Journal Club on Topological Data Analysis

Edge Collapse and Persistence of Flag Complexes (by Boissonnat and Pritam)

Tim Mäder

February 8, 2021

Outline

- Motivational introduction

Outline

- Motivational introduction
- Preliminaries

Outline

- Motivational introduction
- Preliminaries
- Simplicial collapses and flag complexes

Outline

- Motivational introduction
- Preliminaries
- Simplicial collapses and flag complexes
- Computational Experiments

Introduction

- Elementary collapse: Remove a maximal simplex together with a free face

Introduction

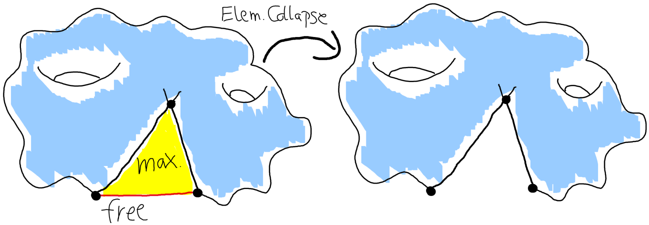
- Elementary collapse: Remove a maximal simplex together with a free face
- Maximal simplex: Not the proper face of any other simplex

Introduction

- Elementary collapse: Remove a maximal simplex together with a free face
- Maximal simplex: Not the proper face of any other simplex
- Free simplex: Only the face of a unique simplex

Introduction

- Elementary collapse: Remove a maximal simplex together with a free face
- Maximal simplex: Not the proper face of any other simplex
- Free simplex: Only the face of a unique simplex



Example

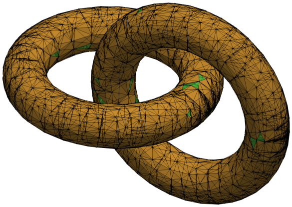


Figure: 23088 Simplices

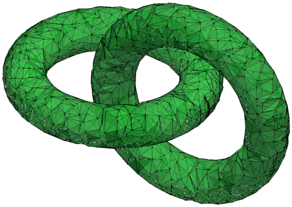


Figure: 9520 Simplices

Example

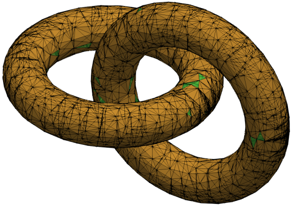


Figure: 23088 Simplices

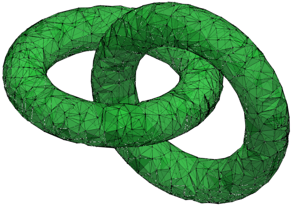


Figure: 9520 Simplices

- Less than half of the simplices, same homotopy type, same homology groups.

Example

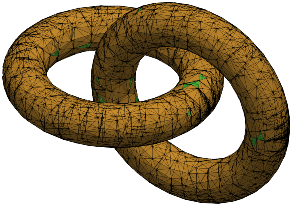


Figure: 23088 Simplices

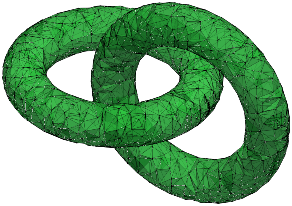


Figure: 9520 Simplices

- Less than half of the simplices, same homotopy type, same homology groups.
- Question: Is this a quick process?

Reference

J-D. Boissonnat, S. Pritam, *Edge Collapse and Persistence of Flag Complexes*, Proceedings of 36th International Symposium on Computational Geometry (2020), 19:1 – 19:15.

Preliminaries

- An **(abstract) simplicial complex** K is a collection of subsets of a non-empty finite set X, such that for every subset A in K, all the subsets of A are in K.

Preliminaries

- An **(abstract) simplicial complex** K is a collection of subsets of a non-empty finite set X, such that for every subset A in K, all the subsets of A are in K.
- A map $f : K \to L$ between simplicial complexes is called a **simplicial map**, if it always maps a simplex in K to a simplex in L.

