An Efficient Representation for Filtrations of Simplicial Complexes

Karthik C. S.

Weizmann Institute of Science.

Joint work with Jean-Daniel Boissonnat (INRIA).

Simplicial Complex (definition)

$V = \{1, 2, 3, 4, 5, 6\}$

Simplicial Complex (definition)

$V = \{1, 2, 3, 4, 5, 6\}$

A simplicial complex $K \subseteq \mathcal{P}(V)$:

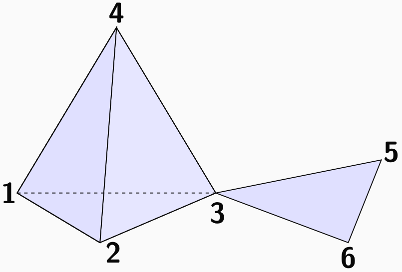
- ◇ $p \in V \Rightarrow \{p\} \in K$.
- ◇ $\sigma \in K, \tau \subseteq \sigma \Rightarrow \tau \in K$.

Simplicial Complex (definition)

$V = \{1, 2, 3, 4, 5, 6\}$

A simplicial complex $K \subseteq \mathcal{P}(V)$:

- $p \in V \Rightarrow \{p\} \in K$.
- $\sigma \in K, \tau \subseteq \sigma \Rightarrow \tau \in K$.

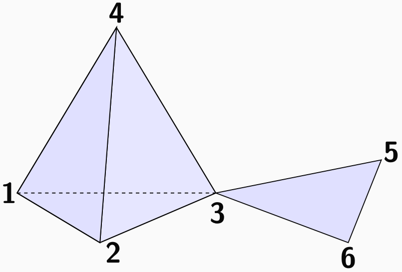


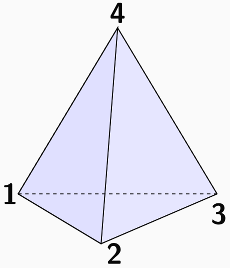
Simplicial Complex (definition)

$V = \{1, 2, 3, 4, 5, 6\}$

A simplicial complex $K \subseteq \mathcal{P}(V)$:

- ◇ $p \in V \Rightarrow \{p\} \in K$.
- ◇ $\sigma \in K, \tau \subseteq \sigma \Rightarrow \tau \in K$.



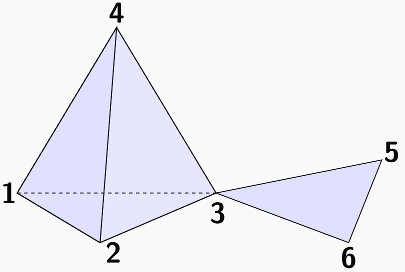


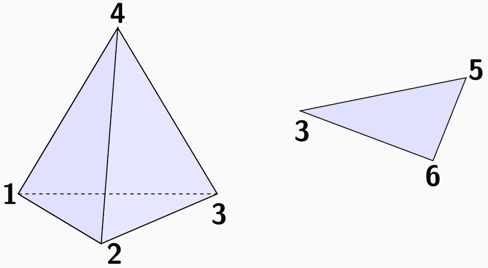
Simplicial Complex (definition)

$V = \{1, 2, 3, 4, 5, 6\}$

A simplicial complex $K \subseteq \mathcal{P}(V)$:

- $\diamond$ $p \in V \Rightarrow \{p\} \in K$.
- $\diamond$ $\sigma \in K, \tau \subseteq \sigma \Rightarrow \tau \in K$.



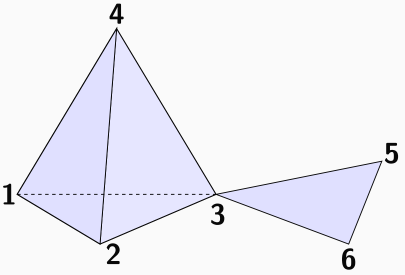


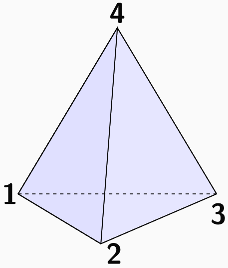
Simplicial Complex (definition)

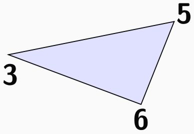
$V = \{1, 2, 3, 4, 5, 6\}$

A simplicial complex $K \subseteq \mathcal{P}(V)$:

- ◇ $p \in V \Rightarrow \{p\} \in K$.
- ◇ $\sigma \in K, \tau \subseteq \sigma \Rightarrow \tau \in K$.







6 vertices
3 dimensional
2 maximal simplices
22 simplices

Data structures for simplicial complexes:

Data structures for simplicial complexes:

⋆ Hasse Diagram

Representation of Simplicial Complexes

Data structures for simplicial complexes:

- ~~Hasse Diagram~~
- Simplex Tree [Boissonnat and Maria, 2014]

Representation of Simplicial Complexes

Data structures for simplicial complexes:

- ~~Hasse Diagram~~
- Simplex Tree [Boissonnat and Maria, 2014]
- Skeleton Blockers [Attali, Lieutier, Salinas, 2012]

Representation of Simplicial Complexes

Data structures for simplicial complexes:

- ⋆ ~~Hasse Diagram~~
- ⋆ Simplex Tree [Boissonnat and Maria, 2014]
- ⋆ Skeleton Blockers [Attali, Lieutier, Salinas, 2012]
- ⋆ Simplex Array List [Boissonnat, K, Tavenas, 2016]

filtration of

Data structures for ↑simplicial complexes:

- ⋆ ~~Hasse Diagram~~
- ⋆ Simplex Tree [Boissonnat and Maria, 2014]
- ⋆ Skeleton Blockers [Attali, Lieutier, Salinas, 2012]
- ⋆ Simplex Array List [Boissonnat, K, Tavenas, 2016]

Representation of Simplicial Complexes

filtration of

Data structures for ↑simplicial complexes:

- ⋆ ~~Hasse Diagram~~
- ⋆ Simplex Tree [Boissonnat and Maria, 2014]
- ⋆ ~~Skeleton Blockers [Attali, Lieutier, Salinas, 2012]~~
- ⋆ ~~Simplex Array List [Boissonnat, K, Tavenas, 2016]~~

Filtration (definition)

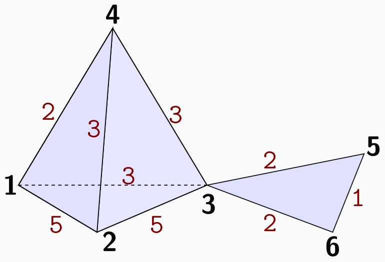
A filtration $f : K \to \mathbb{R}$,

$f(\sigma) \leq f(\tau)$ if $\sigma \subseteq \tau$

Filtration (definition)

