Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

Certified Small Well-Separated Pair Decompositions in Practice

Guilherme D. da Fonseca – Aix-Marseille Université and LIS

2026

Disclaimer: This work has been produced with heavy AI use in all its phases.

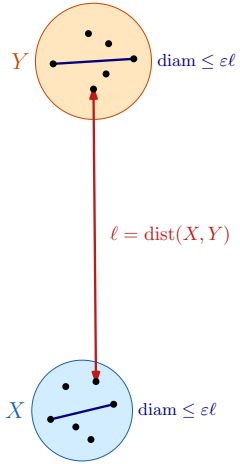
Well-Separated Pair Decomposition

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

Definitions:

- P: set of n points in $\mathbb{R}^d$, $\varepsilon \in (0,1)$
- Dumbbell $\{X, Y\}$: two disjoint nonempty subsets of P, its sides
- ε-separated: $\max\{\mathsf{diam}(X), \mathsf{diam}(Y)\} \leq \varepsilon \cdot \mathsf{dist}(X, Y)$

- A pair is $\{p, q\} \subseteq P$, $p \neq q$
- $\{X, Y\}$ covers the pairs $\{x, y\}$ with $x \in X$, $y \in Y$
- ε-WSPD: a family $\mathcal{W}$ of ε-separated dumbbells covering all $\binom{n}{2}$ pairs; size $|\mathcal{W}|$
- Partition WSPD: every pair covered exactly once; cover WSPD: pairs may be covered several times



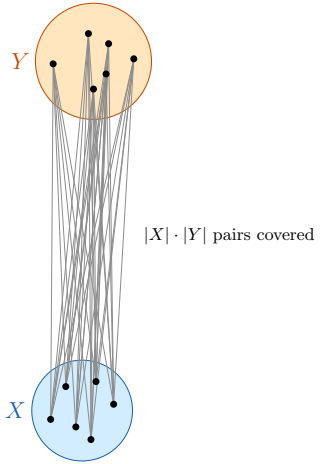
Well-Separated Pair Decomposition

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

Definitions:

- P: set of n points in $\mathbb{R}^d$, $\varepsilon \in (0,1)$
- Dumbbell $\{X, Y\}$: two disjoint nonempty subsets of P, its sides
- ε-separated: $\max\{\mathsf{diam}(X), \mathsf{diam}(Y)\} \leq \varepsilon \cdot \mathsf{dist}(X, Y)$

- A pair is $\{p, q\} \subseteq P$, $p \neq q$
- $\{X, Y\}$ covers the pairs $\{x, y\}$ with $x \in X$, $y \in Y$
- ε-WSPD: a family $\mathcal{W}$ of ε-separated dumbbells covering all $\binom{n}{2}$ pairs; size $|\mathcal{W}|$
- Partition WSPD: every pair covered exactly once; cover WSPD: pairs may be covered several times



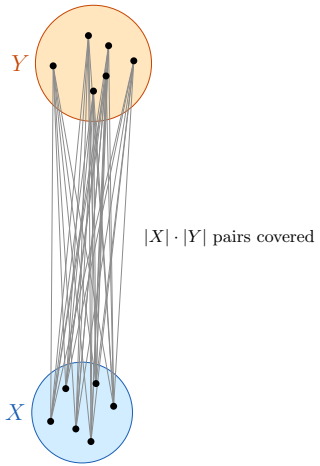
Well-Separated Pair Decomposition

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

Definitions:

- P: set of n points in $\mathbb{R}^d$, $\varepsilon \in (0,1)$
- Dumbbell $\{X, Y\}$: two disjoint nonempty subsets of P, its sides
- ε-separated: $\max\{\mathsf{diam}(X), \mathsf{diam}(Y)\} \leq \varepsilon \cdot \mathsf{dist}(X, Y)$

- A pair is $\{p, q\} \subseteq P$, $p \neq q$
- $\{X, Y\}$ covers the pairs $\{x, y\}$ with $x \in X$, $y \in Y$
- ε-WSPD: a family $\mathcal{W}$ of ε-separated dumbbells covering all $\binom{n}{2}$ pairs; size $|\mathcal{W}|$
- Partition WSPD: every pair covered exactly once; cover WSPD: pairs may be covered several times



Cover or Partition?

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- Callahan–Kosaraju (CK) [CK95]:
  partition ε-WSPD of size $O(n/\varepsilon^d)$ in $O(n \log n + n/\varepsilon^d)$ time

A cover WSPD is enough for:

- t-spanners with $|\mathcal{W}|$ edges [CK93,NS07,Smi18]
- Approximate MST [CK93,Smi18]
- Sparse edge sets for geometric tour solvers [FKKN24]
- Approximate diameter [Har11,Smi18]

A partition WSPD is needed for:

- n-body potential fields [CK95]
- Approximate distance selection [BS02,Smi18]
- Approximate sum of all pairwise distances [Har11]

- Surveys: [NS07,Smi18]
- From now on, WSPD means cover WSPD, as proposed in [BCH+26]

Minimum WSPD

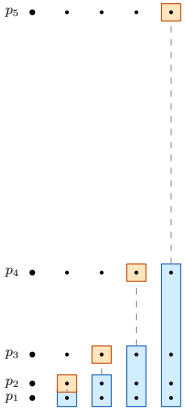
Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- $O(n/\varepsilon^d)$ is a worst-case bound
- Many inputs admit much smaller WSPDs
- Spanners and other structures built from a WSPD inherit its size

Problem MINWSPD:

Given P and $\varepsilon \in (0, 1)$, compute an ε-WSPD of minimum size OPT_ε.