Preliminaries

- An **(abstract) simplicial complex** K is a collection of subsets of a non-empty finite set X, such that for every subset A in K, all the subsets of A are in K.
- A map $f : K \to L$ between simplicial complexes is called a **simplicial map**, if it always maps a simplex in K to a simplex in L.
- A sequence of simplicial complexes

$$\{K_1 \xrightarrow{f_1} K_2 \xrightarrow{f_2} \dots \xrightarrow{f_{n-1}} K_n\}$$

  connected through simplicial inclusion maps f_i is called a **(simplicial) filtration**. We often write $\{K_i, f_i\}$ for such a filtration.

Preliminaries

- Computing simplicial homology over field coeffiecients for all K_i, yields a **persistence module**

$$\{H_p(K_1) \xrightarrow{f_{1,p}} H_p(K_2) \xrightarrow{f_{2,p}} \dots \xrightarrow{f_{n-1,p}} H_p(K_n)\}$$

in every degree p.

Preliminaries

- Computing simplicial homology over field coeffiecients for all K_i, yields a **persistence module**

$$\{H_p(K_1) \xrightarrow{f_{1,p}} H_p(K_2) \xrightarrow{f_{2,p}} \dots \xrightarrow{f_{n-1,p}} H_p(K_n)\}$$

  in every degree p.
- Let X be a finte metric space and let $\epsilon > 0$. The *Vietoris-Rips complex of X at scale ϵ* is defined as the set

$$\mathrm{VR}_\epsilon(X) := \{\sigma \subset X | d(x,y) \leq 2\epsilon \text{ for all } x, y \in \sigma\}.$$

Simplicial Collapses

Definition

- A simplex in K is called *maximal* in K, if it is not a proper face of any simplex in K.

Definition

- A simplex in K is called *maximal* in K, if it is not a proper face of any simplex in K.
- Let τ be a maximal simplex in K and suppose σ is a proper face of τ in K.

Definition

- A simplex in K is called *maximal* in K, if it is not a proper face of any simplex in K.
- Let τ be a maximal simplex in K and suppose σ is a proper face of τ in K.
  If σ is not a proper face of any simplex in K other than τ, then σ is called *free* (in K).

Definition

- A simplex in K is called *maximal* in K, if it is not a proper face of any simplex in K.
- Let τ be a maximal simplex in K and suppose σ is a proper face of τ in K.
  If σ is not a proper face of any simplex in K other than τ, then σ is called *free* (in K).

- The set $K' = K - \{\tau, \sigma\}$ is a simplicial complex.

Definition

- A simplex in K is called *maximal* in K, if it is not a proper face of any simplex in K.
- Let τ be a maximal simplex in K and suppose σ is a proper face of τ in K.
  If σ is not a proper face of any simplex in K other than τ, then σ is called *free* (in K).

- The set $K' = K - \{\tau, \sigma\}$ is a simplicial complex.
- The associated polyhedra $|K'|$ and $|K|$ are homotopy equivalent.

Definition

- A simplex in K is called *maximal* in K, if it is not a proper face of any simplex in K.
- Let τ be a maximal simplex in K and suppose σ is a proper face of τ in K.
  If σ is not a proper face of any simplex in K other than τ, then σ is called *free* (in K).

- The set $K' = K - \{\tau, \sigma\}$ is a simplicial complex.
- The associated polyhedra $|K'|$ and $|K|$ are homotopy equivalent.
- J. H. C. Whitehead, *Simplicial Spaces, Nuclei and m-groups*, Proceedings of the London mathematical society **2** (1939), no. 1, 243-327.
- M. M. Cohen, *A Course in Simple-Homotopy Theory*, Graduate Texts in Mathematics, Vol. 10, Springer-Verlag, New York, 1973.