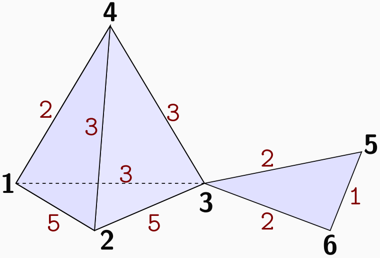
A filtration $f : K \to \mathbb{R}$,

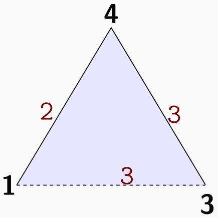
$f(\sigma) \leq f(\tau)$ if $\sigma \subseteq \tau$



A filtration $f : K \to \mathbb{R}$,

$f(\sigma) \leq f(\tau)$ if $\sigma \subseteq \tau$

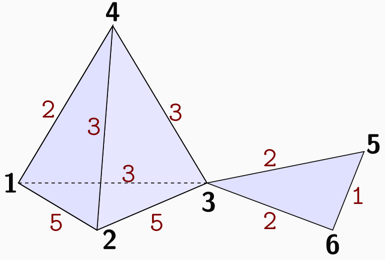


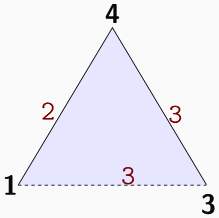


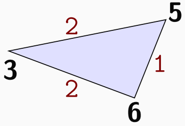
Filtration (definition)

A filtration $f : K \to \mathbb{R}$,

$f(\sigma) \leq f(\tau)$ if $\sigma \subseteq \tau$



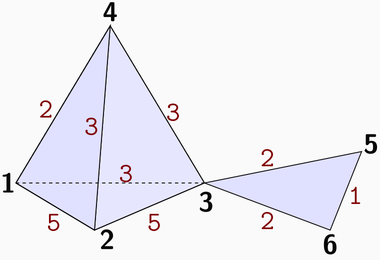


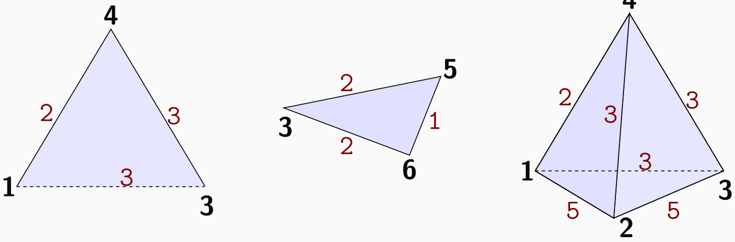


Filtration (definition)

A filtration $f : K \to \mathbb{R}$,

$f(\sigma) \leq f(\tau)$ if $\sigma \subseteq \tau$





Agenda

Find a representation for filtrations of simplicial complexes:

Agenda

Find a representation for filtrations of simplicial complexes:

⋆ Small size.

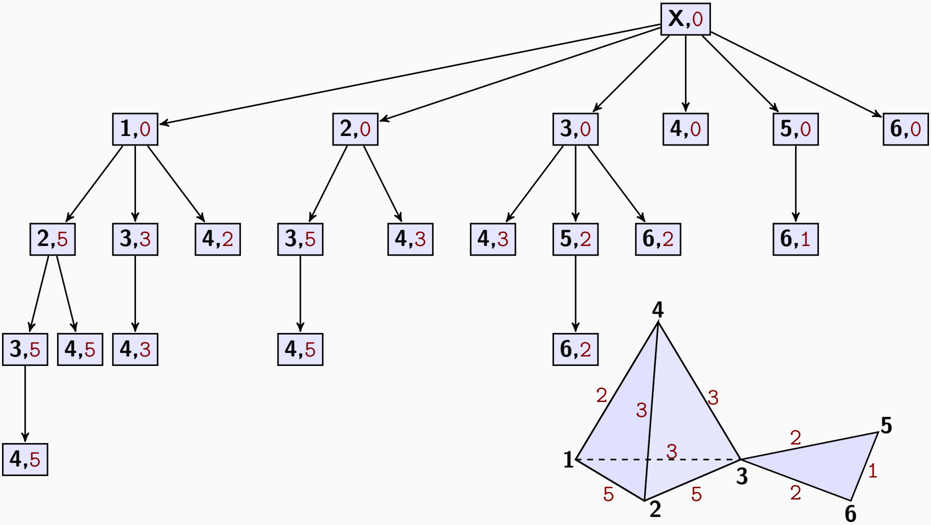
Agenda

Find a representation for filtrations of simplicial complexes:

- ⋆ Small size.
- ⋆ Perform queries quickly:
  - Access filtration value.
  - Simplex Insertion.
  - Simplex Removal.

Simplex Tree

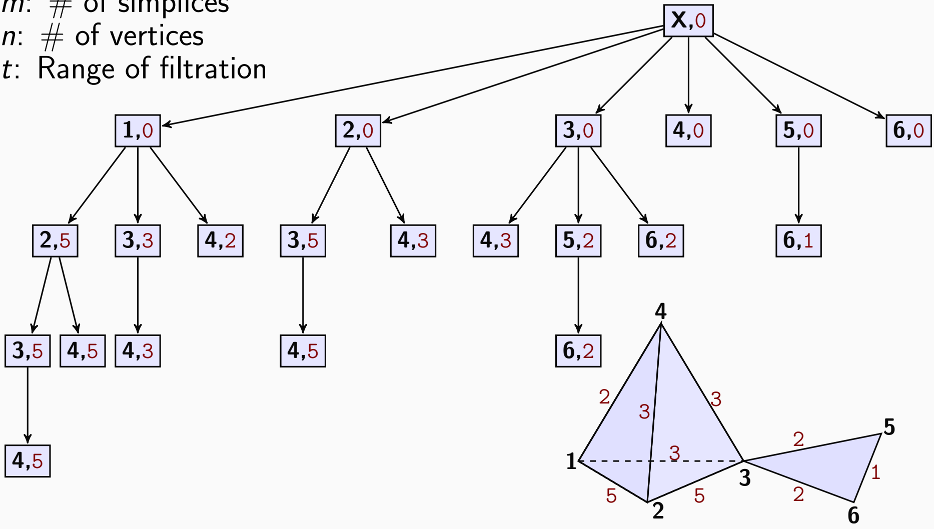
Introduced by Boissonnat and Maria [ESA '12, Algorithmica '14].



