) = \left(1 - \frac{\varepsilon/2}{1+\varepsilon/2}\right) \underset{\varepsilon<2}{<} \left(1-\frac{\varepsilon}{4}\right)\mathrm{wd}(v,P)$$



↯ □

Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.



Comparison to ε-samples:

ε-sample:

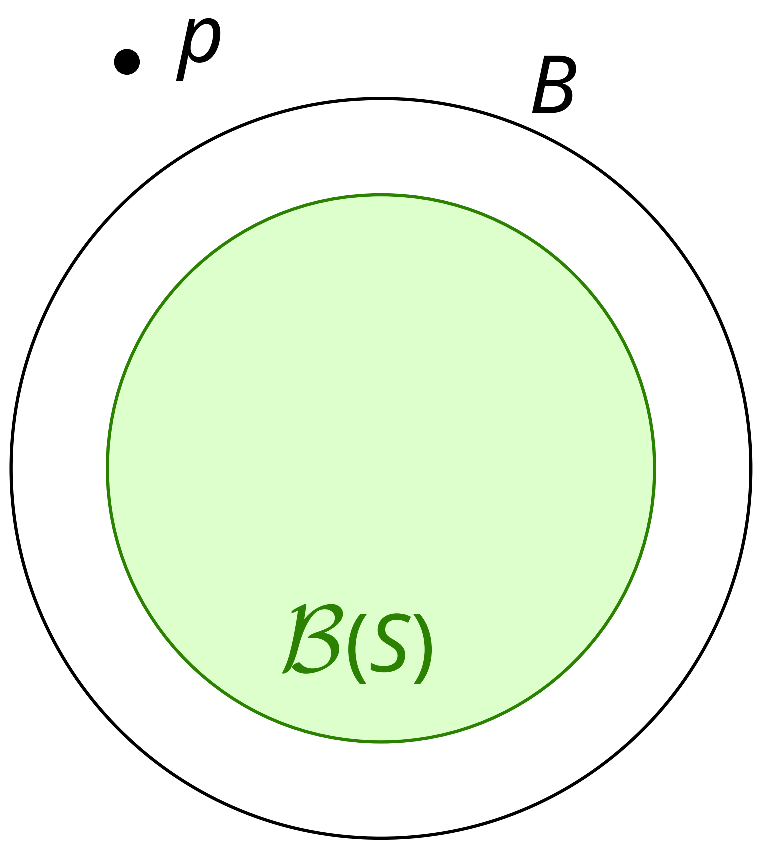
Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.

Comparison to ε-samples:

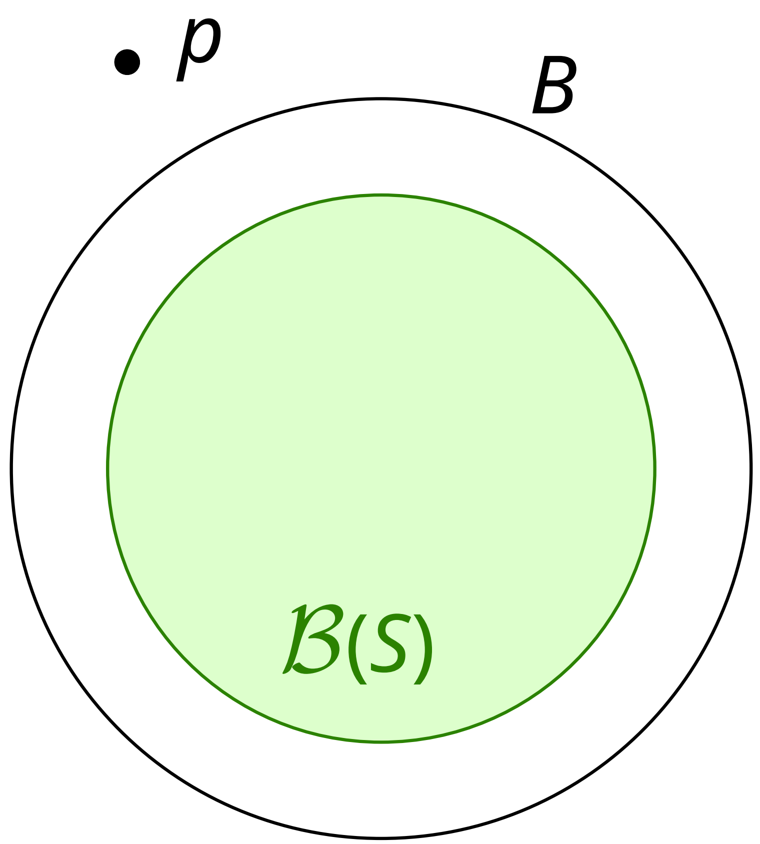
ε-sample:

guarantees most points in $\mathcal{B}(S)$



Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.



Comparison to ε-samples:

ε-sample:

guarantees most points in $\mathcal{B}(S)$

combinatorial/statistical error

Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.

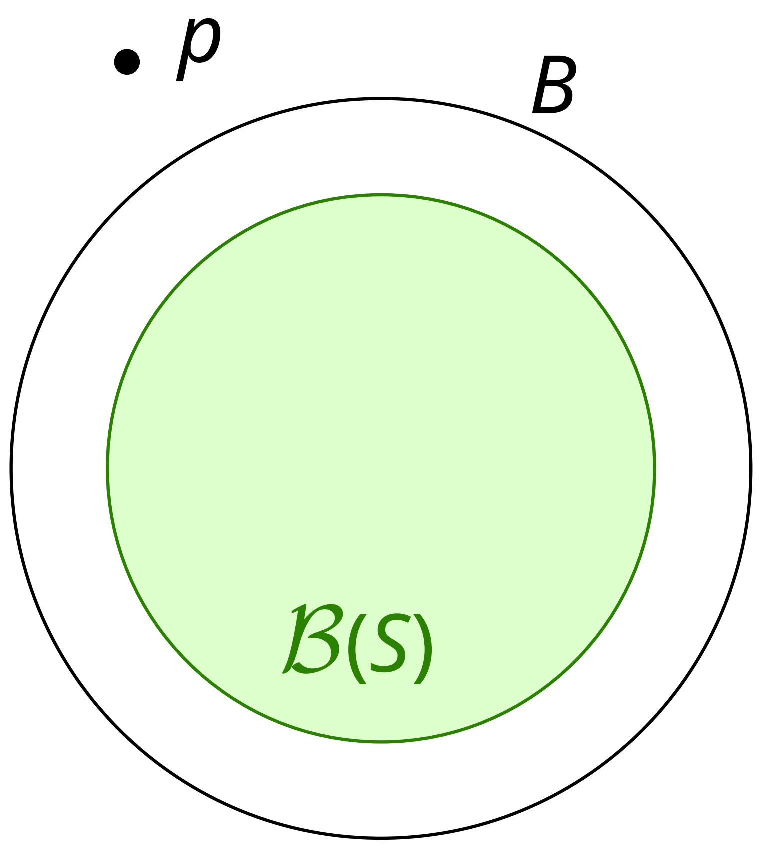
Comparison to ε-samples:

ε-sample:

guarantees most points in $\mathcal{B}(S)$

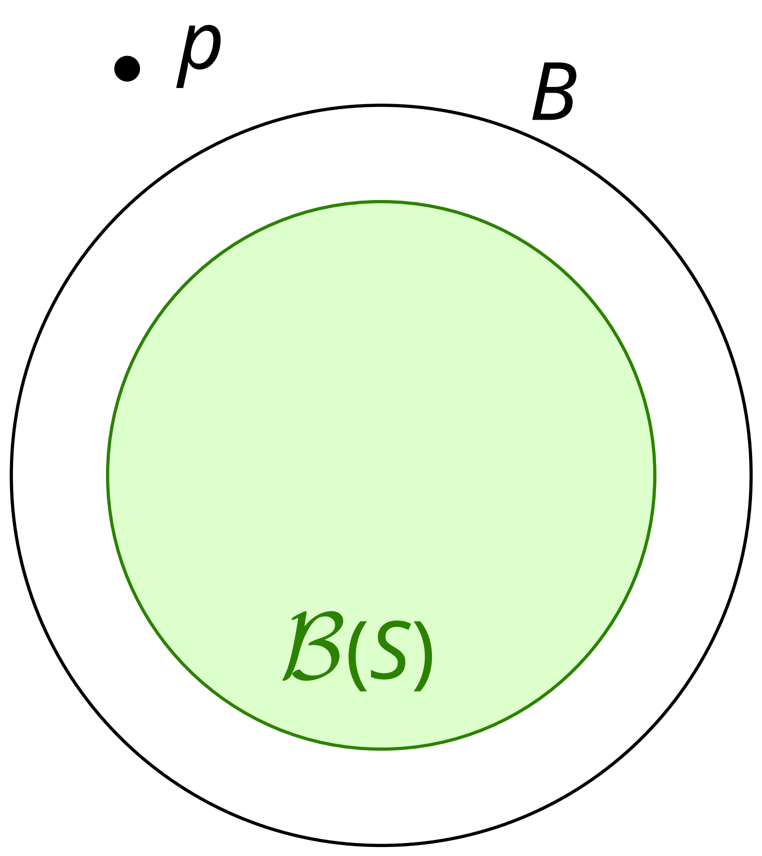
combinatorial/statistical error

does not guarantee $p \in B$



Use Case: minimum enclosing ball

If S is an $\varepsilon/4$-coreset of P for directional width, then the smallest enclosing ball of S, $\mathcal{B}(S)$, scaled by $(1 + \varepsilon)$ around its center, B, contains P.



Comparison to ε-samples:

ε-sample:

guarantees most points in $\mathcal{B}(S)$

combinatorial/statistical error

does not guarantee $p \in B$

coreset: geometric error (bounded for all points)