- A partition WSPD is a cover WSPD: OPT_ε is at most the partition optimum, so every factor w.r.t. OPT_ε also holds w.r.t. the partition optimum



Exponentially spread points:
$n - 1$ dumbbells
$\{\{p_1, \ldots, p_i\}, \{p_{i+1}\}\}$

What Was Known

- MINWSPD is NP-hard in the plane when ε depends on n [BCH+26]
- General metrics: no n^c-approximation for some $c > 0$, unless P = NP [BCH+26]
- Greedy set cover over all maximal ε-separated dumbbells: $O(\log n)$-approximation if their number is polynomial; known on the line, open for $d \geq 2$ [BCH+26]
- $\mathbb{R}^d$: the net-tree WSPD [HM06] is a $2^{O(d)}$-approximation [BCH+26], the only one stated for the separation ε
  The factor is unspecified; we estimate it at 10^{11} for $d = 2$
- Every metric: size $7\,\text{OPT}_\varepsilon$, but separation relaxed to 33ε [HRR25]
  In $\mathbb{R}^2$, cutting the sides with a grid restores ε: factor $7 \cdot 47^4 \approx 3.4 \cdot 10^7$
- Bicriteria: size $(1+\delta)\text{OPT}_\varepsilon$, separation $(1+\delta)\varepsilon$, time doubly exponential in d/δ [Fon26]

Question [BCH+26]:

Is there a $(1+\delta)$-approximation, or a simple $O(1)$-approximation, in $\mathbb{R}^d$ for $d \geq 2$?

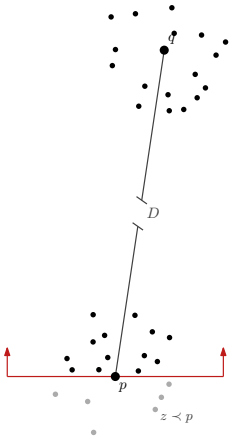
Our Results

Setting	Problem	Size	Separation	Time	Reference
$\mathbb{R}^1$	cover	$(1+\delta)\,\mathrm{OPT}_\varepsilon$	ε	$n^{O(1/\delta^2)}$	[BCH+26]
$\mathbb{R}^1$	cover	$3\,\mathrm{OPT}_\varepsilon$	ε	$O(n \log n)$	[BCH+26]
$\mathbb{R}^1$	both	$10\,\mathrm{OPT}_\varepsilon$	ε	polynomial	[BCH+26]
$\mathbb{R}^d$	both	$2^{O(d)}\,\mathrm{OPT}_\varepsilon$	ε	$2^{O(d)}(n \log n + \mathrm{OPT})$ exp.	[BCH+26,HM06]
$\mathbb{R}^d$	cover	$(1+\delta)\,\mathrm{OPT}_\varepsilon$	$(1+\delta)\varepsilon$	$f(d,\delta)\,n^6$	[Fon26]
Any metric	cover	$7\,\mathrm{OPT}_\varepsilon$	33ε	polynomial	[HRR25]
$\mathbb{R}^2$, $\varepsilon \le 1/16$	cover	$40\,\mathrm{OPT}_\varepsilon$	ε	$O(n^2 \log n)$	sweep
$\mathbb{R}^2$, $\varepsilon \le 1/5$	both	$1024\,\mathrm{OPT}_\varepsilon$	ε	$O(n \log n + \mathrm{OPT})$	CK, fair split tree
$\mathbb{R}^2$, $\varepsilon \le 1/5$	both	$529\,\mathrm{OPT}_\varepsilon$	ε	$O(n \log n + \mathrm{OPT})$	quadtree
$\mathbb{R}^2$, $\varepsilon \le 1/5$	both	$252\,\mathrm{OPT}_\varepsilon$ exp.	ε	$O(n \log n + \mathrm{OPT})$	random quadtree
$\mathbb{R}^d$, $\varepsilon \le 1/16$	cover	$2^{O(d)}\,\mathrm{OPT}_\varepsilon$	ε	$O(n^2 \log n)$	sweep

- *Both*: the output is a partition WSPD; $\mathbb{R}^3$: factor 2048
- Lower bounds: sweep steps, enlarged by local search to $\mathrm{LB} \ge 0.99\,\mathrm{OPT}$ for $n = 500$
- Local search *reshape*: never increases the size, uses no solver
- $\varepsilon = 1/20$, $n = 500$: sweep $1.03\,\mathrm{OPT}$, CK $1.17\,\mathrm{OPT}$, sweep + reshape within 0.02% of OPT
  $n = 10000$: sweep $1.22\,\mathrm{LB}$, CK $1.55\,\mathrm{LB}$, sweep + reshape $1.05\,\mathrm{LB}$

The Sweep

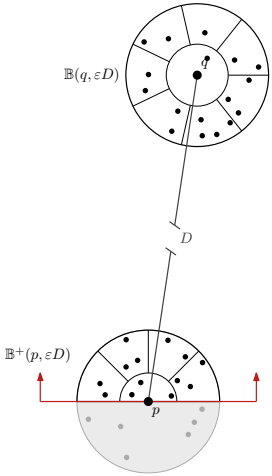
- Process $p \in P$ in lexicographic order $\prec$ (by y, then x), and the points $q \succ p$
- If $\{p, q\}$ is uncovered: a step with anchor $\{p, q\}$, $D = \|pq\|$
- Pieces: 5 in the half-disk $\mathbb{B}^+(p, \varepsilon D)$ (diameter $\leq 0.8\,\varepsilon D$), 8 in the disk $\mathbb{B}(q, \varepsilon D)$ (diameter $\leq 0.868\,\varepsilon D$)
- Output every dumbbell $\{A_i, C_k\}$ of two nonempty pieces: at most $5 \cdot 8 = 40$
- ε-separated: diam $\leq 0.868\,\varepsilon D$ and dist $\geq (1 - 2\varepsilon)D \geq 0.868\,D$ for $\varepsilon \leq 1/16$