Simplex Tree

Storage: $\Theta\left(m\left(\log n + \log t\right)\right)$

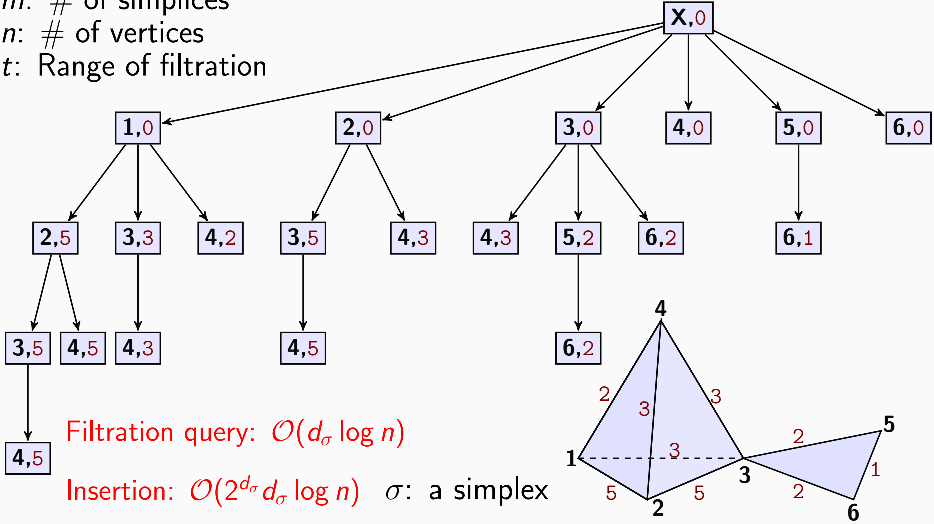
m: # of simplices
n: # of vertices
t: Range of filtration



Simplex Tree

Storage: $\Theta\left(m\left(\log n + \log t\right)\right)$

m: # of simplices
n: # of vertices
t: Range of filtration



Filtration query: $\mathcal{O}(d_\sigma \log n)$

Insertion: $\mathcal{O}(2^{d_\sigma} d_\sigma \log n)$ σ: a simplex

Removal: $\mathcal{O}(m \log n)$ d_σ: dimension of σ

A Change in Perspective

A simplicial complex $K : \{0,1\}^n \to \{0,1\}$:

- $|x| \leq 1 \implies K(x) = 0$.
- Monotonicity: A directed path from y to $x \Rightarrow K(y) \leq K(x)$.

Monotonicity $\longleftrightarrow$ Closed under subsets

A Change in Perspective

A simplicial complex $K : \{0,1\}^n \to \{0,1\}$:

- $|x| \leq 1 \implies K(x) = 0$.
- Monotonicity: A directed path from y to $x \Rightarrow K(y) \leq K(x)$.

Monotonicity $\longleftrightarrow$ Closed under subsets

Boundary: All elements of $\{0,1\}^n$ mapped to 0, all of whose out-neighbors are mapped to 1.

Boundary = Set of Maximal simplices

A Change in Perspective

A simplicial complex $K : \{0,1\}^n \to \{0,1\}$:

- $|x| \leq 1 \implies K(x) = 0$.
- Monotonicity: A directed path from y to $x \Rightarrow K(y) \leq K(x)$.

$$\text{Monotonicity} \longleftrightarrow \text{Closed under subsets}$$

Boundary: All elements of $\{0,1\}^n$ mapped to 0, all of whose out-neighbors are mapped to 1.

$$\text{Boundary} = \text{Set of Maximal simplices}$$

A compact representation: store only the boundary!

Critical Simplex (definition)

A filtration $f : \{0,1\}^n \to \{0,\ldots,t+1\}$:

- $|x| \leq 1 \implies K(x) < t+1$.
- Monotonicity: A directed path from y to $x \Rightarrow K(y) \leq K(x)$.

Critical Simplex (definition)

A filtration $f : \{0,1\}^n \to \{0,\ldots,t+1\}$:

- $|x| \leq 1 \implies K(x) < t+1$.
- Monotonicity: A directed path from y to $x \Rightarrow K(y) \leq K(x)$.

i-Boundary: All elements of $\{0,1\}^n$ mapped to i, all of whose out-neighbors are mapped to values greater than i.

$$\bigcup_{0 \leq i \leq t} i\text{-Boundary} = \text{Set of Critical simplices}$$

Critical Simplex (definition)

A filtration $f : \{0,1\}^n \to \{0,\dots,t+1\}$:

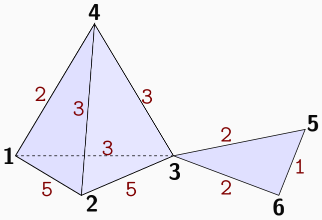
- $|x| \leq 1 \implies K(x) < t+1$.
- Monotonicity: A directed path from y to $x \Rightarrow K(y) \leq K(x)$.

i-Boundary: All elements of $\{0,1\}^n$ mapped to i, all of whose out-neighbors are mapped to values greater than i.

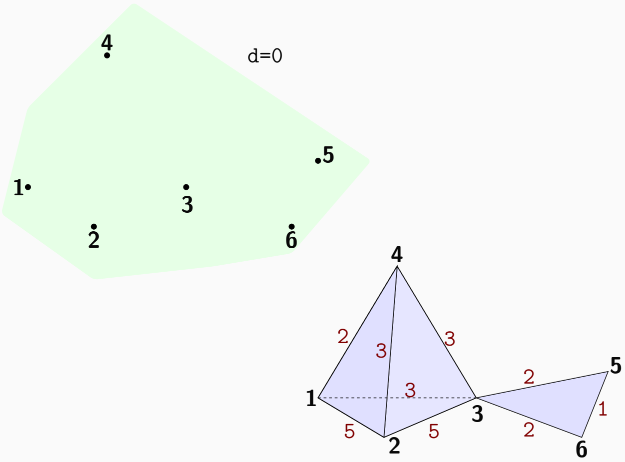
$$\bigcup_{0 \leq i \leq t} i\text{-Boundary} = \text{Set of Critical simplices}$$

A compact representation: store only the critical simplices!