Overview

Coreset for directional width

- definition
- applications
- construction algorithm ←

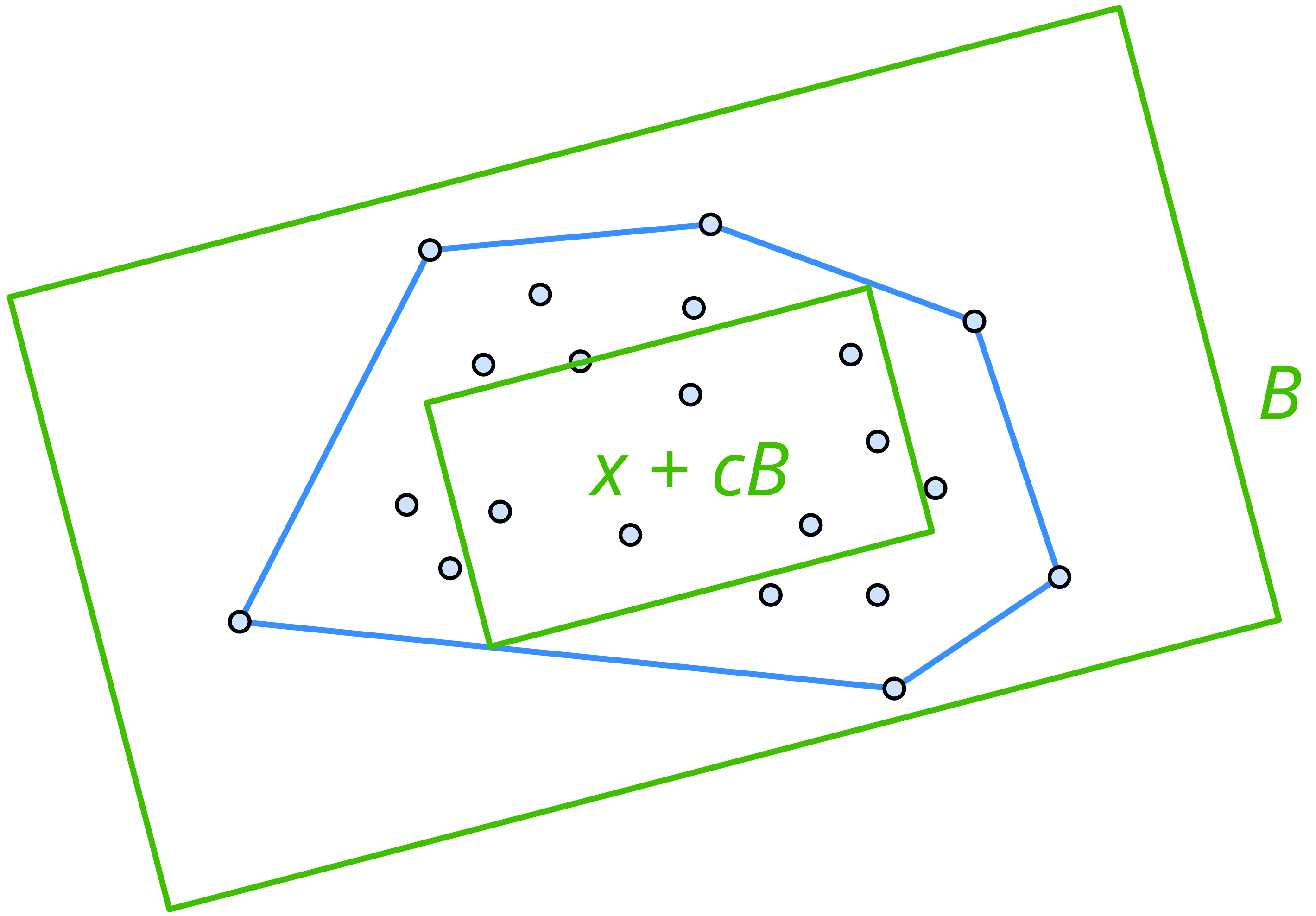
Extra ingredient: Minimum volume bounding box ←

Computing a tight (enough) bounding box

We can compute a bounding box B of P in $O(d^2 n)$ time s.t.

$$(i)\ \mathrm{Vol}(B_{opt}(P)) \leq \mathrm{Vol}(B) \leq 2^d d! \mathrm{Vol}(B_{opt}(P))$$

and (ii) there is a shift $x \in \mathbb{R}^d$ and $c > 0$ that depends only on d, s.t. $x + cB \subset \mathrm{conv}(P)$.



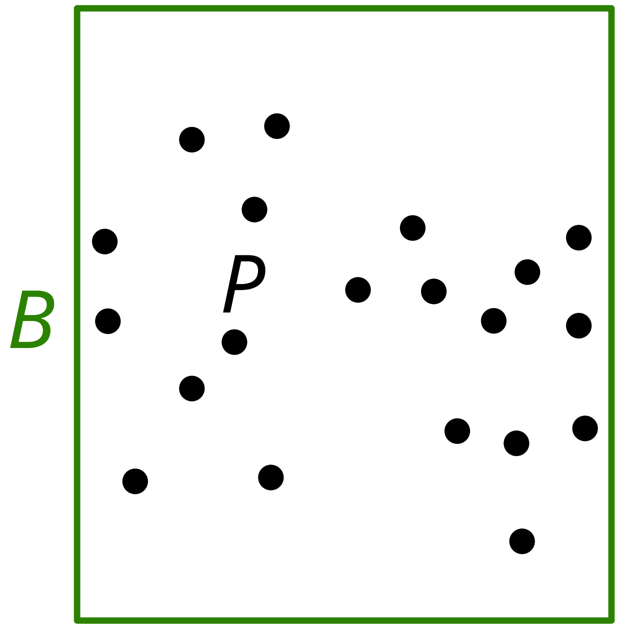
(without proof, for now)

Constructing a coreset

Input: $P \subset \mathbb{R}^d$, $\varepsilon > 0$ (and bounding box B s.t. $c_d B \subset \text{conv}(P) \subset B$)

Output: an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{d-1})$

Construction time: $O(n)$ (also depends on d and ε).

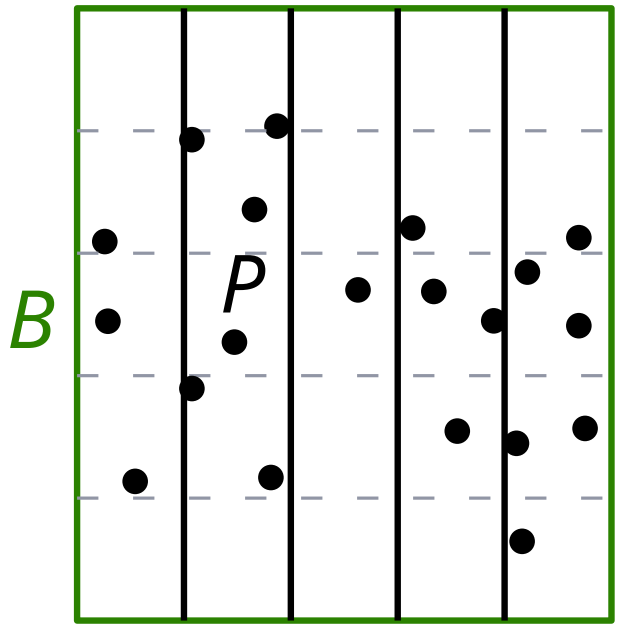


Constructing a coreset

Input: $P \subset \mathbb{R}^d$, $\varepsilon > 0$ (and bounding box B s.t. $c_d B \subset \text{conv}(P) \subset B$)

Output: an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{d-1})$

Construction time: $O(n)$ (also depends on d and ε).



Algorithm

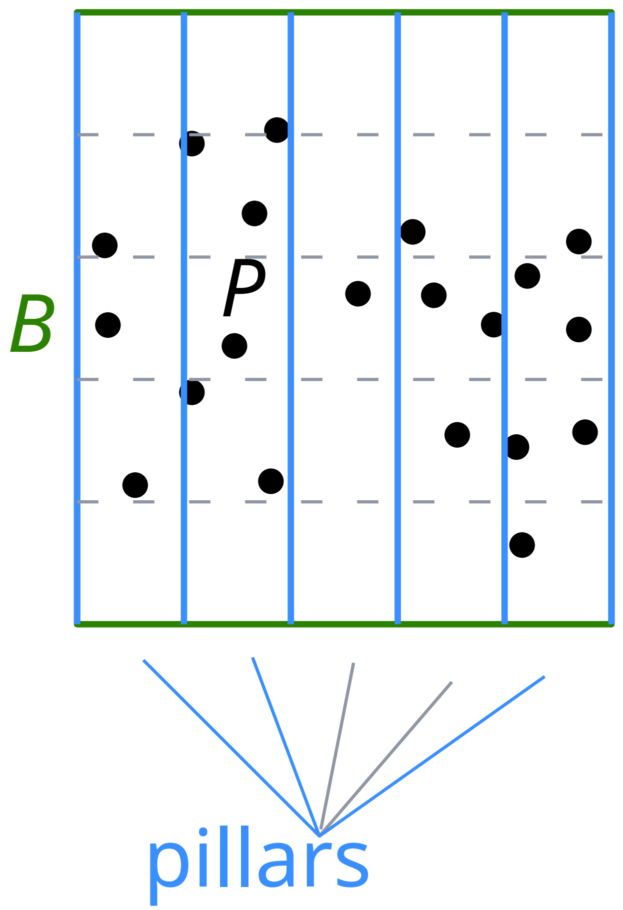
1. Divide B into $M \times \cdots \times M$ grid cells with $M = \lceil \frac{2}{\varepsilon c_d} \rceil$

Constructing a coreset

Input: $P \subset \mathbb{R}^d$, $\varepsilon > 0$ (and bounding box B s.t. $c_d B \subset \text{conv}(P) \subset B$)

Output: an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{d-1})$

Construction time: $O(n)$ (also depends on d and ε).



Algorithm

1. Divide B into $M \times \cdots \times M$ grid cells with $M = \lceil \frac{2}{\varepsilon c_d} \rceil$

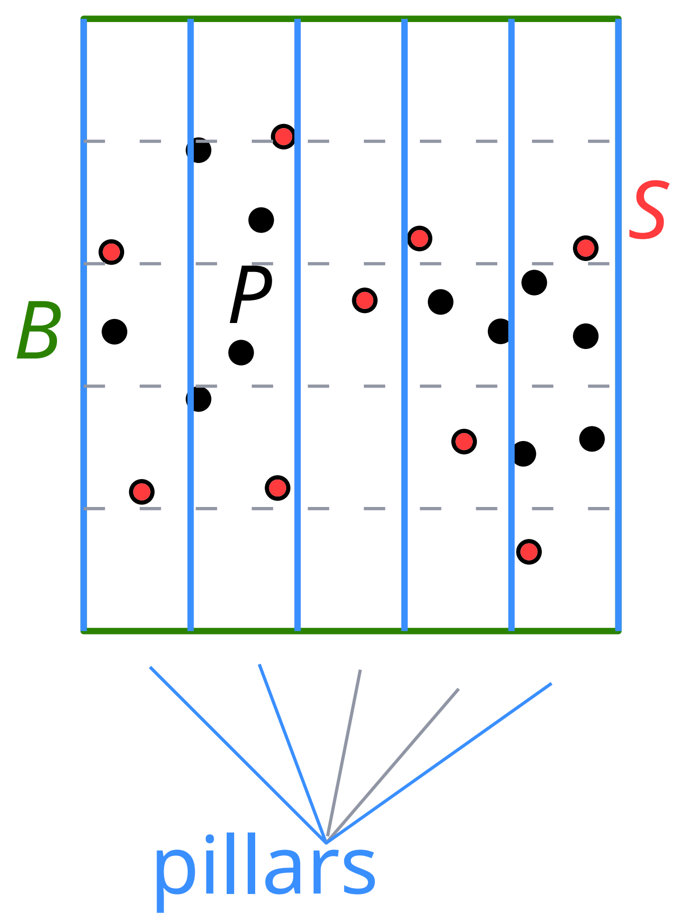
Pillar of cell $(i_1, \dots, i_d)$: $(i_1, \dots, i_{d-1})$

Constructing a coreset

Input: $P \subset \mathbb{R}^d$, $\varepsilon > 0$ (and bounding box B s.t. $c_d B \subset \text{conv}(P) \subset B$)

Output: an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{d-1})$

Construction time: $O(n)$ (also depends on d and ε).