The Sweep

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

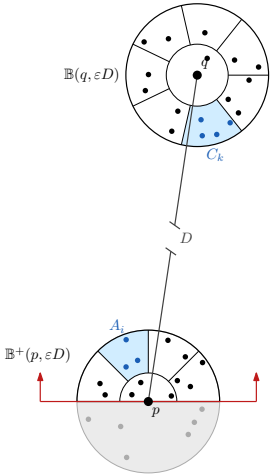
- Process $p \in P$ in lexicographic order $\prec$ (by y, then x), and the points $q \succ p$
- If $\{p, q\}$ is uncovered: a step with anchor $\{p, q\}$, $D = \|pq\|$
- Pieces: 5 in the half-disk $\mathbb{B}^+(p, \varepsilon D)$ (diameter $\leq 0.8\,\varepsilon D$), 8 in the disk $\mathbb{B}(q, \varepsilon D)$ (diameter $\leq 0.868\,\varepsilon D$)
- Output every dumbbell $\{A_i, C_k\}$ of two nonempty pieces: at most $5 \cdot 8 = 40$
- ε-separated: diam $\leq 0.868\,\varepsilon D$ and dist $\geq (1 - 2\varepsilon)D \geq 0.868\,D$ for $\varepsilon \leq 1/16$



The Sweep

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- Process $p \in P$ in lexicographic order $\prec$ (by y, then x), and the points $q \succ p$
- If $\{p, q\}$ is uncovered: a step with anchor $\{p, q\}$, $D = \|pq\|$
- Pieces: 5 in the half-disk $\mathbb{B}^+(p, \varepsilon D)$ (diameter $\leq 0.8\,\varepsilon D$), 8 in the disk $\mathbb{B}(q, \varepsilon D)$ (diameter $\leq 0.868\,\varepsilon D$)
- Output every dumbbell $\{A_i, C_k\}$ of two nonempty pieces: at most $5 \cdot 8 = 40$
- ε-separated: diam $\leq 0.868\,\varepsilon D$ and dist $\geq (1 - 2\varepsilon)D \geq 0.868\,D$ for $\varepsilon \leq 1/16$



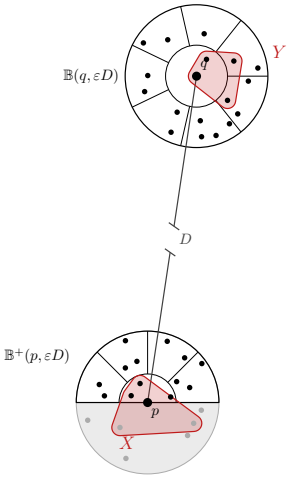
Approximation Factor 40

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- Invariant: at the step with anchor $\{p, q\}$, every pair $\{x, y\}$ with $\min_{\prec}\{x, y\} \prec p$ is covered
- Locality: if $\{X, Y\}$ is ε-separated with $\ell = \mathsf{dist}(X, Y)$, each side lies in the disk of radius $\varepsilon\ell$ around any of its points
- Charge the step to an optimal $Z = \{X, Y\}$ covering $\{p, q\}$: $X \subseteq \mathbb{B}(p, \varepsilon D)$, $Y \subseteq \mathbb{B}(q, \varepsilon D)$
- Uncovered pairs of Z have $x \succeq p$: all covered by the step
- Z is charged once

Theorem:

For $P \subset \mathbb{R}^2$ and $\varepsilon \leq 1/16$, the sweep performs at most OPT_ε steps and outputs an ε-WSPD of size at most $40\,\mathrm{OPT}_\varepsilon$, in $O(n^2 \log n)$ time and $O(n^2)$ space.



The Number of Steps Is a Lower Bound

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- Pairs a, b are compatible if one ε-separated dumbbell covers both
- Heredity: subdumbbells of ε-separated dumbbells are ε-separated
  $\Rightarrow$ compatibility is tested on 4 points
- R_{pq}: pairs $\{x, y\}$ uncovered at the step, compatible with $\{p, q\}$, with $x, y \succeq p$

Observation:

If every step covers R_{pq}, then the anchors are pairwise incompatible, so there are at most OPT_ε steps.

- A later anchor $\{p', q'\}$ has $p', q' \succeq p$ and was uncovered: if compatible, it would lie in R_{pq}
- The fixed collection of 40 dumbbells covers R_{pq} for every $\varepsilon < 1/2$
- The number of steps is a lower bound LB_0 on OPT_ε, enlarged later by local search

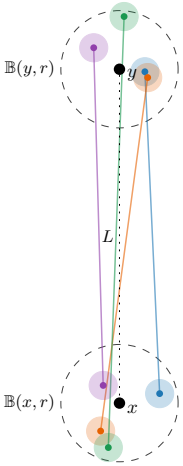
Running Time

- Per point, the list of all points sorted by distance: $O(n^2 \log n)$ time, $O(n^2)$ space
- A step scans two prefixes of these lists; every scanned point lies in an output dumbbell
- $n \times n$ table of covered pairs

Lemma:

For $\varepsilon \leq 1/16$, every pair is covered by the output of at most 360 steps, and by one output dumbbell per step.