- The deformation retraction associated with the removal of σ and τ as before can be realized as follows:

$$
\begin{aligned}
H : |\tau| \times I &\to |\tau| \\
((x_1, \ldots, x_n), t) &\mapsto (1-t)(x_1, \ldots, x_n) \\
&\qquad + t(x_1 - \min_i x_i, \ldots, x_n - \min_i x_i)
\end{aligned}
$$

with simplices τ and σ parametrised as subsets of $\mathbb{R}^n$

$$
\begin{aligned}
|\tau| &= \{(x_1, \ldots, x_n) \in \mathbb{R}^n | x_i \geq 0, \sum_{i=1}^{n} x_i \leq 1\}, \\
|\sigma| &= \{(x_1, \ldots, x_n) \in \mathbb{R}^n | x_i \geq 0, \sum_{i=1}^{n} x_i = 1\}.
\end{aligned}
$$

Definition

- We say K' has been obtained from K by an *elementary collapse* (of τ using its free face σ).

Definition

- We say K' has been obtained from K by an *elementary collapse* (of τ using its free face σ).
- A simplicial complex K *collapses* (simplicially) to a subcomplex $L \subset K$, if L can be obtained by a finite sequence of elementary collapses. In that case we write $K \searrow L$.

We need some notation to formulate a class of collapses that are easy to identify algorithmically.

Star, Link, Simplicial Cone

- Let σ be a simplex in K. The **closed star** of σ in K, denoted as $st_K(\sigma)$ is defined as

$$st_K(\sigma) := \{\tau \in K | \tau \cup \sigma \in K\}.$$

Star, Link, Simplicial Cone

- Let σ be a simplex in K. The **closed star** of σ in K, denoted as $st_K(\sigma)$ is defined as

$$st_K(\sigma) := \{\tau \in K | \tau \cup \sigma \in K\}.$$

- The **link** of σ in K, $lk_K(\sigma)$ is defined as

$$lk_K(\sigma) := \{\tau \in st_K(\sigma) | \tau \cap \sigma = \emptyset\}.$$

Star, Link, Simplicial Cone

- Let σ be a simplex in K. The **closed star** of σ in K, denoted as $st_K(\sigma)$ is defined as

$$st_K(\sigma) := \{\tau \in K | \tau \cup \sigma \in K\}.$$

- The **link** of σ in K, $lk_K(\sigma)$ is defined as

$$lk_K(\sigma) := \{\tau \in st_K(\sigma) | \tau \cap \sigma = \emptyset\}.$$

- The **open star** of σ in K, denoted as $st_K^\circ(\sigma)$ is defined as

$$st_K^\circ(\sigma) := st_K(\sigma) \setminus lk_K(\sigma).$$

Star, Link, Simplicial Cone

- Let σ be a simplex in K. The **closed star** of σ in K, denoted as $st_K(\sigma)$ is defined as

$$st_K(\sigma) := \{\tau \in K | \tau \cup \sigma \in K\}.$$

- The **link** of σ in K, $lk_K(\sigma)$ is defined as

$$lk_K(\sigma) := \{\tau \in st_K(\sigma) | \tau \cap \sigma = \emptyset\}.$$

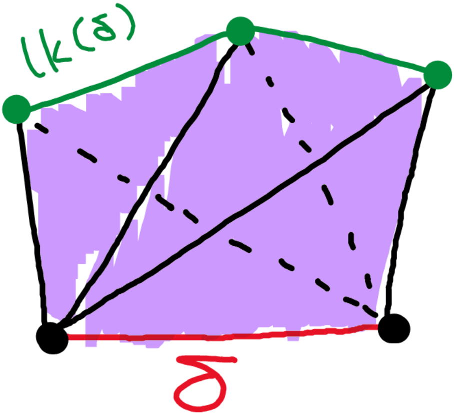
- The **open star** of σ in K, denoted as $st^\circ_K(\sigma)$ is defined as

$$st^\circ_K(\sigma) := st_K(\sigma) \backslash lk_K(\sigma).$$

- Let L be a subcomplex of K and let v be a vertex in K but not in L. Then the set

$$vL := \{\tau | \tau \in L \text{ or } \tau = \sigma \cup v \text{ for } \sigma \in L\} \cup \{v\}$$

is called a **simplicial cone**.


lk(δ)
δ

Definition

- A simplex σ in K is called a **dominated** if the link of σ in K is a simplicial cone, i.e. if there exists a vertex $v \notin \sigma$ and a subcomplex L of K, such that $lk_K(\sigma) = vL$.