Algorithm

1. Divide B into $M \times \cdots \times M$ grid cells with $M = \lceil \frac{2}{\varepsilon c_d} \rceil$

Pillar of cell $(i_1, \ldots, i_d)$: $(i_1, \ldots, i_{d-1})$

2. For each of the M^{d-1} pillars, find point with max and point with min x_d-coordinate

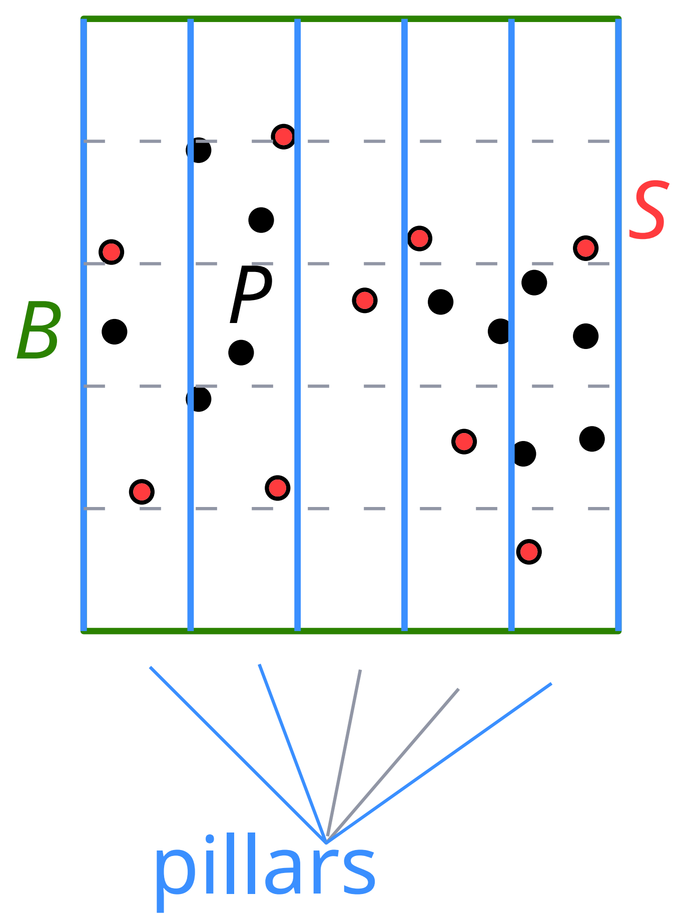
$\underbrace{\hphantom{xxxxxxxxxx}}_{S}$

Constructing a coreset

Input: $P \subset \mathbb{R}^d$, $\varepsilon > 0$ (and bounding box B s.t. $c_d B \subset \text{conv}(P) \subset B$)

Output: an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{d-1})$

Construction time: $O(n)$ (also depends on d and ε).



Algorithm

1. Divide B into $M \times \cdots \times M$ grid cells with $M = \lceil \frac{2}{\varepsilon c_d} \rceil$

Pillar of cell $(i_1, \ldots, i_d)$: $(i_1, \ldots, i_{d-1})$

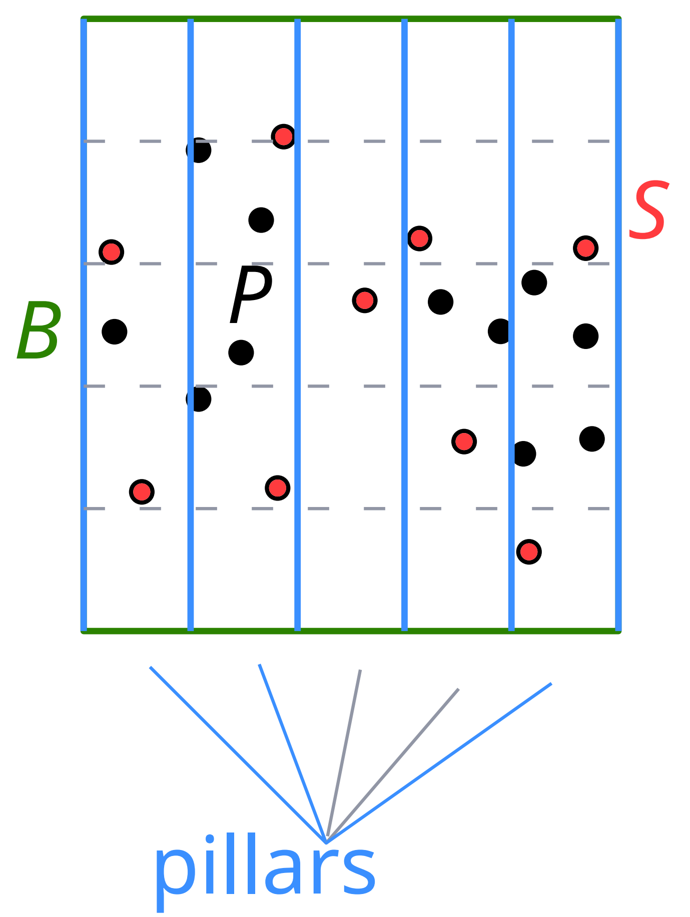
2. For each of the M^{d-1} pillars, find point with max and point with min x_d-coordinate

$\underbrace{\hspace{10em}}_{S}$

$|S| \leq 2M^{d-1} = O(1/\varepsilon^{d-1})$, still need: S is coreset

Constructing a coreset

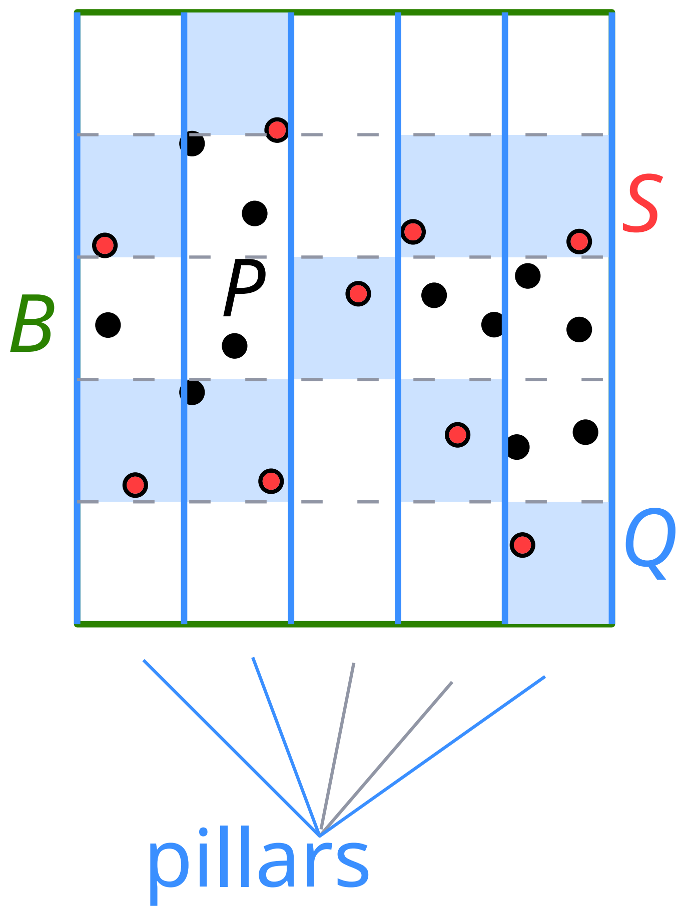
proof (S is coreset):



Constructing a coreset

proof (S is coreset):

Let Q = union of cells containg a point of S.

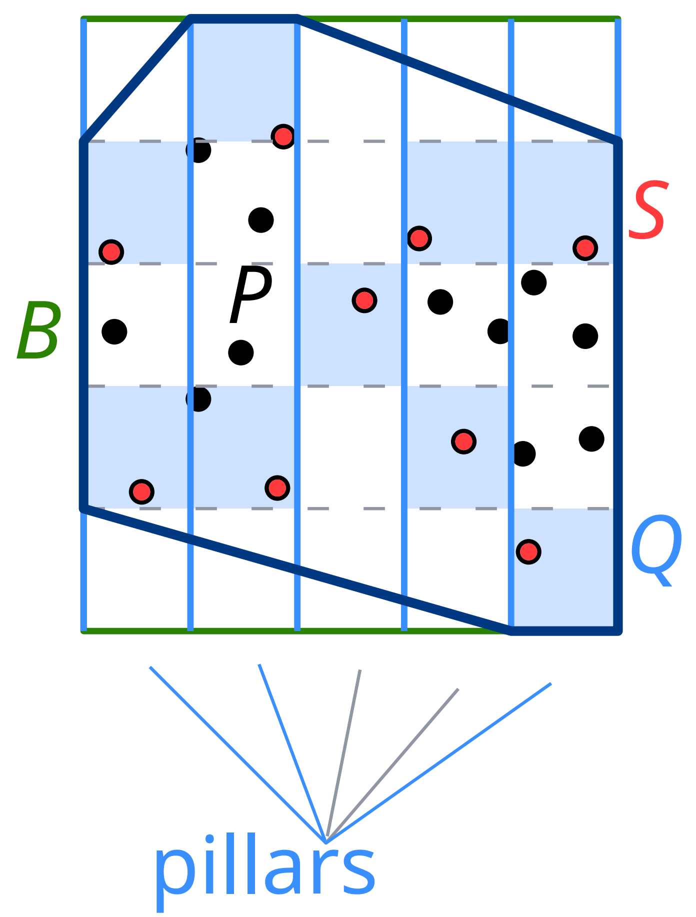


Constructing a coreset

proof (S is coreset):

Let Q = union of cells containg a point of S.

$P \subset \text{conv}(Q) \Rightarrow \text{wd}(v, P) \leq \text{wd}(v, \text{conv}(Q))$ for all v



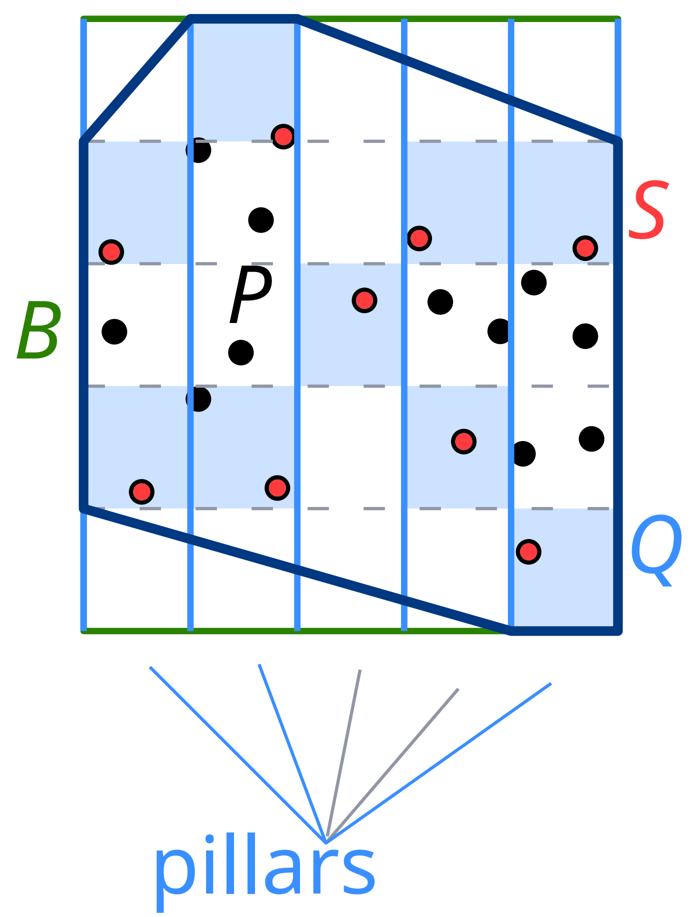
Constructing a coreset

proof (S is coreset):

Let Q = union of cells containg a point of S.