- Anchors $\{p_j, q_j\}$: $p_j \in \mathbb{B}(x, r)$, $q_j \in \mathbb{B}(y, r)$
- Incompatible: $\max\{\|p_j p_k\|, \|q_j q_k\|\} > s \geq 3r/4$
- Products $\mathbb{B}(p_j, \frac{s}{2}) \times \mathbb{B}(q_j, \frac{s}{2}) \subset \mathbb{R}^4$ are disjoint: at most $(11/3)^4 < 181$ per orientation



The Greedy Sweep

- The fixed sweep outputs a dumbbell for every two nonempty pieces, even if one dumbbell covers R_{pq}: of theoretical interest only
- The proofs only use that a step covers R_{pq} with at most 40 ε-separated dumbbells

Greedy step:

- Grow a dumbbell from a seed $\{a, b\}$: add a point within $\varepsilon\|ab\|$ of one side and at least $\|ab\|$ from the other, maximizing the newly covered pairs of R_{pq}
- Seeds: all pairs of R_{pq} if $|R_{pq}| \leq 16$ (then greedy set cover), otherwise $\{p, q\}$ and then the closest uncovered pair of R_{pq}
- Fall back to the fixed collection if it is smaller (cap: 40 grown dumbbells)
- At the end, delete redundant dumbbells one at a time

Theorem:

For $\varepsilon \leq 1/16$, the greedy sweep performs at most OPT_ε steps and outputs an ε-WSPD of size at most $40\,\mathrm{OPT}_\varepsilon$ in $O(n^3)$ time and $O(n^2)$ space.

- In the experiments, "the sweep" is the greedy sweep

Extensions

Higher dimensions:

- k pieces in the ball, h in the half-ball, diameter $\leq (1-2\varepsilon)\rho$: factor kh
- Grid pieces: $64 \cdot 32 = 2048$ for $d = 3$; net pieces: less than 32^d
- Same time $O(n^2 \log n)$ (volumes in $\mathbb{R}^{2d}$); greedy sweep $O(n^3)$
- Isodiametric inequality: this analysis always gives $kh > 2^{2d-1}$

Larger ε:

- Pieces that depend on ε, with a separation test for each output dumbbell
- Constant factor for every $\varepsilon > 0$: less than $4 \cdot 10^5 \max\{1, \varepsilon^8\}$
- The number of steps is a lower bound for every ε

Partition WSPD Constructions

- Quadtree: maximal ε-separated pairs of same-level cells
- CK on a fair split tree [CK95], testing bounding boxes (CK box) or point sets (CK)

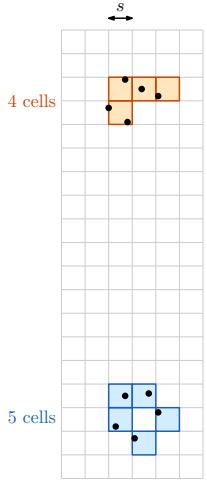
Theorem:

For $P \subset \mathbb{R}^2$ and $\varepsilon \leq 1/5$, these partition ε-WSPDs have size at most

- $529\,\mathrm{OPT}_\varepsilon$ (quadtree), $252\,\mathrm{OPT}_\varepsilon$ in expectation (randomly shifted and scaled quadtree)
- $1024\,\mathrm{OPT}_\varepsilon$ (CK, both tests; $576\,\mathrm{OPT}_\varepsilon$ for $\varepsilon \leq 1/16$)

in $O(n \log n + \mathrm{OPT}_\varepsilon)$ time, with dumbbells as pairs of tree nodes.

- Charge to an optimal $\{X, Y\}$: at the level with cell diameter $\approx \varepsilon\,\mathsf{dist}(X, Y)$, a side meets at most 23 cells (5×5 block without two opposite corners): $23^2 = 529$
- Random: expected number of cells met, averaged over the shift and the scale



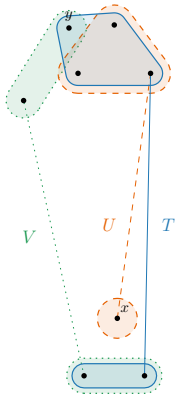
Local Search: Reshape

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- Improves any given ε-WSPD; never increases the size
- Private pair of T: only T covers it; no private pairs: redundant
- Free growth: p joins X if $\{X \cup \{p\}, Y\}$ stays ε-separated; dist may decrease

Reshape a target T:

1. Shrink T to the subdumbbell spanned by its private pairs
2. Grow it freely, favoring pairs private to other dumbbells
3. Delete the dumbbells that became redundant



$\varepsilon = 1/3$.
With y in T, x cannot join T.

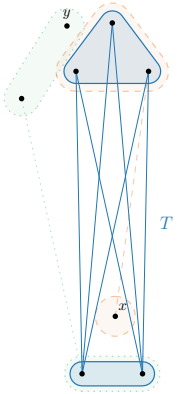
Local Search: Reshape

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- Improves any given ε-WSPD; never increases the size
- Private pair of T: only T covers it; no private pairs: redundant
- Free growth: p joins X if $\{X \cup \{p\}, Y\}$ stays ε-separated; dist may decrease

Reshape a target T:

1. Shrink T to the subdumbbell spanned by its private pairs
2. Grow it freely, favoring pairs private to other dumbbells
3. Delete the dumbbells that became redundant



$\varepsilon = 1/3$.
Private pairs of T; y is released.

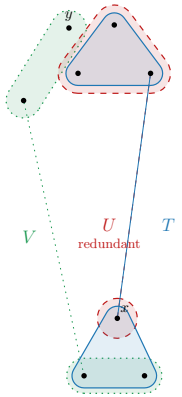
Local Search: Reshape

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- Improves any given ε-WSPD; never increases the size
- Private pair of T: only T covers it; no private pairs: redundant
- Free growth: p joins X if $\{X \cup \{p\}, Y\}$ stays ε-separated; dist may decrease

Reshape a target T:

1. Shrink T to the subdumbbell spanned by its private pairs
2. Grow it freely, favoring pairs private to other dumbbells
3. Delete the dumbbells that became redundant



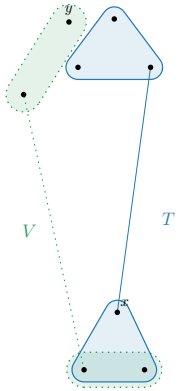
$\varepsilon = 1/3$.
x joins; U is now redundant.