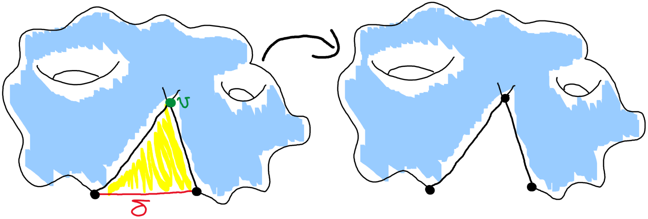
Definition

- A simplex σ in K is called a **dominated** if the link of σ in K is a simplicial cone, i.e. if there exists a vertex $v \notin \sigma$ and a subcomplex L of K, such that $lk_K(\sigma) = vL$. In that case we say σ is dominated by the vertex v and write $\sigma \prec v$.

Definition

- A simplex σ in K is called a **dominated** if the link of σ in K is a simplicial cone, i.e. if there exists a vertex $v \notin \sigma$ and a subcomplex L of K, such that $lk_K(\sigma) = vL$. In that case we say σ is dominated by the vertex v and write $\sigma \prec v$.

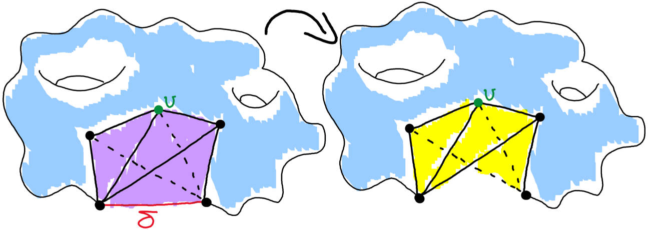
- A free simplex σ in K is dominated by $v = \tau \backslash \sigma$ where τ is the maximal coface:



Definition

- A simplex σ in K is called a **dominated** if the link of σ in K is a simplicial cone, i.e. if there exists a vertex $v \notin \sigma$ and a subcomplex L of K, such that $lk_K(\sigma) = vL$. In that case we say σ is dominated by the vertex v and write $\sigma \prec v$.

- Detecting dominated simplices in a complex can trigger a sequence of elementary collapses:



Lemma

Let K be a simplicial complex and let σ be a simplex of K. If the link of σ is a cone, then there is a sequence of elementary collapses from K to $K \backslash st^{\circ}_{K}(\sigma)$.

Lemma

Let K be a simplicial complex and let σ be a simplex of K. If the link of σ is a cone, then there is a sequence of elementary collapses from K to $K \backslash st_K^\circ(\sigma)$.

- A cone vL is collapsable to its apex v by sequentially removing pairs of simplices of the form $(\alpha \cup \{v\}, \alpha)$ with $v \notin \alpha \subset L$ and $\alpha \cup \{v\}$ maximal.

Lemma

Let K be a simplicial complex and let σ be a simplex of K. If the link of σ is a cone, then there is a sequence of elementary collapses from K to $K \backslash st_K^{\circ}(\sigma)$.

- A cone vL is collapsable to its apex v by sequentially removing pairs of simplices of the form $(\alpha \cup \{v\}, \alpha)$ with $v \notin \alpha \subset L$ and $\alpha \cup \{v\}$ maximal.
- For $lk_K(\sigma) = vL$ in K, these collapses can be associated with the removal of pairs $(\sigma \cup \alpha \cup \{v\}, \sigma \cup \alpha)$ in K that define elementary collapses.

How to identify the dominated simplices?

Lemma 1

A simplex $\sigma \in K$ is dominated by a vertex $v \in K$, $v \notin \sigma$, if and only if all the maximal simplices of K that contain σ also contain v.