$P \subset \text{conv}(Q) \Rightarrow \text{wd}(v, P) \leq \text{wd}(v, \text{conv}(Q))$ for all v

$\text{wd}(v, P) \leq \text{wd}(v, Q) \leq \text{wd}(v, S) + 2\text{wd}(v, B/M)$



Constructing a coreset

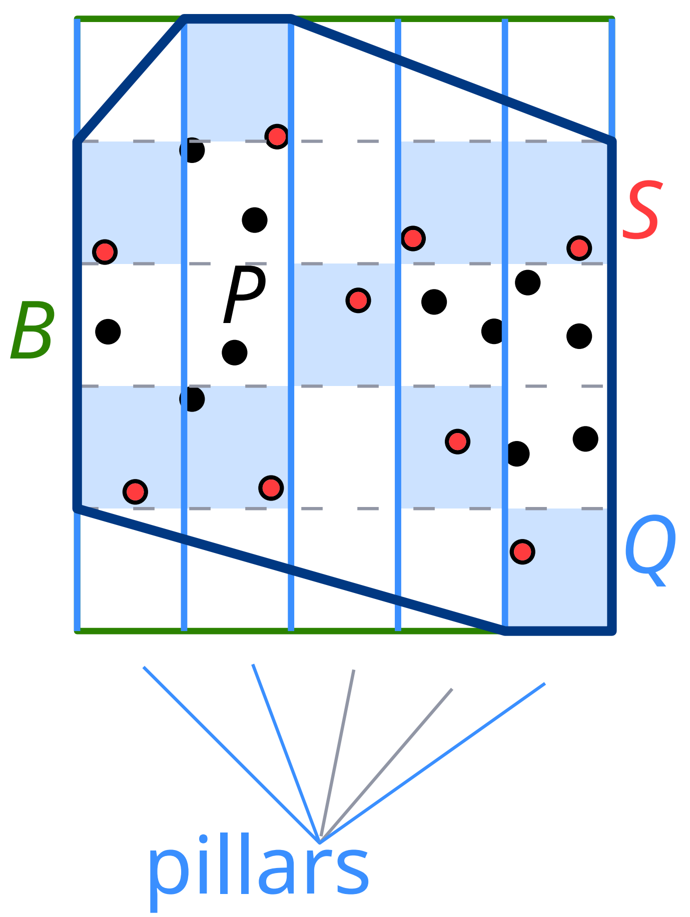
proof (S is coreset):

Let Q = union of cells containg a point of S.

$P \subset \text{conv}(Q) \Rightarrow \text{wd}(v, P) \leq \text{wd}(v, \text{conv}(Q))$ for all v

$\text{wd}(v, P) \leq \text{wd}(v, Q) \leq \text{wd}(v, S) + 2\text{wd}(v, B/M)$

$\text{wd}(v, B/M) = \frac{\text{wd}(v,B)}{M} = \frac{\text{wd}(v,c_d B)}{c_d M}$



Constructing a coreset

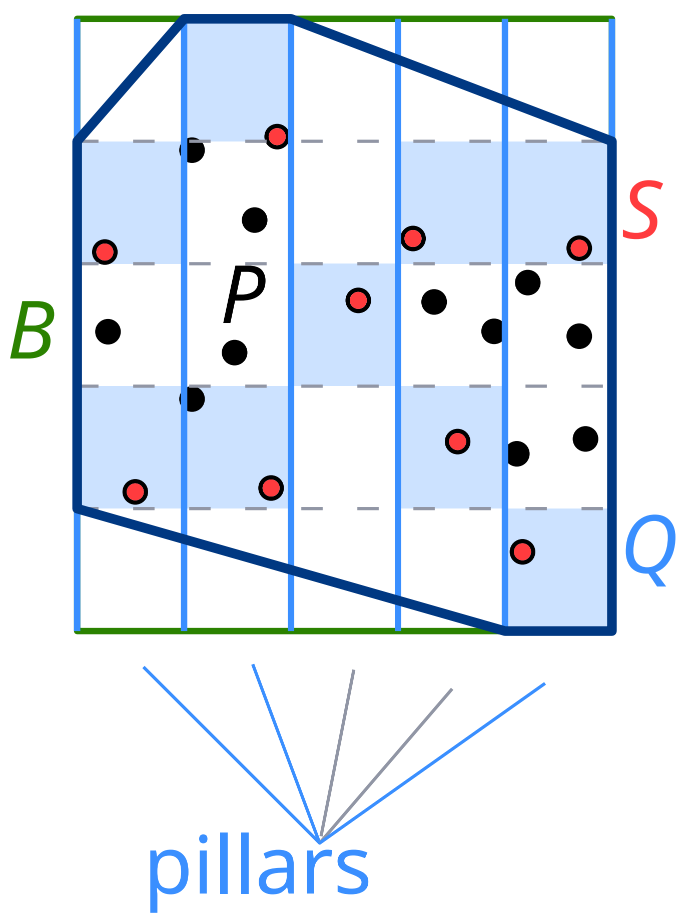
proof (S is coreset):

Let Q = union of cells containg a point of S.

$P \subset \text{conv}(Q) \Rightarrow \text{wd}(v, P) \leq \text{wd}(v, \text{conv}(Q))$ for all v

$\text{wd}(v, P) \leq \text{wd}(v, Q) \leq \text{wd}(v, S) + 2\text{wd}(v, B/M)$

$$\text{wd}(v, B/M) = \frac{\text{wd}(v,B)}{M} = \frac{\text{wd}(v,c_d B)}{c_d M}$$
$$\leq \frac{\text{wd}(v,P)}{c_d M} \leq \frac{\text{wd}(v,P)}{2/\varepsilon} = \frac{\varepsilon}{2}\text{wd}(v, P)$$



Constructing a coreset

proof (S is coreset):

Let Q = union of cells containg a point of S.

$P \subset \text{conv}(Q) \Rightarrow \text{wd}(v, P) \leq \text{wd}(v, \text{conv}(Q))$ for all v

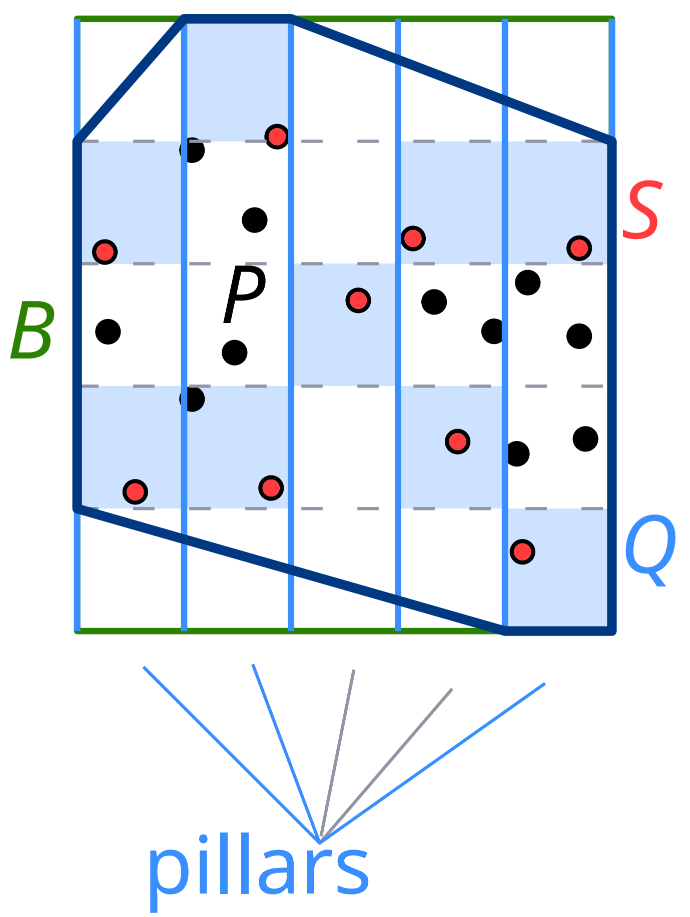
$\text{wd}(v, P) \leq \text{wd}(v, Q) \leq \text{wd}(v, S) + 2\text{wd}(v, B/M)$

$$\text{wd}(v, B/M) = \frac{\text{wd}(v,B)}{M} = \frac{\text{wd}(v,c_d B)}{c_d M}$$

$$\leq \frac{\text{wd}(v,P)}{c_d M} \leq \frac{\text{wd}(v,P)}{2/\varepsilon} = \frac{\varepsilon}{2}\text{wd}(v, P)$$

$$\Rightarrow \text{wd}(v, P)(1 - 2\tfrac{\varepsilon}{2})\text{wd}(v, P) \leq \text{wd}(v, S)$$

□



Constructing a smaller coreset

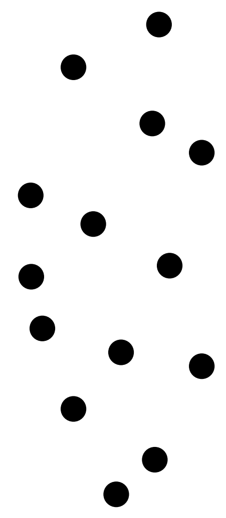
Given $\varepsilon > 0$ and $P \subset \mathbb{R}^d$, we can compute an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{(d-1)/2})$ in $O(n + 1/\varepsilon^{3(d-1)/2})$ time (where d is a fixed constant).

Constructing a smaller coreset

Given $\varepsilon > 0$ and $P \subset \mathbb{R}^d$, we can compute an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{(d-1)/2})$ in $O(n + 1/\varepsilon^{3(d-1)/2})$ time (where d is a fixed constant).

algorithm ideas/proof sketch:

- two stages: first the previous algo. for $\varepsilon/2$ gives S', then this (slower) algorithm for $\varepsilon/2$ on S' gives S
- make conv(P) fat via affine transformation into unit hypercube
- find small enclosing ball B (radius $\sqrt{d}$)
- X := a $c\sqrt{\varepsilon}$-packing in ∂B
- Let S = nearest point to each $x \in X$

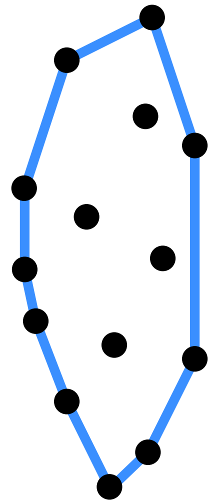


Constructing a smaller coreset

Given $\varepsilon > 0$ and $P \subset \mathbb{R}^d$, we can compute an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{(d-1)/2})$ in $O(n + 1/\varepsilon^{3(d-1)/2})$ time (where d is a fixed constant).

algorithm ideas/proof sketch:

- two stages: first the previous algo. for $\varepsilon/2$ gives S', then this (slower) algorithm for $\varepsilon/2$ on S' gives S
- make conv(P) fat via affine transformation into unit hypercube
- find small enclosing ball B (radius $\sqrt{d}$)
- X := a $c\sqrt{\varepsilon}$-packing in ∂B
- Let S = nearest point to each $x \in X$

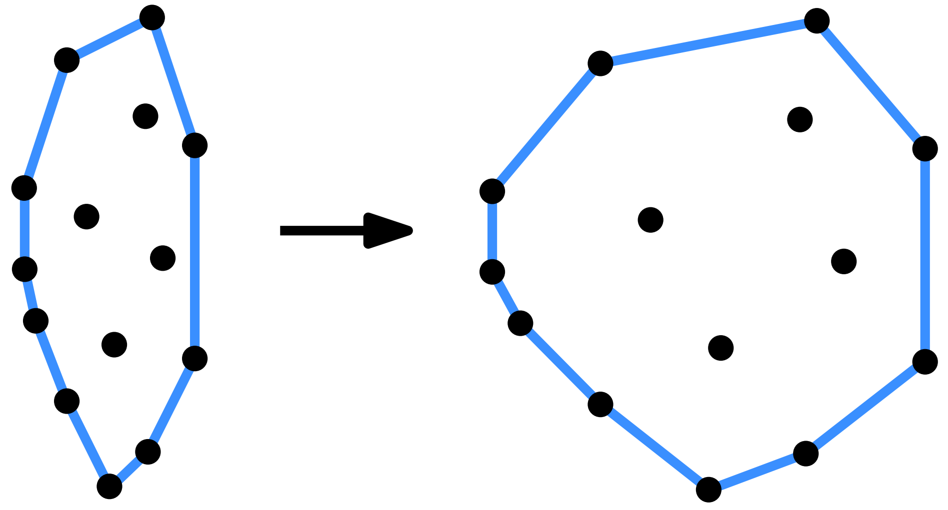


Constructing a smaller coreset

Given $\varepsilon > 0$ and $P \subset \mathbb{R}^d$, we can compute an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{(d-1)/2})$ in $O(n + 1/\varepsilon^{3(d-1)/2})$ time (where d is a fixed constant).