Local Search: Reshape

- Improves any given ε-WSPD; never increases the size
- Private pair of T: only T covers it; no private pairs: redundant
- Free growth: p joins X if $\{X \cup \{p\}, Y\}$ stays ε-separated; dist may decrease

Reshape a target T:

1. Shrink T to the subdumbbell spanned by its private pairs
2. Grow it freely, favoring pairs private to other dumbbells
3. Delete the dumbbells that became redundant



$\varepsilon = 1/3$.
One dumbbell fewer.

Driver and Comparison

- Targets in passes over $\mathcal{W}$, each pass in a new random order, until the time budget is exhausted
- Every reshape keeps $\mathcal{W}$ a valid ε-WSPD
- Works directly on the points: no solver, never builds the set-cover instance
- No guarantee of its own, but never worse than the start: factor 40 from the sweep, 576 from CK ($\varepsilon \leq 1/16$)

For comparison, large neighborhood search (LNS):

- Remove T and its k nearest dumbbells: a set U of uncovered pairs
- Re-cover U exactly with CPLEX, among the removed dumbbells and, if $|U| \leq u$, maximal ε-separated dumbbells whose closest cross pair is in U
- Accept only if the size decreases
- Tuned: $(k, u) = (16, 1000)$ for $n \leq 1000$, $(32, 400)$ for larger n

Implementation

- Three global questions, each answered in $O(1)$ per pair or point:

1. Is a pair private? $n \times n$ table of coverage counts
2. Which dumbbells become redundant? Only owners of pairs whose count goes from 1 to 2
   - Second $n \times n$ table: XOR of the ids of the dumbbells covering the pair
   - Insert/delete dumbbell k: apply $\oplus k$, in the same loop as the counts
   - Count $1 \Rightarrow$ the entry is the id of the owner
   - No point-to-dumbbell lists: $> 10\times$ faster for $n = 10000$
3. Which points can join? Candidates within $\varepsilon\mathsf{dist}(X, Y)$ from a kd-tree
   - A point that cannot join never can later: discard it for good

- Points relabeled in Hilbert order: the pairs of a dumbbell are close in memory

Local Search for Larger Lower Bounds

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

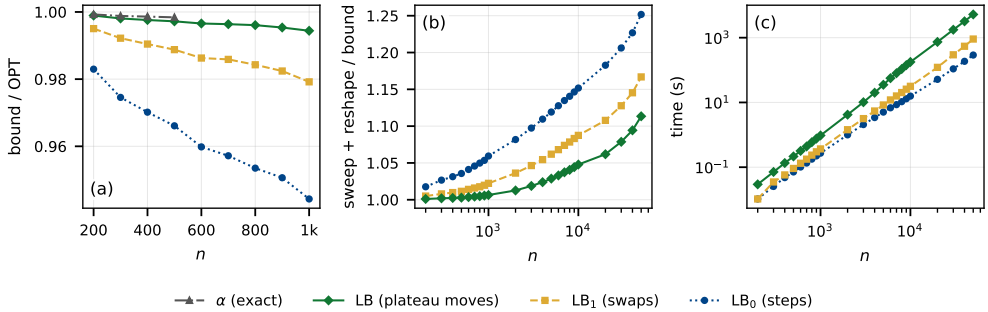
- Compatibility graph G_ε: the pairs, with an edge between compatible pairs
- Every independent set I of G_ε gives $|I| \leq \mathrm{OPT}_\varepsilon$; the anchors give LB_0
- For I, a pair is free if it has no neighbor in I, tight for $e \in I$ if e is its only neighbor [ARW12]

Start from the anchors of the sweep:

1. Add free pairs until none is left
2. 1-for-2 swaps [ARW12]: replace $e \in I$ by two incompatible pairs tight for e, add the new free pairs; repeat until no swap: LB_1
3. 30 more passes with plateau moves [PH06,GLP08]: replace e by a random tight pair with probability $1/2$: LB

- Pairs compatible with $\{x, y\}$: one point in $\mathbb{B}(x, \varepsilon\|xy\|)$, the other in $\mathbb{B}(y, \varepsilon\|xy\|)$: prefixes of the sorted lists
- For $\varepsilon \leq 1/16$: $O(n^2 \log n + kn^3)$ time for k passes, $k \leq \mathrm{LB}_1 - \mathrm{LB}_0 + 31$