Lemma 1

A simplex $\sigma \in K$ is dominated by a vertex $v \in K$, $v \notin \sigma$, if and only if all the maximal simplices of K that contain σ also contain v.

- Identifiying dominated simplices requires knowledge about maximal simplices in K.

Flag Complex and Neighborhood

- A complex K is a flag or a clique complex if, when a subset of its vertices form a clique (i.e. any pair of vertices is joined by an edge), they span a simplex

Flag Complex and Neighborhood

- A complex K is a flag or a clique complex if, when a subset of its vertices form a clique (i.e. any pair of vertices is joined by an edge), they span a simplex
- For a vertex $v \in K$ the **open neighborhood** and the **closed neighborhood** are defined as $N_K(v) := \{u | [u, v] \in K\}$ and $N_K[v] := N_K(v) \cup \{v\}$ respectively.

Flag Complex and Neighborhood

- A complex K is a flag or a clique complex if, when a subset of its vertices form a clique (i.e. any pair of vertices is joined by an edge), they span a simplex
- For a vertex $v \in K$ the **open neighborhood** and the **closed neighborhood** are defined as $N_K(v) := \{u | [u, v] \in K\}$ and $N_K[v] := N_K(v) \cup \{v\}$ respectively.
- The open and closed neighborhoods of a k-simplex $\sigma = [v_1, \ldots, v_k]$ in K are defined as $N_K(\sigma) := \bigcap_{v_i \in \sigma} N_K(v_i)$ and $N_K[\sigma] := \bigcap_{v_i \in \sigma} N_K[v_i]$

Lemma 2

Let σ be a simplex of a flag complex K. Then σ will be dominated by a vertex $v \in K$ if and only if $N_K[\sigma] \subseteq N_K[v]$.

Lemma 2

Let σ be a simplex of a flag complex K. Then σ will be dominated by a vertex $v \in K$ if and only if $N_K[\sigma] \subseteq N_K[v]$.

Proof.

$'' \Leftarrow ''$:

- Let τ be a maximal simplex, s.t. $\sigma \subset \tau$.

Lemma 2

Let σ be a simplex of a flag complex K. Then σ will be dominated by a vertex $v \in K$ if and only if $N_K[\sigma] \subseteq N_K[v]$.

Proof.

" $\Leftarrow$ " :

- Let τ be a maximal simplex, s.t. $\sigma \subset \tau$.
- For any vertex $x \in \tau$ we have $x \in N_K[\sigma] \subseteq N_K[v]$.

Lemma 2

Let σ be a simplex of a flag complex K. Then σ will be dominated by a vertex $v \in K$ if and only if $N_K[\sigma] \subseteq N_K[v]$.

Proof.

" $\Leftarrow$ " :

- Let τ be a maximal simplex, s.t. $\sigma \subset \tau$.
- For any vertex $x \in \tau$ we have $x \in N_K[\sigma] \subseteq N_K[v]$.
- Because K is a flag complex and τ is maximal, v must lie in τ. With Lemma 1 we conclude that $\sigma \prec v$.

Lemma 2

Let σ be a simplex of a flag complex K. Then σ will be dominated by a vertex $v \in K$ if and only if $N_K[\sigma] \subseteq N_K[v]$.

Proof.

" $\Leftarrow$ " :

- Let τ be a maximal simplex, s.t. $\sigma \subset \tau$.
- For any vertex $x \in \tau$ we have $x \in N_K[\sigma] \subseteq N_K[v]$.
- Because K is a flag complex and τ is maximal, v must lie in τ. With Lemma 1 we conclude that $\sigma \prec v$.

" $\Rightarrow$ " :

- $\sigma \prec v \xLongrightarrow{Lemma 1}$ all maximal simplices that contain σ also contain v. This implies $N_K[\sigma] \subseteq N_K[v]$.