algorithm ideas/proof sketch:

- two stages: first the previous algo. for $\varepsilon/2$ gives S', then this (slower) algorithm for $\varepsilon/2$ on S' gives S
- make conv(P) fat via affine transformation into unit hypercube
- find small enclosing ball B (radius $\sqrt{d}$)
- X := a $c\sqrt{\varepsilon}$-packing in ∂B
- Let S = nearest point to each $x \in X$

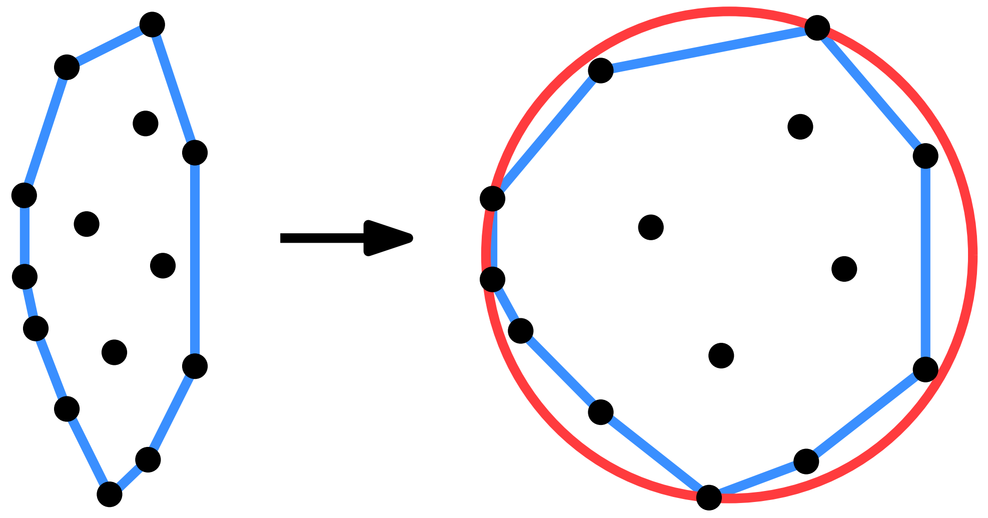


Constructing a smaller coreset

Given $\varepsilon > 0$ and $P \subset \mathbb{R}^d$, we can compute an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{(d-1)/2})$ in $O(n + 1/\varepsilon^{3(d-1)/2})$ time (where d is a fixed constant).

algorithm ideas/proof sketch:

- two stages: first the previous algo. for $\varepsilon/2$ gives S', then this (slower) algorithm for $\varepsilon/2$ on S' gives S
- make conv(P) fat via affine transformation into unit hypercube
- find small enclosing ball B (radius $\sqrt{d}$)
- X := a $c\sqrt{\varepsilon}$-packing in ∂B
- Let S = nearest point to each $x \in X$

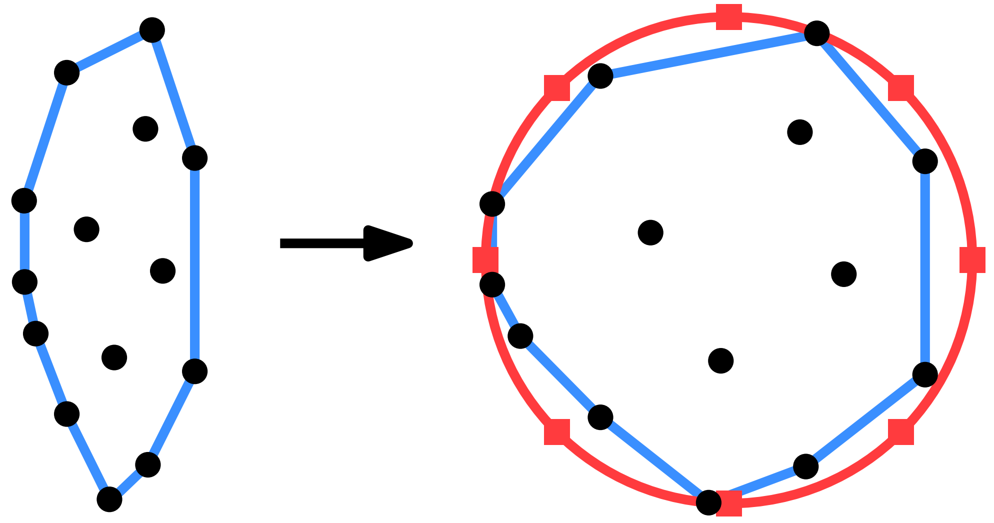


Constructing a smaller coreset

Given $\varepsilon > 0$ and $P \subset \mathbb{R}^d$, we can compute an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{(d-1)/2})$ in $O(n + 1/\varepsilon^{3(d-1)/2})$ time (where d is a fixed constant).

algorithm ideas/proof sketch:

- two stages: first the previous algo. for $\varepsilon/2$ gives S', then this (slower) algorithm for $\varepsilon/2$ on S' gives S
- make conv(P) fat via affine transformation into unit hypercube
- find small enclosing ball B (radius $\sqrt{d}$)
- X := a $c\sqrt{\varepsilon}$-packing in ∂B
- Let S = nearest point to each $x \in X$

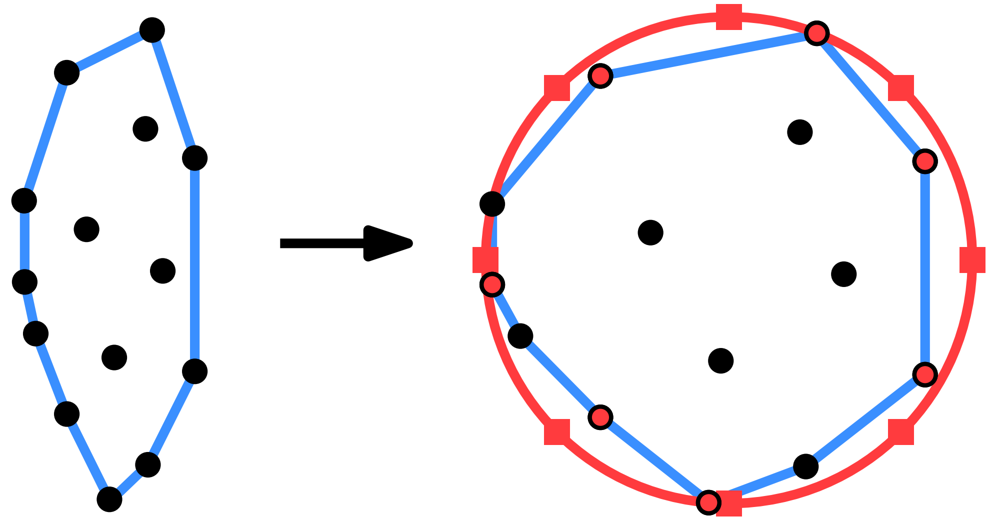


Constructing a smaller coreset

Given $\varepsilon > 0$ and $P \subset \mathbb{R}^d$, we can compute an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{(d-1)/2})$ in $O(n + 1/\varepsilon^{3(d-1)/2})$ time (where d is a fixed constant).

algorithm ideas/proof sketch:

- two stages: first the previous algo. for $\varepsilon/2$ gives S', then this (slower) algorithm for $\varepsilon/2$ on S' gives S
- make conv(P) fat via affine transformation into unit hypercube
- find small enclosing ball B (radius $\sqrt{d}$)
- X := a $c\sqrt{\varepsilon}$-packing in ∂B
- Let S = nearest point to each $x \in X$

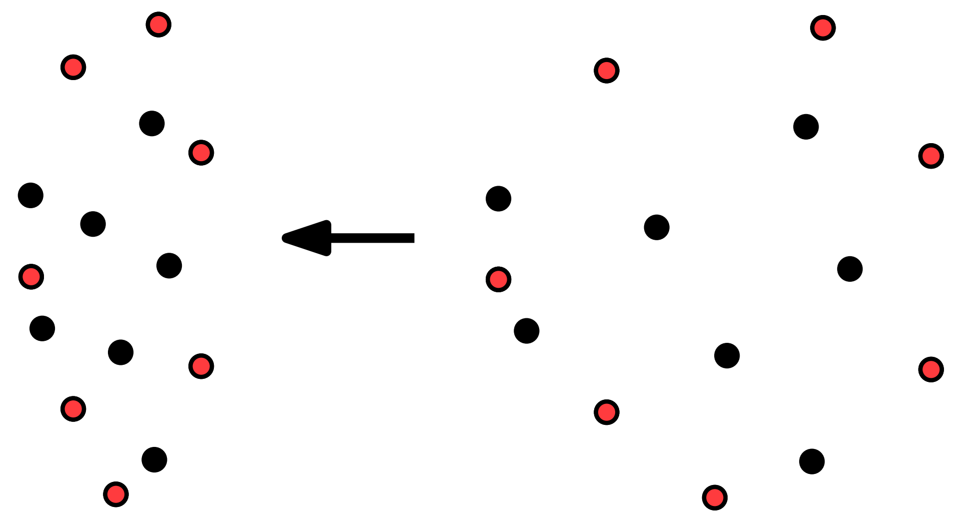


Constructing a smaller coreset

Given $\varepsilon > 0$ and $P \subset \mathbb{R}^d$, we can compute an ε-coreset $S \subseteq P$ of size at most $|S| = O(1/\varepsilon^{(d-1)/2})$ in $O(n + 1/\varepsilon^{3(d-1)/2})$ time (where d is a fixed constant).

algorithm ideas/proof sketch:

- two stages: first the previous algo. for $\varepsilon/2$ gives S', then this (slower) algorithm for $\varepsilon/2$ on S' gives S
- make conv(P) fat via affine transformation into unit hypercube
- find small enclosing ball B (radius $\sqrt{d}$)
- X := a $c\sqrt{\varepsilon}$-packing in ∂B
- Let S = nearest point to each $x \in X$