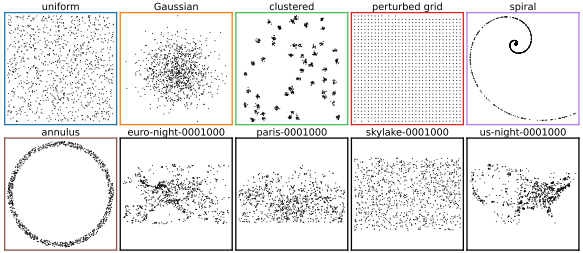
Lower Bounds in Practice



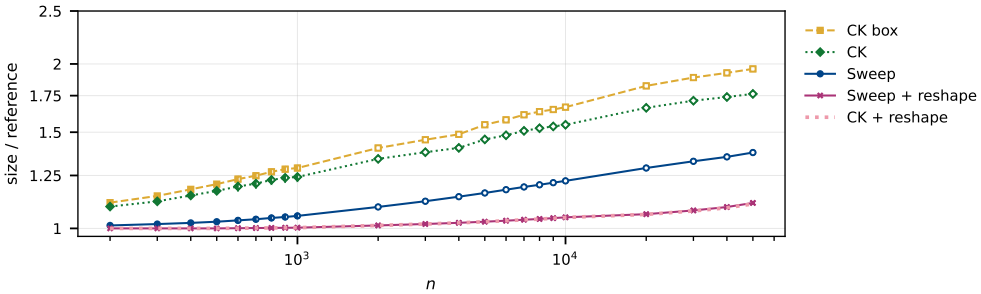
- Where OPT is known: LB is 0.999, 0.997, $0.994\,\mathrm{OPT}$ for $n = 200$, 500, 1000 (at least 0.989; LB_0: 0.98, 0.97, 0.94)
- Independence number α of G_ε by CPLEX ($n \leq 500$): at least $0.992\,\mathrm{OPT}$; $LB = \alpha$ on 31 of 88
- LB/LB_0: 1.05, 1.10, 1.12 for $n = 1000$, 10000, 50000
- After the sweep, LB takes 2.7 min for $n = 10000$ and 82 min for $n = 50000$

Setup

- C++, integer coordinates, $\varepsilon = 1/20$ rational: exact separation tests; every output checked by a brute-force verifier
- Sweep: the greedy sweep; CK and CK box on a fair split tree
- Inputs: 6 synthetic families and the non-uniform CG:SHOP 2019 instances [DFK+22], $n = 200, \ldots, 50000$
- Reference: OPT where CPLEX proves it (all $n \leq 500$, some $n \leq 1000$), otherwise LB
- Code and data: `github.com/gfonsecabr/small-wspd-experiments`

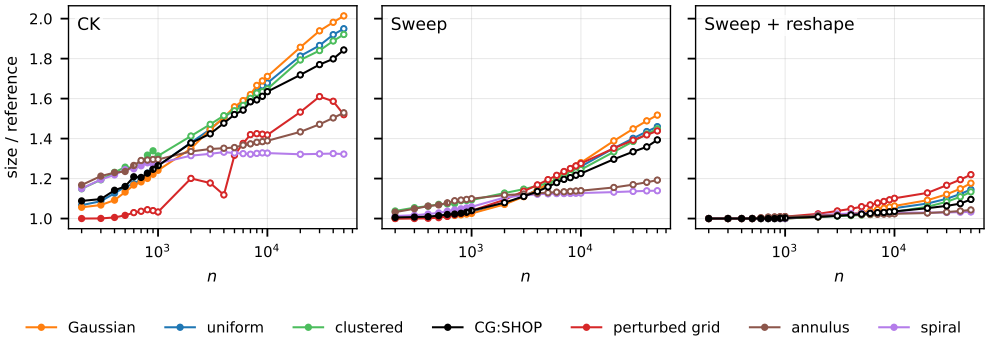


Quality



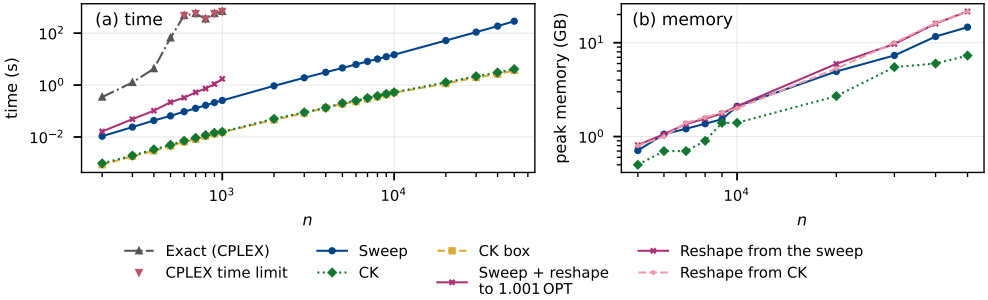
- Open markers: the reference is LB (half-filled: on some instances)
- $n = 500$: sweep $+3\%$ (at most 9%), CK $+17\%$, CK box $+21\%$ over OPT
- $n = 10000$: sweep $1.22\,\mathrm{LB}$, CK $1.55\,\mathrm{LB}$; $n = 50000$: $1.38\,\mathrm{LB}$ and $1.76\,\mathrm{LB}$
- The sweep is smallest on every input and n, except the perturbed grid for $n = 200$, 1000, 4000

Quality per Input



- Mean size divided by the reference, per input
- Sweep + reshape, $n = 50000$: from 23 dumbbells per point (spiral) to 981 (perturbed grid)

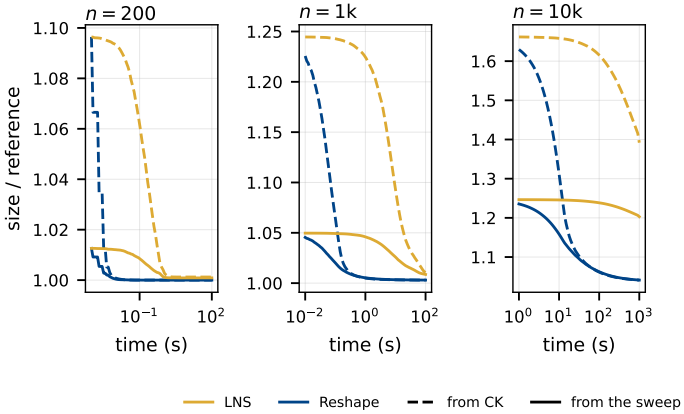
Running Times



- CK: 0.5 s for $n = 10000$, 4 s for $n = 50000$
- Sweep: 0.3 s for $n = 1000$, 15 s for $n = 10000$, 5 min for $n = 50000$
- $n = 50000$: CK, the sweep, and reshape use up to 8, 15, and 22 GB
- OPT: 15 s median for $n = 1000$ where CPLEX proves optimality