□

Theorem

Let $f : K \to L$ be a simplicial map between two complexes K and L and let $K' \subset K$ and $L' \subset L$ be subcomplexes of K and L such that $K \searrow K'$ and $L \searrow L'$. Then there exists a map $f' : K' \to L'$, induced by f, such that the persistence of $f_* : H_p(K) \to H_p(L)$ and $f'_* : H_p(K') \to H_p(L')$ are the same for any integer $p \geq 0$.

Theorem

Let $f : K \to L$ be a simplicial map between two complexes K and L and let $K' \subset K$ and $L' \subset L$ be subcomplexes of K and L such that $K \searrow K'$ and $L \searrow L'$. Then there exists a map $f' : K' \to L'$, induced by f, such that the persistence of $f_* : H_p(K) \to H_p(L)$ and $f'_* : H_p(K') \to H_p(L')$ are the same for any integer $p \geq 0$.

Proof

$$
\begin{array}{ccc}
|K| & \xrightarrow{|f|} & |L| \\
|i_K| \uparrow \downarrow |r_K| & & |i_L| \uparrow \downarrow |r_L| \\
|K'| & \xrightarrow{|f'|} & |L'|
\end{array}
\quad \rightsquigarrow \quad
\begin{array}{ccc}
H_p(|K|) & \xrightarrow{|f|_*} & H_p(|L|) \\
|i_K|_* \uparrow \downarrow |r_K|_* & & |i_L|_* \uparrow \downarrow |r_L|_* \\
H_p(|K'|) & \xrightarrow{|f'|_*} & H_p(|L'|)
\end{array}
$$

- Advantages of using flag complexes:

- Advantages of using flag complexes:
  - Complex is fully determined by the 1-skeleton

- Advantages of using flag complexes:
  - Complex is fully determined by the 1-skeleton
  - Dominated edges can easily be recognized on the 1-skeleton

Edge Collapse Algorithm

Edge Collapse Algorithm

- Let $\{K_1 \to K_2 \to \cdots \to K_n\}$ be a filtration of flag complexes and let $\{G_1 \to G_2 \to \cdots \to G_n\}$ be the associated filtration of 1-skeleta.

Edge Collapse Algorithm

- Let $\{K_1 \to K_2 \to \cdots \to K_n\}$ be a filtration of flag complexes and let $\{G_1 \to G_2 \to \cdots \to G_n\}$ be the associated filtration of 1-skeleta.
- Assume $G_i = \{e_1, \ldots, e_i\}$ and $G_{i+1} = G_i \cup \{e_{i+1}\}$.

Edge Collapse Algorithm

- Let $\{K_1 \to K_2 \to \cdots \to K_n\}$ be a filtration of flag complexes and let $\{G_1 \to G_2 \to \cdots \to G_n\}$ be the associated filtration of 1-skeleta.
- Assume $G_i = \{e_1, \ldots, e_i\}$ and $G_{i+1} = G_i \cup \{e_{i+1}\}$.
- Iterate over all edges in increasing order and check if an edge e_i is dominated in G_i or not:

Edge Collapse Algorithm

- Let $\{K_1 \to K_2 \to \cdots \to K_n\}$ be a filtration of flag complexes and let $\{G_1 \to G_2 \to \cdots \to G_n\}$ be the associated filtration of 1-skeleta.
- Assume $G_i = \{e_1, \ldots, e_i\}$ and $G_{i+1} = G_i \cup \{e_{i+1}\}$.
- Iterate over all edges in increasing order and check if an edge e_i is dominated in G_i or not:
  - If non-dominated $\to$ include e_i in a collection of edges E^c with filtration index i