Overview

Coreset for directional width

- definition
- applications
- construction algorithm

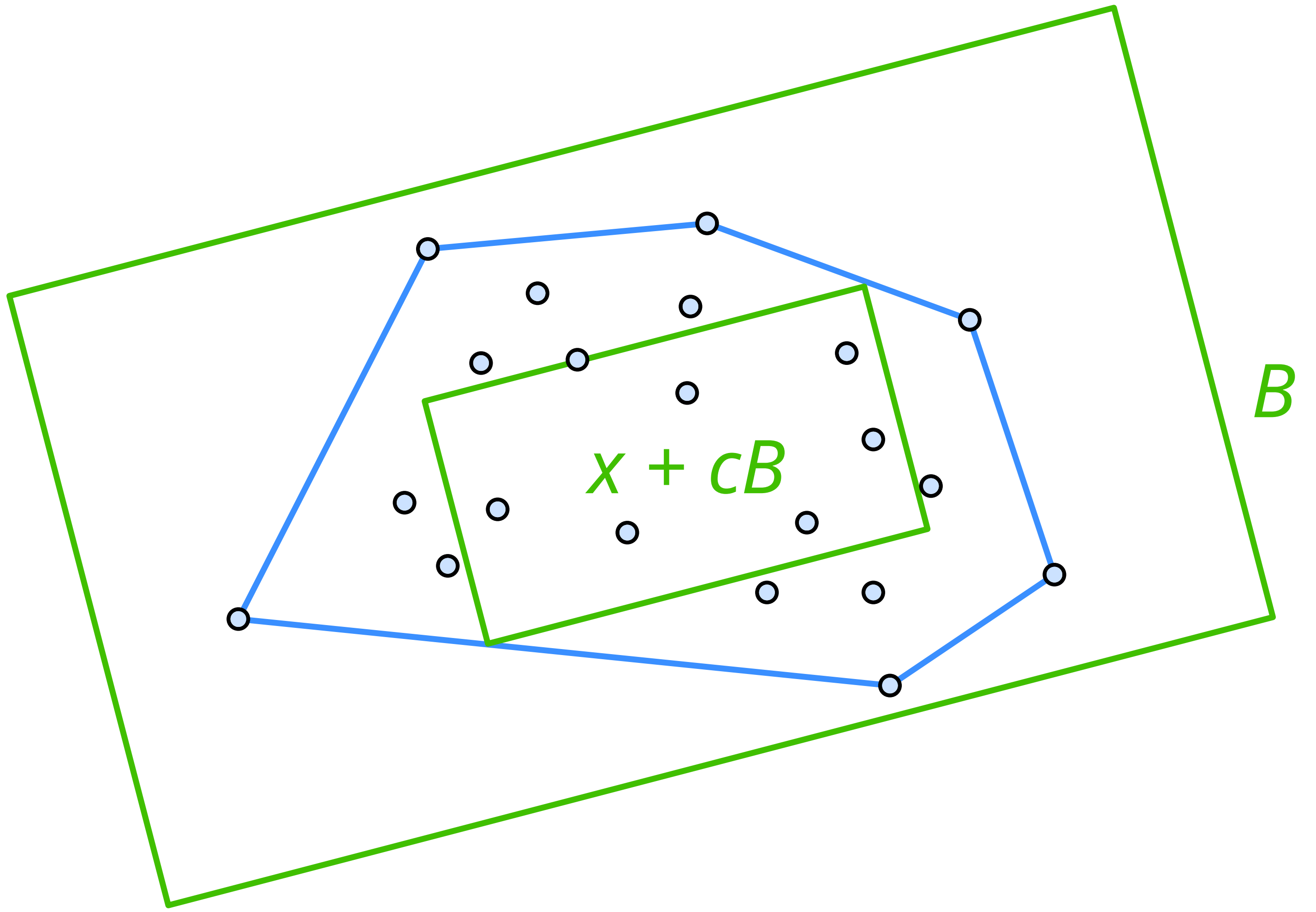
Extra ingredient: Minimum volume bounding box ←

Computing a tight (enough) bounding box

We can compute a bounding box B of P in $O(d^2 n)$ time s.t.

$$(i)\ \text{Vol}(B_{opt}(P)) \leq \text{Vol}(B) \leq 2^d d!\text{Vol}(B_{opt}(P))$$

and (ii) there is a shift $x \in \mathbb{R}^d$ and $c > 0$ that depends only on d, s.t. $x + cB \subset \text{conv}(P)$.



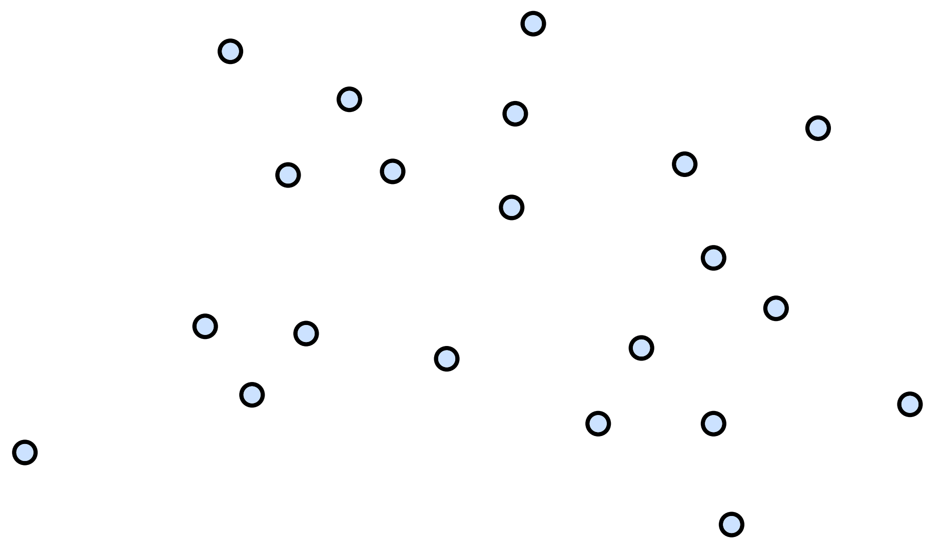
Computing a tight (enough) bounding box

We can compute a bounding box B of P in $O(d^2 n)$ time s.t.

$$(i)\ \mathrm{Vol}(B_{opt}(P)) \leq \mathrm{Vol}(B) \leq 2^d d!\mathrm{Vol}(B_{opt}(P))$$

and (ii) there is a shift $x \in \mathbb{R}^d$ and $c > 0$ that depends only on d, s.t. $x + cB \subset \mathrm{conv}(P)$.

1. Approximate diameter:



Computing a tight (enough) bounding box

We can compute a bounding box B of P in $O(d^2 n)$ time s.t.

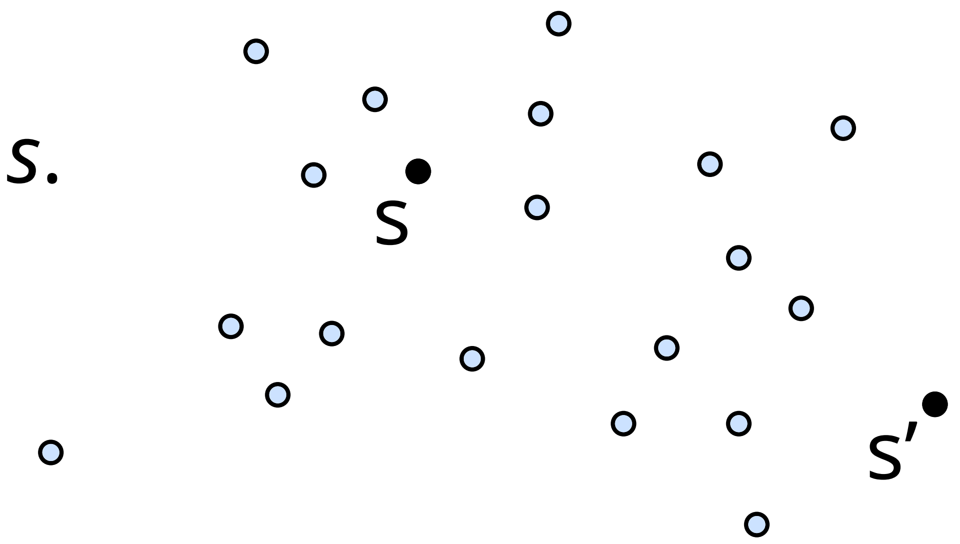
$$(i)\ \mathrm{Vol}(B_{opt}(P)) \leq \mathrm{Vol}(B) \leq 2^d d! \mathrm{Vol}(B_{opt}(P))$$

and (ii) there is a shift $x \in \mathbb{R}^d$ and $c > 0$ that depends only on d, s.t. $x + cB \subset \mathrm{conv}(P)$.

1. Approximate diameter:

Let $s \in P$ arbitrary and let $s' \in P$ most distant form s.
If t, t' realize the diameter of P, then

$$\mathrm{diam}(P) = |tt'| \leq |ts| + |st'| \leq 2|ss'|$$



Computing a tight (enough) bounding box

We can compute a bounding box B of P in $O(d^2 n)$ time s.t.

$$(i)\ \mathrm{Vol}(B_{opt}(P)) \leq \mathrm{Vol}(B) \leq 2^d d! \mathrm{Vol}(B_{opt}(P))$$

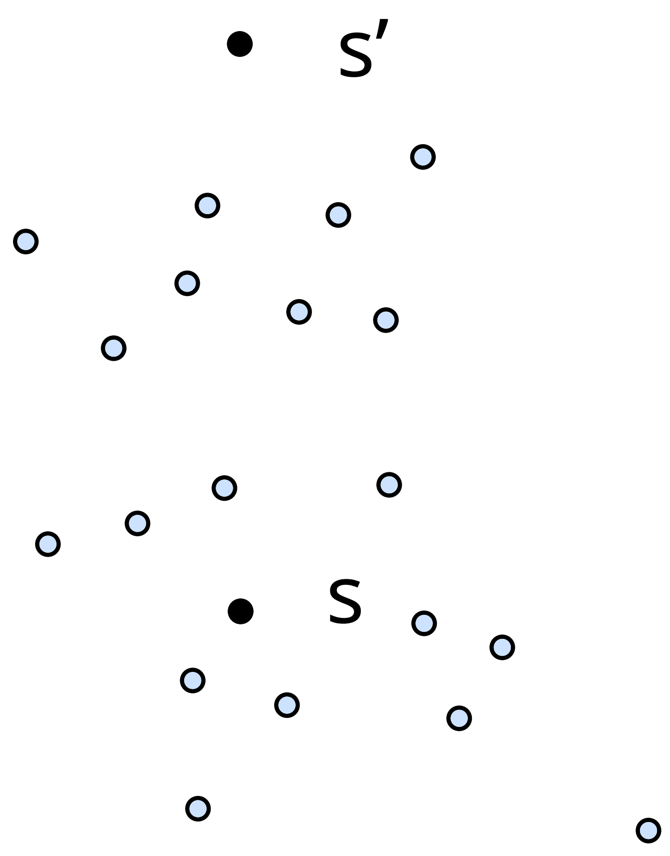
and (ii) there is a shift $x \in \mathbb{R}^d$ and $c > 0$ that depends only on d, s.t. $x + cB \subset \mathrm{conv}(P)$.

1. Approximate diameter:

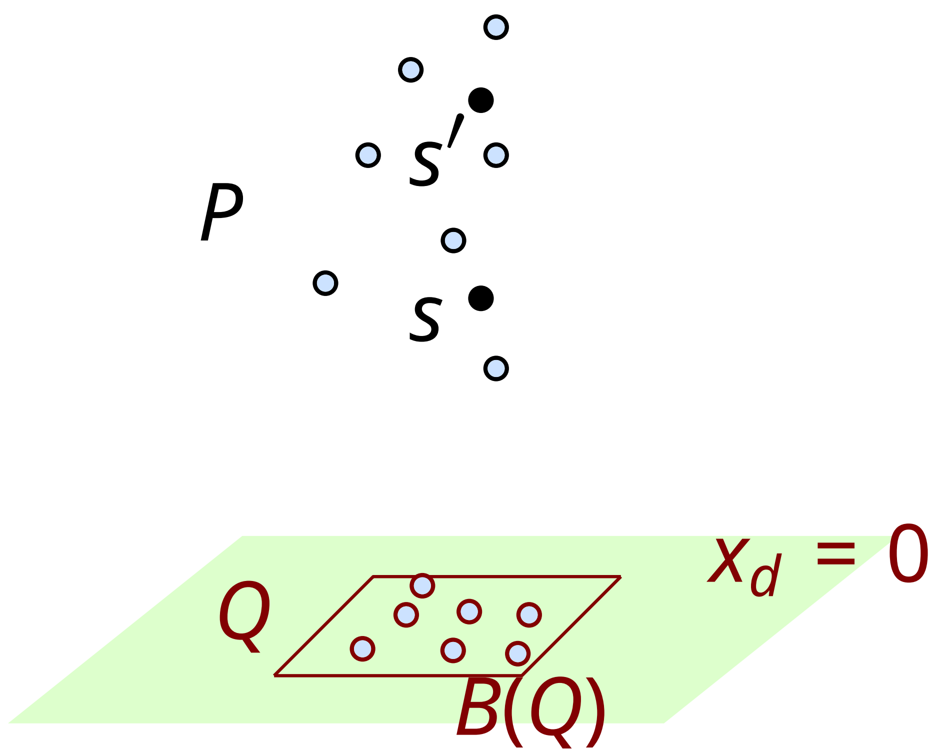
Let $s \in P$ arbitrary and let $s' \in P$ most distant form s.
If t, t' realize the diameter of P, then

$$\mathrm{diam}(P) = |tt'| \leq |ts| + |st'| \leq 2|ss'|$$

Wlog. ss' parallel to x_d axis.
π := perpendicular projection to $x_d = 0$.