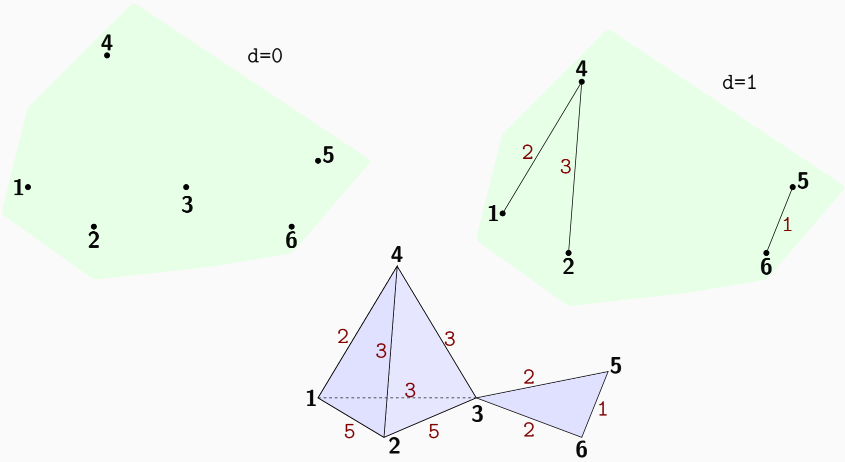
Demonstration

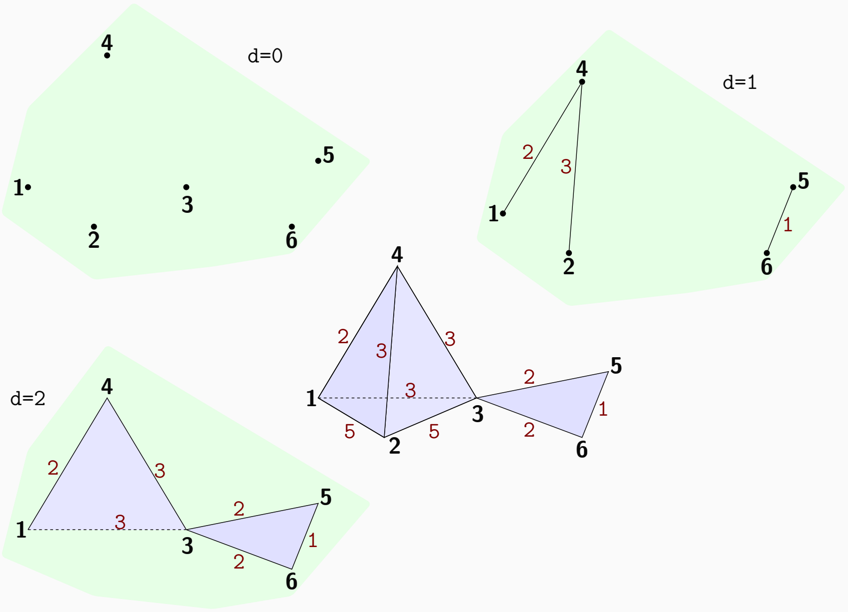


Demonstration

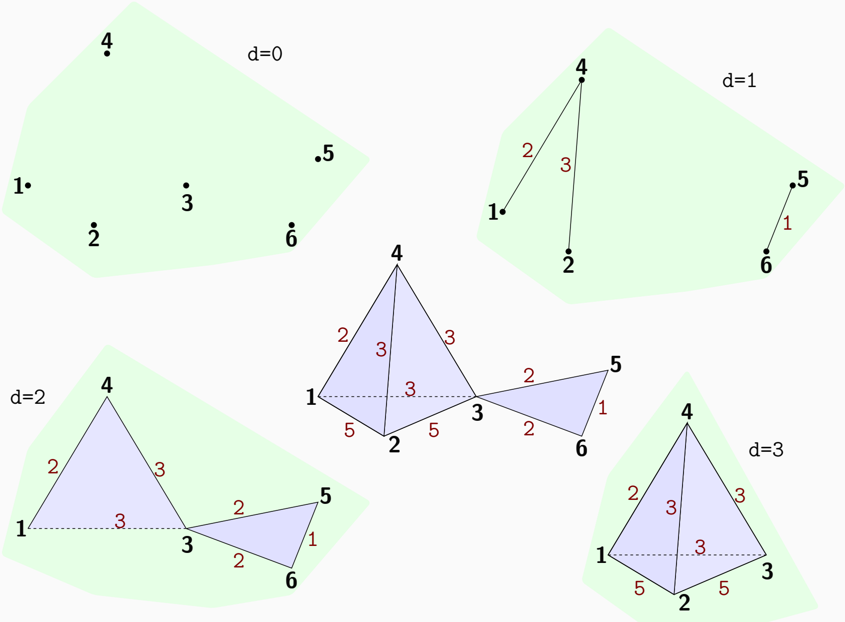


Demonstration




d=0
4
5
1
3
2
6
d=1
4
2
3
5
1
1
2
6
4
2
3
3
3
5
2
1
1
3
5
5
2
2
6
d=2
4
2
3
2
5
1
3
1
3
2
6

Demonstration



Lower Bound on Data Structures

Theorem

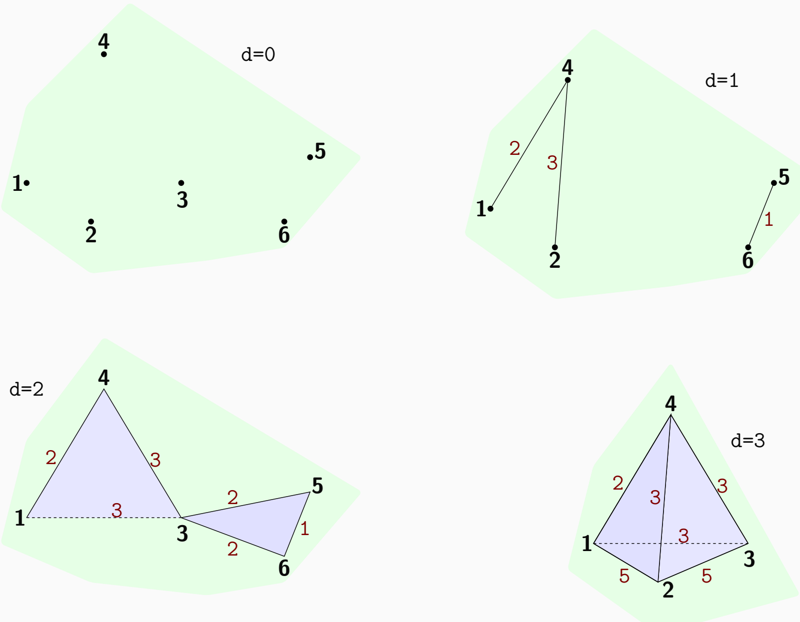
Consider the class of all simplicial complexes on n vertices of dimension d, associated with a filtration over the range of $\{0, \ldots, t\}$, such that the number of critical simplices is κ, where $d \geq 2$ and $\kappa \geq n+1$, and consider any data structure that can represent the simplicial complexes of this class. Such a data structure requires $\log \left(\binom{\binom{n/2}{d+1}}{\kappa - n} t^{\kappa - n} \right)$ bits to be stored. For any constant $\varepsilon \in (0,1)$ and for $\frac{2}{\varepsilon} n \leq \kappa \leq n^{(1-\varepsilon)d}$ and $d \leq n^{\varepsilon/3}$, the bound becomes $\Omega(\kappa(d \log n + \log t))$.