Local Search over Time

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography



- Reshape ends at most as large as LNS from the same start on every instance, and reaches the final size of LNS within 1% of the budget
- LNS improves CK from 1.66 to 1.39 times the reference for $n = 10000$ (1000 s), and from 1.89 to 1.82 for $n = 50000$ (5000 s)

Local Search: Summary

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

Sweep + reshape, $n \leq 1000$ with known OPT (164 instances):

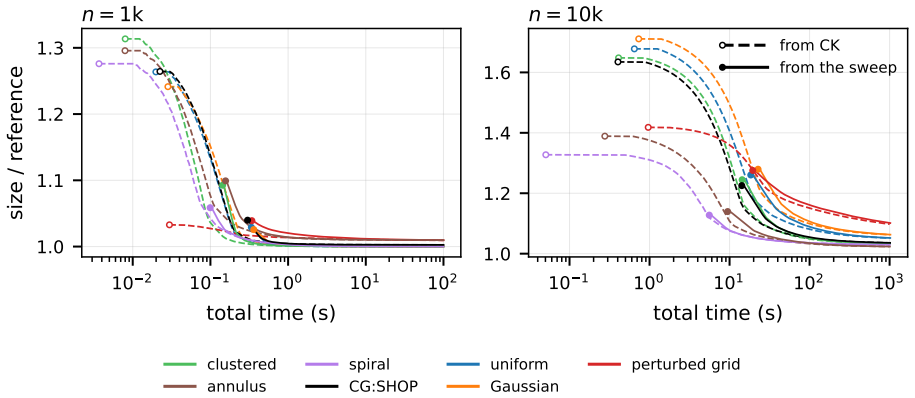
- On average 0.002%, at most 0.02% above OPT
- Reaches OPT on 101: all 22 for $n = 200$, 13 of 22 for $n = 500$, 1 of 13 for $n = 1000$
- LNS from the sweep: 0.23% on average, up to 1.5%; reaches OPT on 18

Larger n:

- $n = 10000$: $1.05\,\mathrm{LB}$ on average, at most $1.10\,\mathrm{LB}$ (sweep alone: $1.22\,\mathrm{LB}$)
- $n = 50000$: $1.11\,\mathrm{LB}$ on average, at most $1.22\,\mathrm{LB}$ (sweep alone: $1.38\,\mathrm{LB}$)

- Budgets: 100 s for $n \leq 1000$, 1000 s for $n \leq 10000$, 5000 s above

Which Start?



- Same total time (build + reshape): best is CK + reshape, although CK is larger than the sweep
- It beats the sweep's output before the sweep ends (418 of 436 instances); afterwards, sweep + reshape is at most 0.73% smaller
- But the proven factor is 576 for CK + reshape, against 40 for sweep + reshape

From Cover to Partition

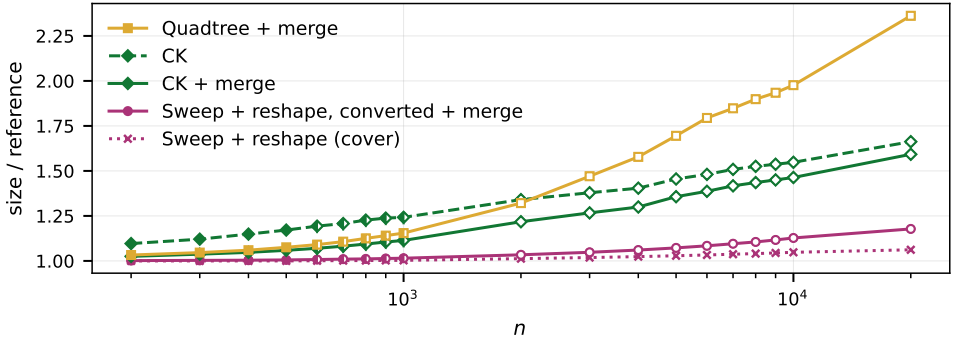
Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

- Process the dumbbells of a cover WSPD one by one
- U: pairs of the current $\{X, Y\}$ that no earlier dumbbell covers
- Group the points of X with the same partners in Y among U: each group X' with partners Y' gives $\{X', Y'\}$, separated by heredity (same with X and Y exchanged; keep the smaller split)
- Order: fewest pieces first (lazy priority queue), ties by size
- Merge pass: replace $\{S, O_1\}$, $\{S, O_2\}$ by $\{S, O_1 \cup O_2\}$ if ε-separated; applies to every partition WSPD (CK, quadtree)
- No size guarantee

Optimal WSPDs, $n = 500$:

Converted, they give partition WSPDs of size at most $1.0125\,\mathrm{OPT}_\varepsilon$, which bounds the partition optimum on these instances