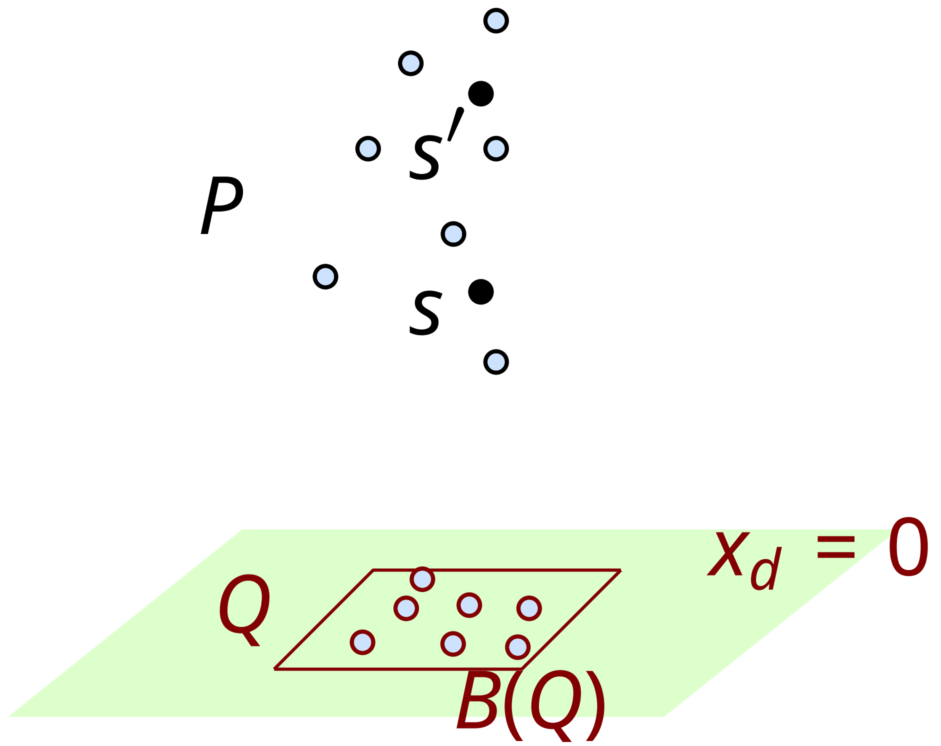


Recursive step



Recursive step

$d = 1$: return interval containing points



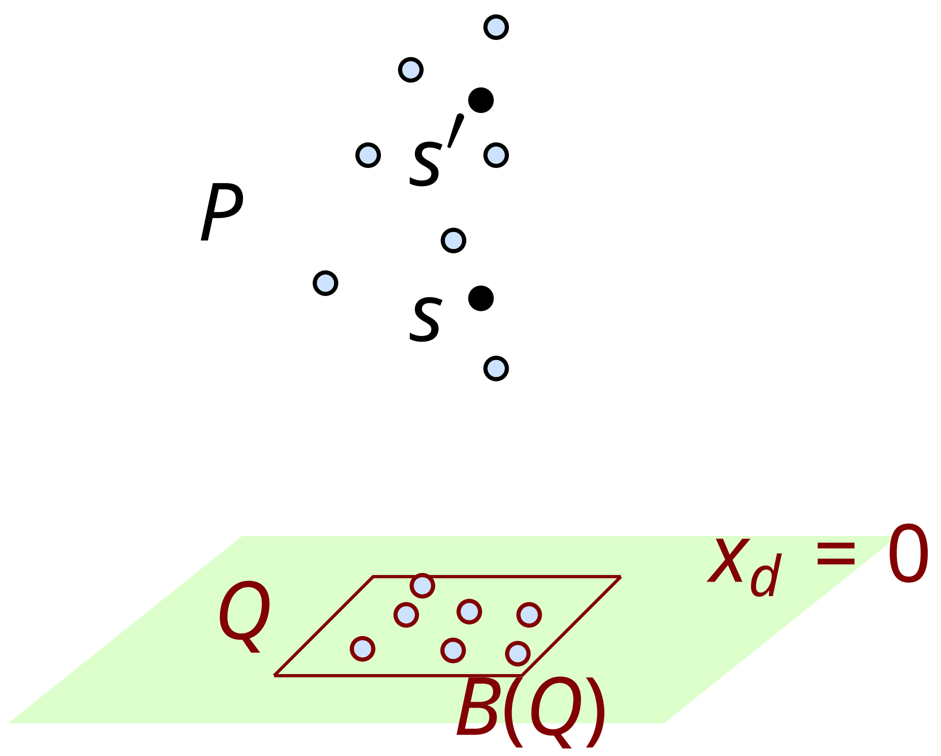
Recursive step

$d = 1$: return interval containing points

$d > 1$:

$Q := \pi(P)$

$B(Q) :=$ bounding box of Q (recursion)



Recursive step

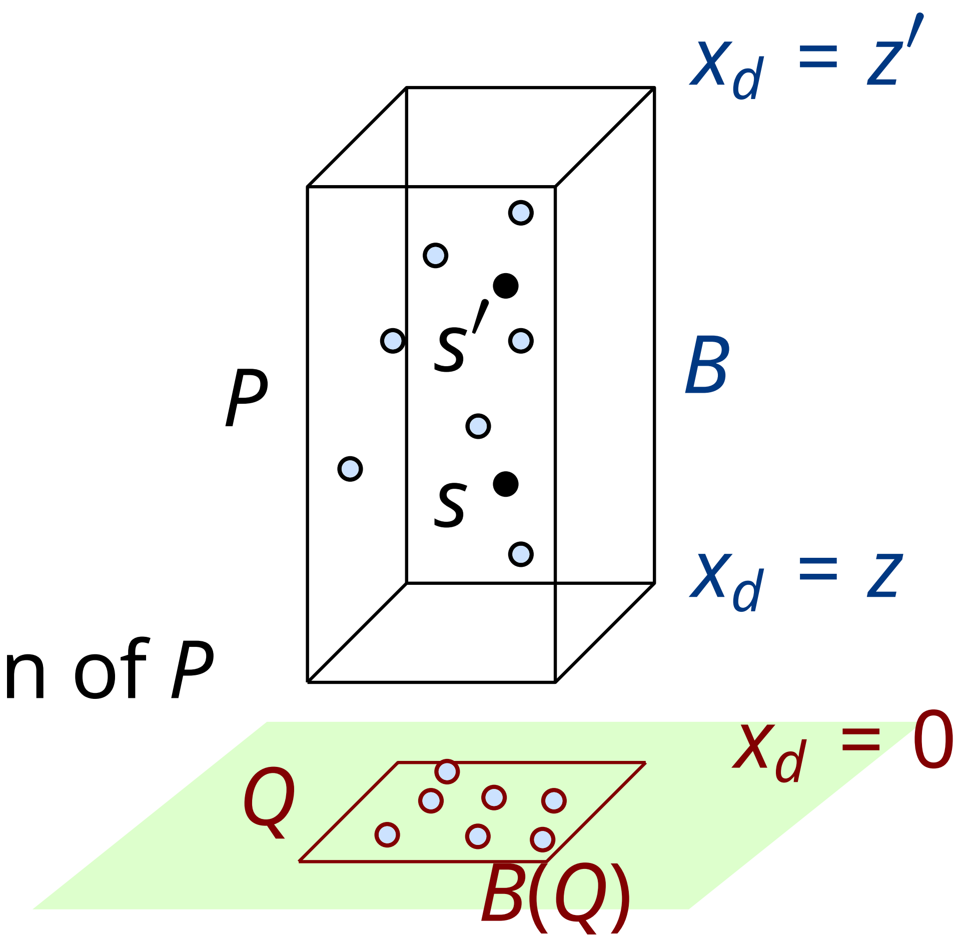
$d = 1$: return interval containing points

$d > 1$:

$Q := \pi(P)$
$B(Q) :=$ bounding box of Q (recursion)

$[z, z'] :=$ shortest interval on x_d axis covering projection of P

$$B := B(Q) \times [z, z']$$



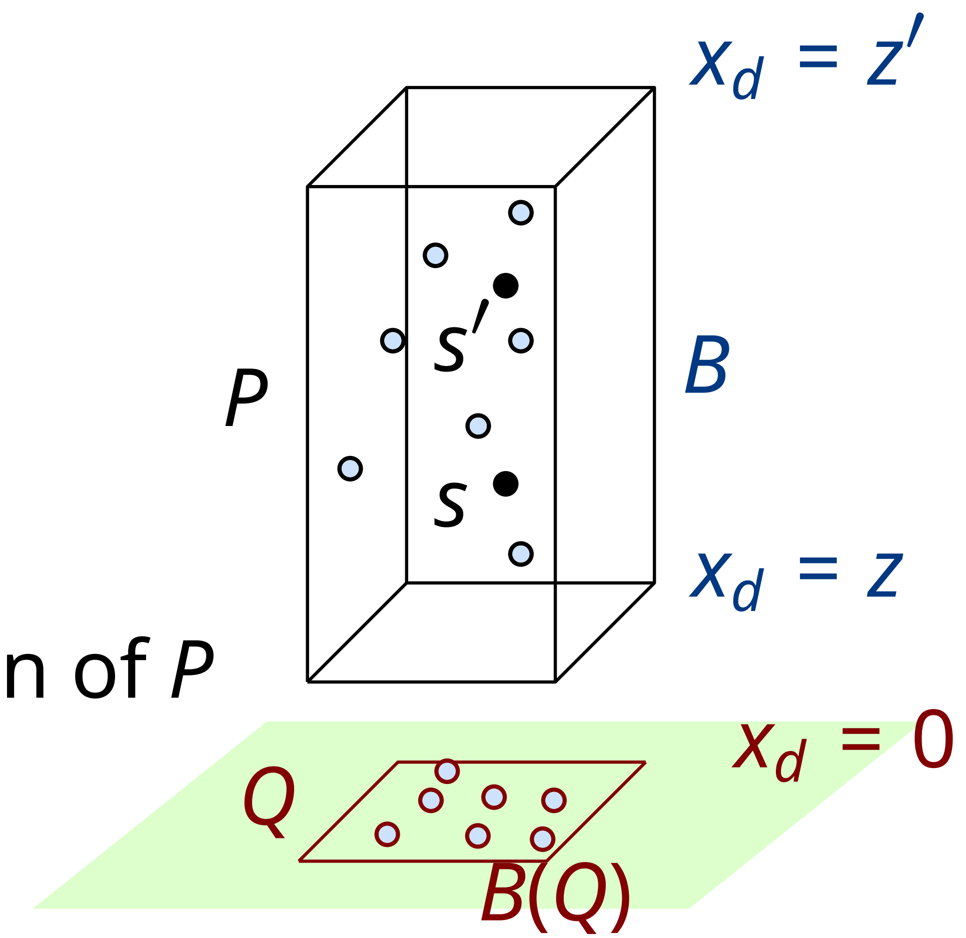
Recursive step

$d = 1$: return interval containing points

$d > 1$:

$Q := \pi(P)$
$B(Q) :=$ bounding box of Q (recursion)

$[z, z'] :=$ shortest interval on x_d axis covering projection of P



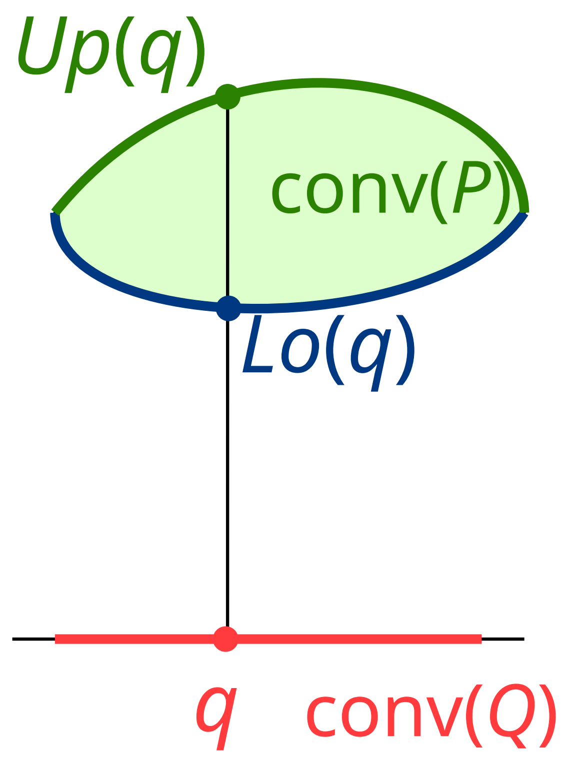
$$B := B(Q) \times [z, z']$$

Still need: $\mathrm{Vol}_d(\mathrm{conv}(P)) \geq \mathrm{Vol}_d(B)/(2^d d!)$

Volume bound

Upper hull $\mathrm{conv}^{\uparrow}(P)$ as function: $Up : \mathrm{conv}(Q) \rightarrow \mathbb{R}^d$ is concave

Lower hull $\mathrm{conv}^{\downarrow}(P)$ as function: $Lo : \mathrm{conv}(Q) \rightarrow \mathbb{R}^d$ is convex



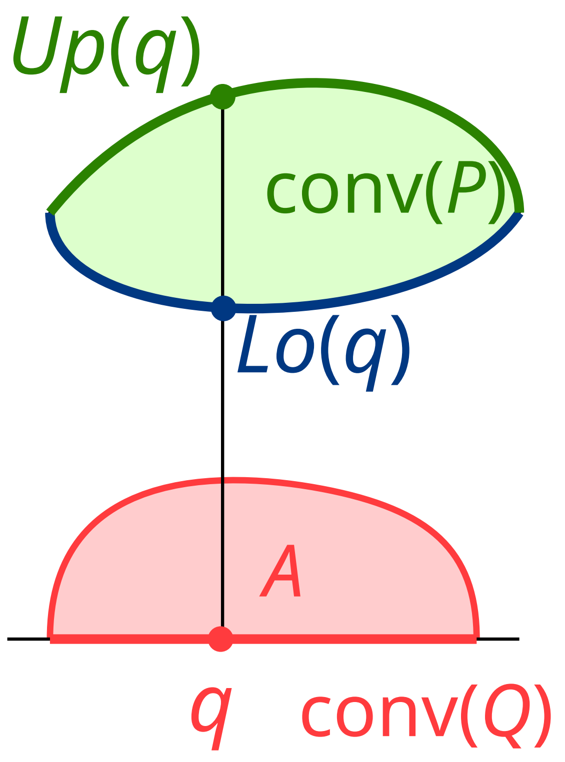
Volume bound

Upper hull $\text{conv}^{\uparrow}(P)$ as function: $Up : \text{conv}(Q) \to \mathbb{R}^d$ is concave

Lower hull $\text{conv}^{\downarrow}(P)$ as function: $Lo : \text{conv}(Q) \to \mathbb{R}^d$ is convex

$Up - Lo$ is concave

$\Rightarrow A := \bigcup_{q \in \text{conv}(Q)} [0, Up(q) - Lo(q)]$ is convex



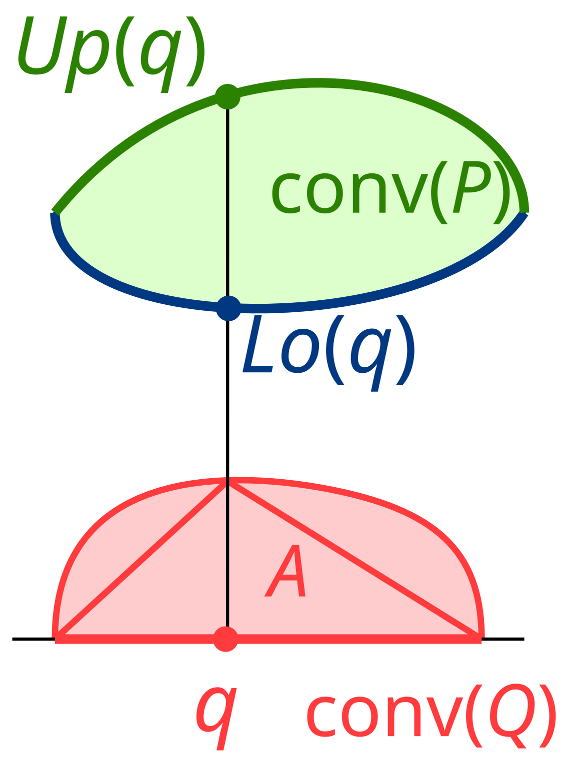
Volume bound

Upper hull $\text{conv}^{\uparrow}(P)$ as function: $Up : \text{conv}(Q) \to \mathbb{R}^d$ is concave

Lower hull $\text{conv}^{\downarrow}(P)$ as function: $Lo : \text{conv}(Q) \to \mathbb{R}^d$ is convex

$Up - Lo$ is concave

$$\Rightarrow A := \bigcup_{q \in \text{conv}(Q)} [0, Up(q) - Lo(q)] \text{ is convex}$$

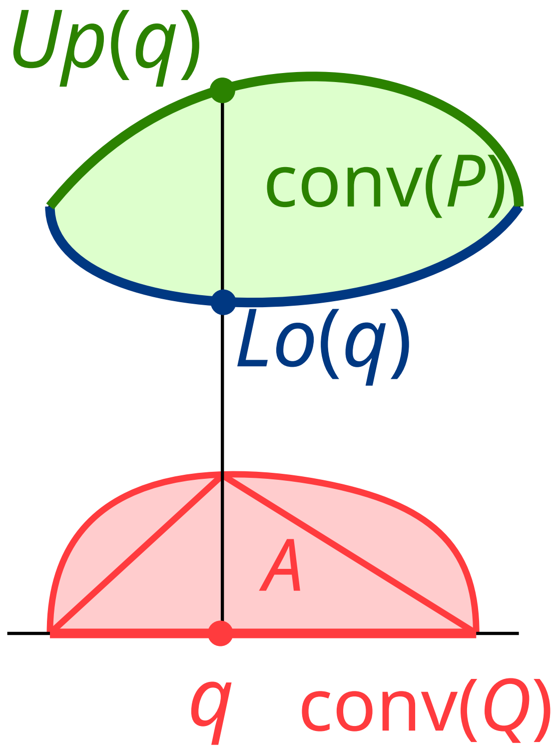


At $\pi(s)$, height of A is at least $|ss'|$.

A contains *pyramid* with base $\text{conv}(Q)$ and pole length $\geq |ss'|$.

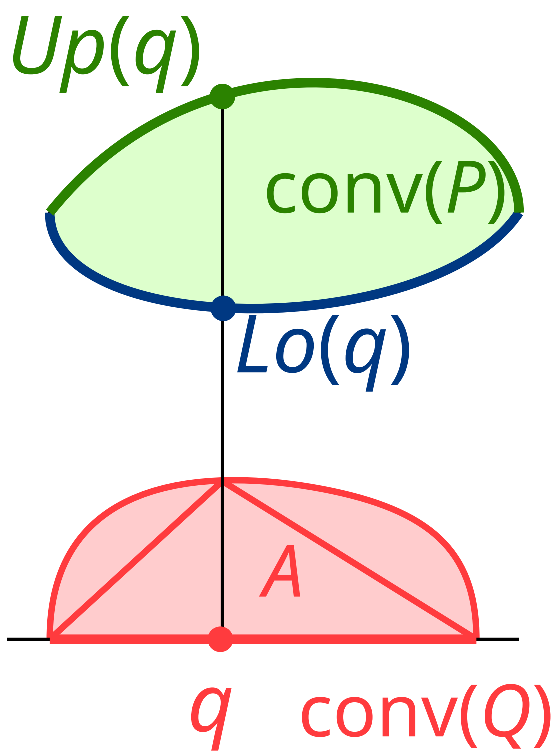
Volume bound

$$
\begin{aligned}
\mathrm{Vol}_d(B) &\geq \mathrm{Vol}_d(B_{opt}) \\
&\geq \mathrm{Vol}_d(\mathrm{conv}(P)) = \mathrm{Vol}_d(A) \\
&\geq \mathrm{Vol}(pyramid) \\
&\geq \frac{\mathrm{Vol}_{d-1}(\mathrm{conv}(Q))|ss'|}{d} \\
&\geq \frac{\mathrm{Vol}_{d-1}\left(B(Q)/(2^{d-1}(d-1)!)\right)2|ss'|}{2d} \\
&\geq \frac{\mathrm{Vol}_{d-1}(B(Q))|zz'|}{2^d d!} \\
&= \frac{\mathrm{Vol}_d(B)}{2^d d!}
\end{aligned}
$$



Volume bound

$$\begin{aligned}
\mathrm{Vol}_d(B) &\geq \mathrm{Vol}_d(B_{opt}) \\
&\geq \mathrm{Vol}_d(\mathrm{conv}(P)) = \mathrm{Vol}_d(A) \\
&\geq \mathrm{Vol}(pyramid) \\
&\geq \frac{\mathrm{Vol}_{d-1}(\mathrm{conv}(Q))|ss'|}{d} \\
&\geq \frac{\mathrm{Vol}_{d-1}\left(B(Q)/(2^{d-1}(d-1)!)\right)2|ss'|}{2d} \\
&\geq \frac{\mathrm{Vol}_{d-1}(B(Q))|zz'|}{2^d d!} \\
&= \frac{\mathrm{Vol}_d(B)}{2^d d!}
\end{aligned}$$



Running time: $T(n, d) = O(nd) + T(n, d-1) = O(nd^2)$.

Summary

Coreset: small (sub-)set capturing the relevant geometry

slow algorithm + coreset = fast approximation algorithm

a coreset is constructed for specific geometric optimization problem

Coreset for directional width:

construction using grids (+ bounding box)

solves various other problems too: min-volume bounding box, min-enclosing ball, diameter, ...