Lower Bound on Data Structures

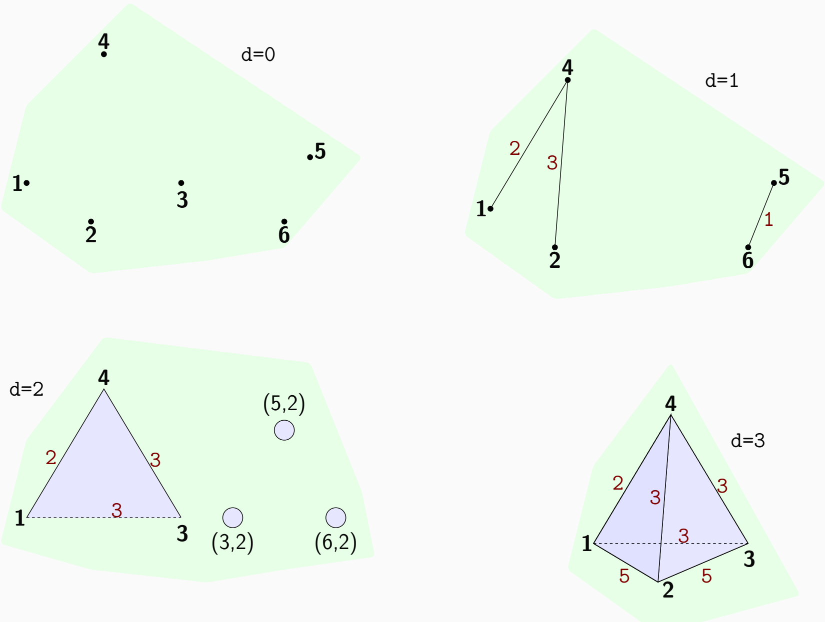
Theorem

Consider the class of all simplicial complexes on n vertices of dimension d, associated with a filtration over the range of $\{0, \ldots, t\}$, such that the number of critical simplices is κ, *where $d \geq 2$ and $\kappa \geq n+1$, and consider any data structure that can represent the simplicial complexes of this class.* **Such a data structure requires** $\log \left(\binom{\binom{n/2}{d+1}}{\kappa - n} t^{\kappa - n} \right)$ *bits to be stored. For any constant $\varepsilon \in (0,1)$ and for $\frac{2}{\varepsilon} n \leq \kappa \leq n^{(1-\varepsilon)d}$ and $d \leq n^{\varepsilon/3}$, the bound becomes* $\Omega(\kappa(d \log n + \log t))$**.**