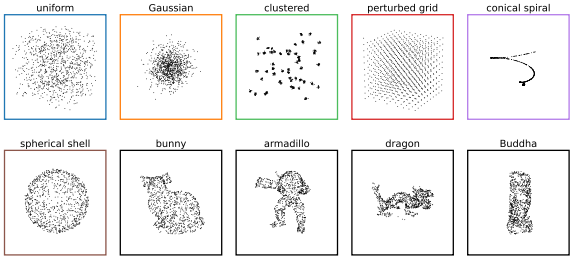
Partition WSPDs



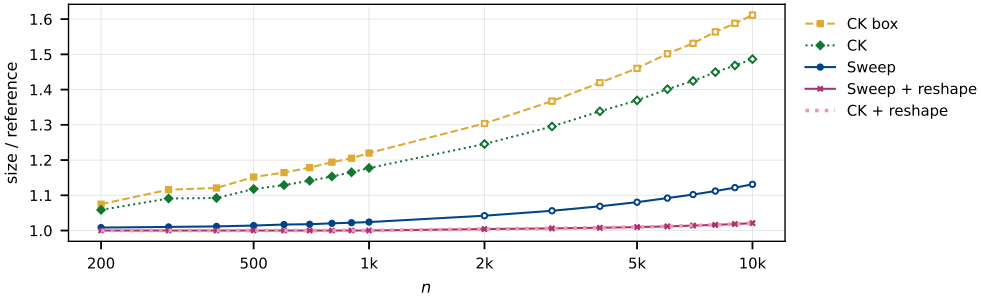
- Sweep + reshape, converted: 1.015, 1.13, 1.18 times the reference for $n = 1000$, 10000, 20000 (cover: 1.003, 1.05, 1.06)
- CK: 1.24, 1.55, 1.66; after merging 1.11, 1.46, 1.59; never smaller than the converted WSPD
- Quadtree: larger than CK despite its smaller proven factor

Experiments in Three Dimensions

- Separate implementation; sweep with the grid pieces (64 cubes in the ball, 32 in the half-ball)
- Inputs: 6 synthetic families and samples of 4 Stanford scans [Sta], $n = 200, \ldots, 10000$, 180 instances
- OPT proven on all 90 instances with $n \leq 1000$; there, $\mathrm{LB} \geq 0.986\,\mathrm{OPT}$



Results in Three Dimensions



- $n = 500$: sweep $+1.4\%$, CK $+12\%$, CK box $+15\%$ over OPT
- $n = 10000$: sweep $1.13\,\mathrm{LB}$, CK $1.49\,\mathrm{LB}$, CK box $1.61\,\mathrm{LB}$
- Reshape: at most 0.03% above OPT where known; $1.021\,\mathrm{LB}$ for $n = 10000$
- Converted partition WSPD: $1.055\,\mathrm{LB}$ for $n = 10000$, against 1.49 for CK

Open Problems

Open problem:

The approximation factor of the greedy sweep for small ε is between 1.503 (observed for `jupiter-0020000`, $\varepsilon = 1/20$) and 40 (proved for $\varepsilon \leq 1/16$). What is it?

- Smaller proven factors in the plane, for cover and partition WSPDs?
- The greedy sweep in $O(n^2 \log n)$ time?
- An approximation guarantee for reshape?
  At most 0.02% above OPT where OPT is known
- How close is LB to OPT for large n?
- A PTAS that keeps the separation ε [BCH+26]?

Bibliography

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

[ARW12] D. V. Andrade, M. G. C. Resende, and R. F. Werneck. Fast local search for the maximum independent set problem. J. Heuristics, 18:525–547, 2012.

[BCH+26] K. Buchin, J. Conradi, S. Har-Peled, A. Kalb, A. Lahiri, L. Plätz, C. Rehs, and S. Wong. On small pair decompositions for point sets. Proc. 34th ESA, 102:1–102:22, 2026.

[BS02] S. Bespamyatnikh and M. Segal. Fast algorithms for approximating distances. Algorithmica, 33:263–269, 2002.

[CK93] P. B. Callahan and S. R. Kosaraju. Faster algorithms for some geometric graph problems in higher dimensions. Proc. 4th SODA, 291–300, 1993.

[CK95] P. B. Callahan and S. R. Kosaraju. A decomposition of multidimensional point sets with applications to k-nearest-neighbors and n-body potential fields. J. ACM, 42:67–90, 1995.

[DFK+22] E. D. Demaine, S. P. Fekete, P. Keldenich, D. Krupke, and J. S. B. Mitchell. Area-optimal simple polygonalizations: The CG Challenge 2019. ACM J. Exp. Algorithmics, 27:2.4:1–2.4:12, 2022.

[FKKN24] S. P. Fekete, P. Keldenich, D. Krupke, and E. Niehs. Edge sparsification for geometric tour problems. Comput. Geom. Topol., 3:1:1–1:23, 2024.

[Fon26] G. D. da Fonseca. A bicriteria approximation scheme for the minimum well-separated pair decomposition. arXiv:2610.13650, 2026.

Bibliography (cont.)

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography

[GLP08] A. Grosso, M. Locatelli, and W. Pullan. Simple ingredients leading to very efficient heuristics for the maximum clique problem. J. Heuristics, 14:587–612, 2008.

[Har11] S. Har-Peled. Geometric Approximation Algorithms. American Mathematical Society, 2011.

[HM06] S. Har-Peled and M. Mendel. Fast construction of nets in low-dimensional metrics and their applications. SIAM J. Comput., 35:1148–1184, 2006.

[HRR25] S. Har-Peled, B. Raichel, and E. W. Robson. Well-separated pairs decomposition revisited. arXiv:2509.05997, 2025.

[NS07] G. Narasimhan and M. Smid. Geometric Spanner Networks. Cambridge University Press, 2007.

[PH06] W. Pullan and H. H. Hoos. Dynamic local search for the maximum clique problem. J. Artif. Intell. Res., 25:159–185, 2006.

[Smi18] M. Smid. The well-separated pair decomposition and its applications. Handbook of Approximation Algorithms and Metaheuristics, 2nd ed., 2018.

[Sta] Stanford Computer Graphics Laboratory. The Stanford 3D Scanning Repository. graphics.stanford.edu/data/3Dscanrep

Problem
Background
Results
Sweep
Partition
Local Search
Lower Bounds
Experiments
In 3D
Open Problems
Bibliography



Thank you!