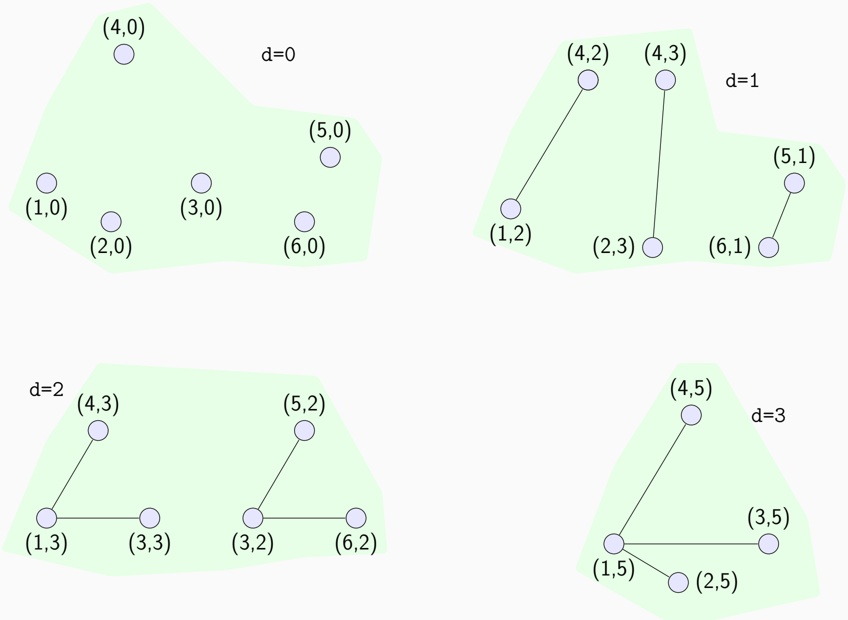
Construction of CSD



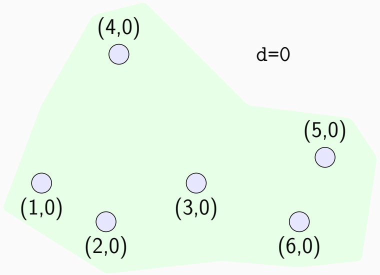
Construction of CSD

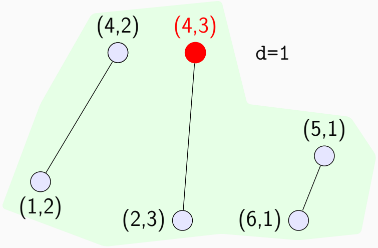


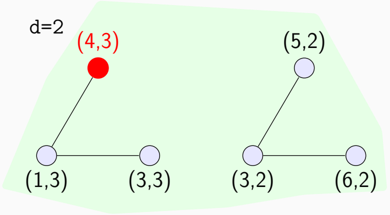
Construction of CSD

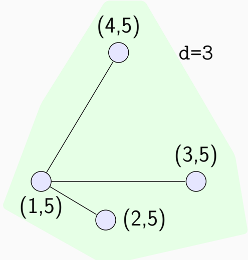


Construction of CSD

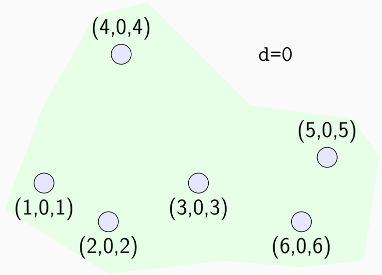


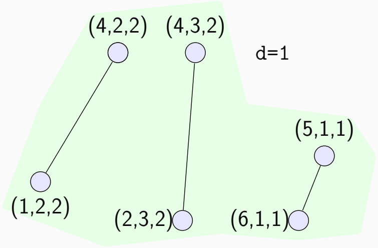


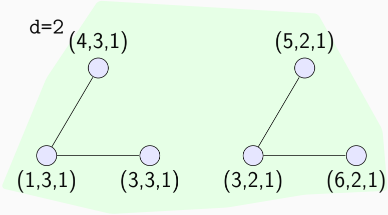


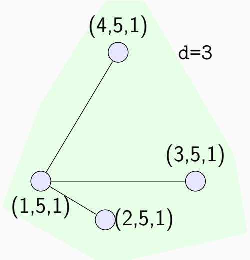


Construction of CSD

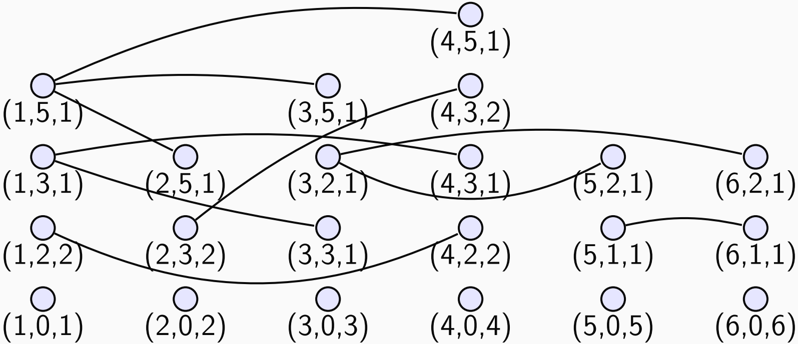




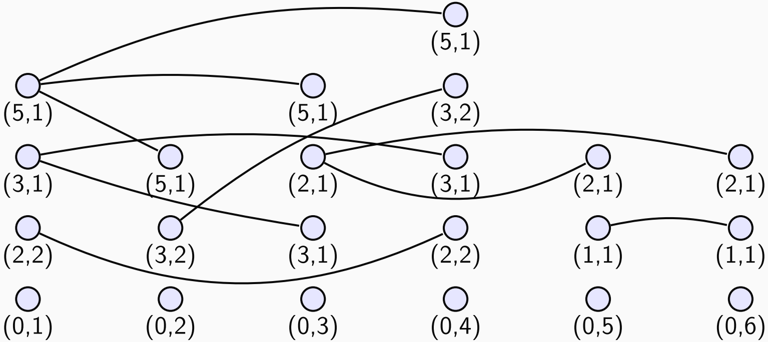


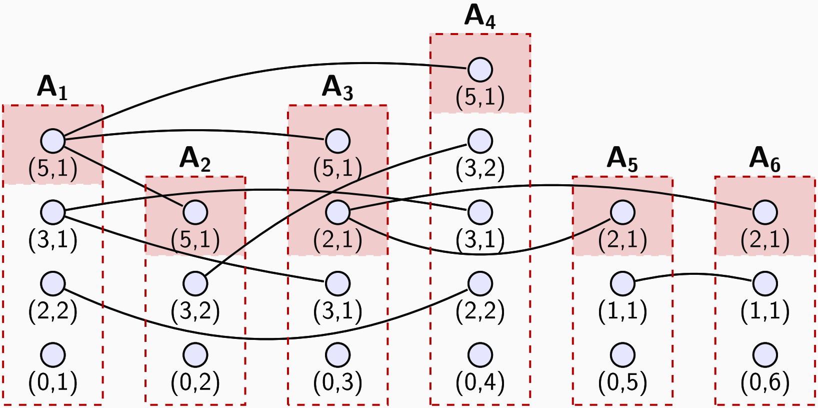


Construction of CSD

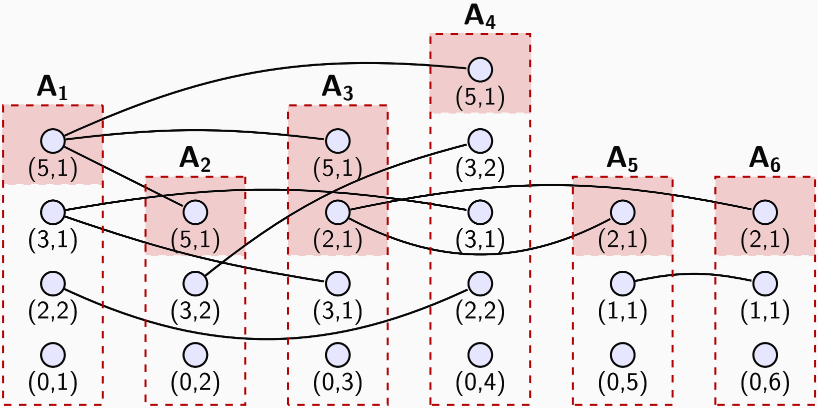


Construction of CSD



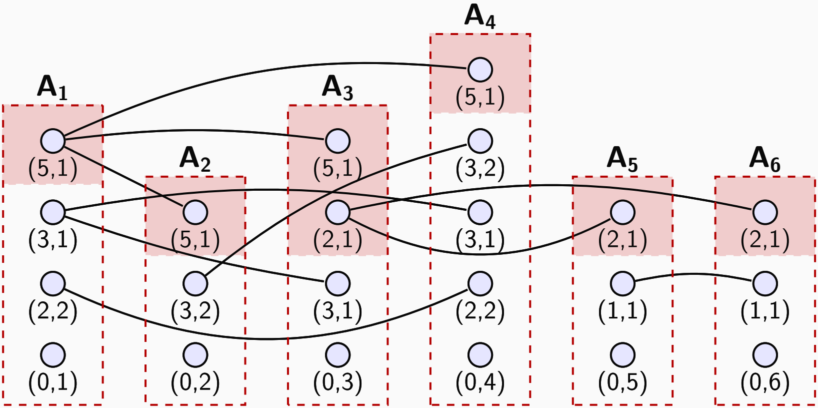

A1
(5,1)
(3,1)
(2,2)
(0,1)
A2
(5,1)
(3,2)
(0,2)
A3
(5,1)
(2,1)
(3,1)
(0,3)
A4
(5,1)
(3,2)
(3,1)
(2,2)
(0,4)
A5
(2,1)
(1,1)
(0,5)
A6
(2,1)
(1,1)
(0,6)

Construction of CSD



Storage$= \mathcal{O}(\kappa d \log(\kappa t))$

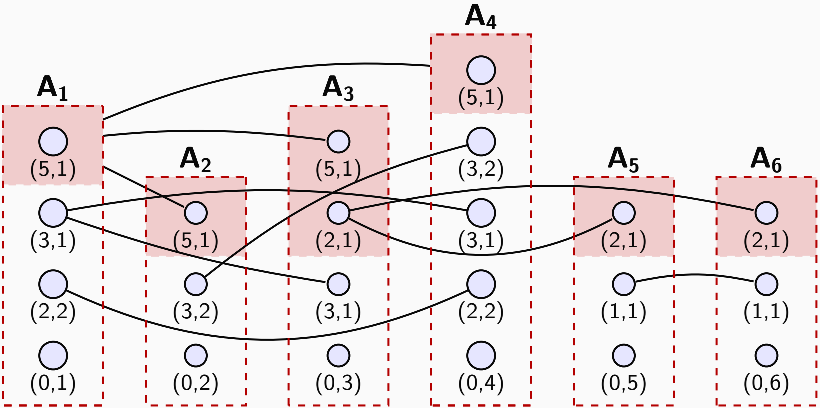
$\kappa :=$ Number of critical simplices



Storage$= \mathcal{O}(\kappa d \log(\kappa t)) \longrightarrow$ "Tight!"

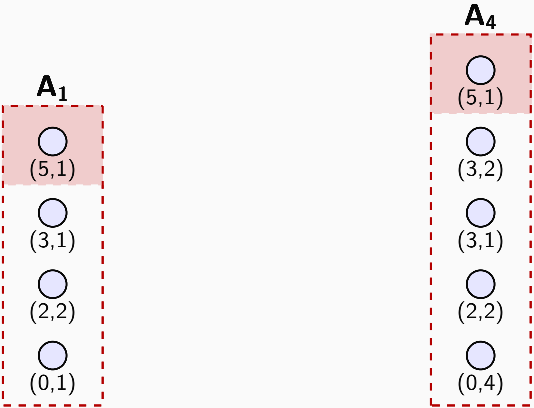
$\kappa :=$ Number of critical simplices

Computing Filtration in CSD

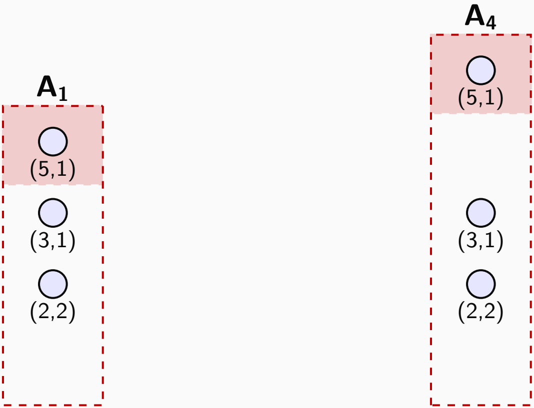


Computing filtration value of $1 - 4$

Computing Filtration in CSD

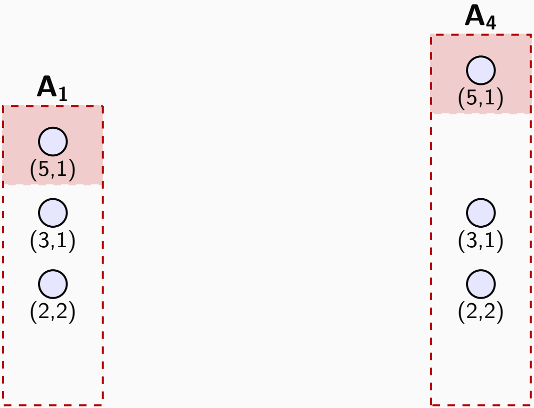


Computing filtration value of $1 - 4$



Computing filtration value of $1 - 4$

Computing Filtration in CSD



Filtration value of $1-4$

$= \min(2,3,5) = 2$

Computing filtration value of $1-4$

Some Relevant Parameters

Γ_j: largest number of maximal simplices of K that a given j-simplex of K may be contained in.

$$1 = \Gamma_d \leq \Gamma_{d-1} \leq \cdots \leq \Gamma_1 \leq \Gamma_0 \leq k := \# \text{ of maximal simplices}$$

Some Relevant Parameters

Γ_j: largest number of maximal simplices of K that a given j-simplex of K may be contained in.

$$1 = \Gamma_d \leq \Gamma_{d-1} \leq \cdots \leq \Gamma_1 \leq \Gamma_0 \leq k := \# \text{ of maximal simplices}$$

$\Psi(v)$: $\#$ of critical simplices that contain vertex v of K.

$$\Psi = \max_{v \in V} \Psi(v).$$

Some Relevant Parameters

Γ_j: largest number of maximal simplices of K that a given j-simplex of K may be contained in.

$$1 = \Gamma_d \leq \Gamma_{d-1} \leq \cdots \leq \Gamma_1 \leq \Gamma_0 \leq k := \#\text{ of maximal simplices}$$

$\Psi(v)$: $\#$ of critical simplices that contain vertex v of K.

$$\Psi = \max_{v \in V} \Psi(v).$$

Ψ is *mostly* much smaller than κ.

Computing Filtration in CSD

Input: A simplex $\sigma = v_{\ell_0} \cdots v_{\ell_{d_\sigma}}$.

Task: Compute $f(\sigma)$.

1. Find $A_{\ell_0}, \ldots, A_{\ell_{d_\sigma}}$.
2. Compute $A_\sigma = \bigcap_{0 \leq i \leq d_\sigma} A_{\ell_i}$.
3. Let $\mathcal{P}$ be projection on to the first coordinate.

$$f(\sigma) = \min \{\mathcal{P}(x) \mid x \in A_\sigma\}.$$

$$\boxed{\mathcal{O}\left(\Psi d_\sigma \log \Psi\right)}$$

Insertion in CSD

Input: A simplex $\sigma = v_{\ell_0} \cdots v_{\ell_{d_\sigma}}$ with filtration value s_σ.

Task: Insert σ in K with filtration value s_σ.

Insertion in CSD

Input: A simplex $\sigma = v_{\ell_0} \cdots v_{\ell_{d_\sigma}}$ with filtration value s_σ.

Task: Insert σ in K with filtration value s_σ.

1. If σ is a maximal simplex in K:
   1.1 Remove or relocate all faces of σ which were previously maximal in K.
   1.2 Introduce the graph component corresponding to σ.

Insertion in CSD

Input: A simplex $\sigma = v_{\ell_0} \cdots v_{\ell_{d_\sigma}}$ with filtration value s_σ.

Task: Insert σ in K with filtration value s_σ.

1. If σ is a maximal simplex in K:
   1.1 Remove or relocate all faces of σ which were previously maximal in K.
   1.2 Introduce the graph component corresponding to σ.
2. If σ is a critical simplex in K:
   2.1 Remove all faces of σ which were previously critical in K but are not any more.
   2.2 Introduce the graph component corresponding to σ.

$$\mathcal{O}\left(\Psi d_\sigma^2 \log \Psi\right)$$

Removal in CSD

Input: A simplex $\sigma = v_{\ell_0} \cdots v_{\ell_{d_\sigma}}$.

Task: Remove σ from K.

Removal in CSD

Input: A simplex $\sigma = v_{\ell_0} \cdots v_{\ell_{d_\sigma}}$.

Task: Remove σ from K.

1. Obtain the set Z_σ of critical simplices in K which contain σ.
2. For every $\tau \in Z_\sigma$, remove τ from K and insert the facets of τ which do not contain σ with appropriate filtration value.

$$\boxed{\mathcal{O}\left((\Psi d_\sigma + \Gamma_{d_\sigma}^2) d \log \Psi\right)}$$

Flag Complexes Construction

Flag Complex of G: All cliques of G are simplices of K.

Flag Complexes Construction

Flag Complex of G: All cliques of G are simplices of K.

Filtration of σ: Maximum of the filtration values of edges in σ.

Flag Complexes Construction

Flag Complex of G: All cliques of G are simplices of K.

Filtration of σ: Maximum of the filtration values of edges in σ.

Input: A weighted graph G.

Output: A simplicial complex K with associated filtration.

Flag Complexes Construction

Flag Complex of G: All cliques of G are simplices of K.

Filtration of σ: Maximum of the filtration values of edges in σ.

Input: A weighted graph G.

Output: A simplicial complex K with associated filtration.

1. Compute all maximal cliques in G. $\boxed{\mathcal{O}(k \cdot n^{\omega})}$

Flag Complexes Construction

Flag Complex of G: All cliques of G are simplices of K.

Filtration of σ: Maximum of the filtration values of edges in σ.

Input: A weighted graph G.

Output: A simplicial complex K with associated filtration.

1. Compute all maximal cliques in G. $\boxed{\mathcal{O}(k \cdot n^{\omega})}$
2. Sort all edges of G according to weight:

$$w(e_1) \geq w(e_2) \geq \cdots \geq w(e_{|E|}).$$

3. For every i from 1 to $|E|$:
   3.1 Find all maximal cliques in induced subgraph of $N(e_i)$.
   3.2 Remove e_i from G.

$\boxed{\mathcal{O}(\kappa \cdot n^{\omega})}$

Summary: ST vs. CSD

	Simplex Tree	Critical Simplex Diagram
Storage	$\mathcal{O}(m \log(nt))$	$\mathcal{O}(\kappa d \log(\kappa t))$
Filtration Query	$\mathcal{O}(d_\sigma \log n)$	$\mathcal{O}(d_\sigma \Psi \log \Psi)$
Insertion	$\mathcal{O}(2^{d_\sigma} d_\sigma \log n)$	$\mathcal{O}(d_\sigma^2 \Psi \log \Psi)$
Removal	$\mathcal{O}(m \log n)$	$\mathcal{O}((\Psi d_\sigma + \Gamma_{d_\sigma}^2) d \log \Psi)$
Construction of Flag Complex	$\mathcal{O}(md \log n)$	$\mathcal{O}(\kappa n^{2.38})$

Summary: ST vs. CSD

	Simplex Tree	Critical Simplex Diagram
Storage	$\mathcal{O}(m\log(nt))$	$\mathcal{O}(\kappa d\log(\kappa t))$
Filtration Query	$\mathcal{O}(d_\sigma \log n)$	$\mathcal{O}(d_\sigma \Psi \log \Psi)$
Insertion	$\mathcal{O}(2^{d_\sigma} d_\sigma \log n)$	$\mathcal{O}(d_\sigma^2 \Psi \log \Psi)$
Removal	$\mathcal{O}(m\log n)$	$\mathcal{O}((\Psi d_\sigma + \Gamma_{d_\sigma}^2)d\log \Psi)$
Construction of Flag Complex	$\mathcal{O}(md\log n)$	$\mathcal{O}(\kappa n^{2.38})$

Summary: ST vs. CSD

	Simplex Tree	Critical Simplex Diagram
Storage	$\mathcal{O}(m\log(nt))$	$\mathcal{O}(\kappa d\log(\kappa t))$
Filtration Query	$\mathcal{O}(d_\sigma \log n)$	$\mathcal{O}(d_\sigma \Psi \log \Psi)$
Insertion	$\mathcal{O}(2^{d_\sigma} d_\sigma \log n)$	$\mathcal{O}(d_\sigma^2 \Psi \log \Psi)$
Removal	$\mathcal{O}(m\log n)$	$\mathcal{O}((\Psi d_\sigma + \Gamma_{d_\sigma}^2)d\log \Psi)$
Construction of Flag Complex	$\mathcal{O}(md\log n)$	$\mathcal{O}(\kappa n^{2.38})$

Summary: ST vs. CSD

	Simplex Tree	Critical Simplex Diagram
Storage	$\mathcal{O}(m \log(nt))$	$\mathcal{O}(\kappa d \log(\kappa t))$
Filtration Query	$\mathcal{O}(d_\sigma \log n)$	$\mathcal{O}(d_\sigma \Psi \log \Psi)$
Insertion	$\mathcal{O}(2^{d_\sigma} d_\sigma \log n)$	$\mathcal{O}(d_\sigma^2 \Psi \log \Psi)$
Removal	$\mathcal{O}(m \log n)$	$\mathcal{O}((\Psi d_\sigma + \Gamma_{d_\sigma}^2) d \log \Psi)$
Construction of Flag Complex	$\mathcal{O}(md \log n)$	$\mathcal{O}(\kappa n^{2.38})$

Summary: ST vs. CSD

	Simplex Tree	Critical Simplex Diagram
Storage	$\mathcal{O}(m \log(nt))$	$\mathcal{O}(\kappa d \log(\kappa t))$
Filtration Query	$\mathcal{O}(d_\sigma \log n)$	$\mathcal{O}(d_\sigma \Psi \log \Psi)$
Insertion	$\mathcal{O}(2^{d_\sigma} d_\sigma \log n)$	$\mathcal{O}(d_\sigma^2 \Psi \log \Psi)$
Removal	$\mathcal{O}(m \log n)$	$\mathcal{O}((\Psi d_\sigma + \Gamma_{d_\sigma}^2) d \log \Psi)$
Construction of Flag Complex	$\mathcal{O}(md \log n)$	$\mathcal{O}(\kappa n^{2.38})$

Summary: ST vs. CSD

	Simplex Tree	Critical Simplex Diagram
Storage	$\mathcal{O}(m \log(nt))$	$\mathcal{O}(\kappa d \log(\kappa t))$
Filtration Query	$\mathcal{O}(d_\sigma \log n)$	$\mathcal{O}(d_\sigma \Psi \log \Psi)$
Insertion	$\mathcal{O}(2^{d_\sigma} d_\sigma \log n)$	$\mathcal{O}(d_\sigma^2 \Psi \log \Psi)$
Removal	$\mathcal{O}(m \log n)$	$\mathcal{O}((\Psi d_\sigma + \Gamma_{d_\sigma}^2) d \log \Psi)$
Construction of Flag Complex	$\mathcal{O}(md \log n)$	$\mathcal{O}(\kappa n^{2.38})$

Thank you!

Value of Ψ

Data Set: Rips Complex from sampling of Klein bottle in $\mathbb{R}^5$.

No	t	$\lvert ST \rvert$	$\lvert CSD \rvert$	Γ_0	Γ_0^{avg}	Ψ	Ψ^{avg}
0	0	10,508,486	179,521	115	17.9	115	17.9
1	10	10,508,486	490,071	115	17.9	329	49.0
2	25	10,508,486	618,003	115	17.9	429	61.8
3	100	10,508,486	728,245	115	17.9	723	72.8
4	500	10,508,486	765,583	115	17.9	839	76.5
5	2,000	10,508,486	774,496	115	17.9	860	77.4
6	10,000	10,508,486	777,373	115	17.9	865	77.7
7	25,000	10,508,486	778,151	115	17.9	865	77.8
8	100,000	10,508,486	778,319	115	17.9	866	77.8
9	1,000,000	10,508,486	778,343	115	17.9	866	77.8
10	10,000,000	10,508,486	778,343	115	17.9	866	77.8