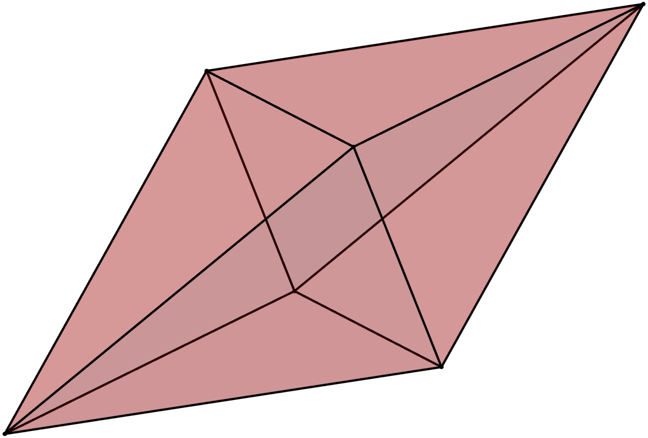
Edge Collapse Algorithm

- Let $\{K_1 \to K_2 \to \cdots \to K_n\}$ be a filtration of flag complexes and let $\{G_1 \to G_2 \to \cdots \to G_n\}$ be the associated filtration of 1-skeleta.
- Assume $G_i = \{e_1, \ldots, e_i\}$ and $G_{i+1} = G_i \cup \{e_{i+1}\}$.
- Iterate over all edges in increasing order and check if an edge e_i is dominated in G_i or not:
    - If non-dominated $\to$ include e_i in a collection of edges E^c with filtration index i
    - Iterate through edges in G_i in reverse order and search for new non-dominated edges e_j, $j < i$.

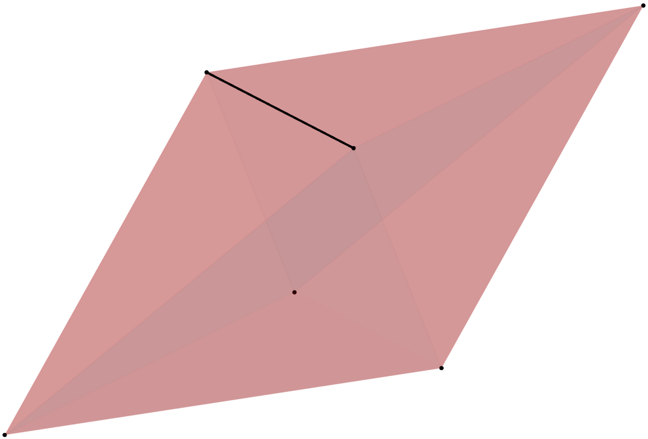
Edge Collapse Algorithm

- Let $\{K_1 \to K_2 \to \cdots \to K_n\}$ be a filtration of flag complexes and let $\{G_1 \to G_2 \to \cdots \to G_n\}$ be the associated filtration of 1-skeleta.
- Assume $G_i = \{e_1, \ldots, e_i\}$ and $G_{i+1} = G_i \cup \{e_{i+1}\}$.
- Iterate over all edges in increasing order and check if an edge e_i is dominated in G_i or not:
  - If non-dominated $\to$ include e_i in a collection of edges E^c with filtration index i
  - Iterate through edges in G_i in reverse order and search for new non-dominated edges e_j, $j < i$.
  - If e_j non-dominated $\to$ include e_j in E^c with filtration index i.

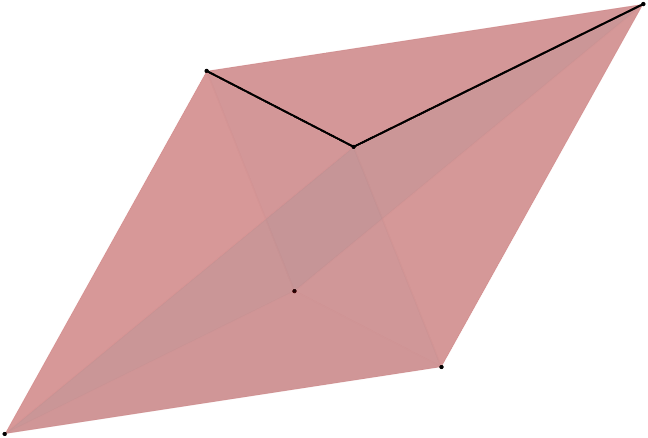
Example Edge Collapse



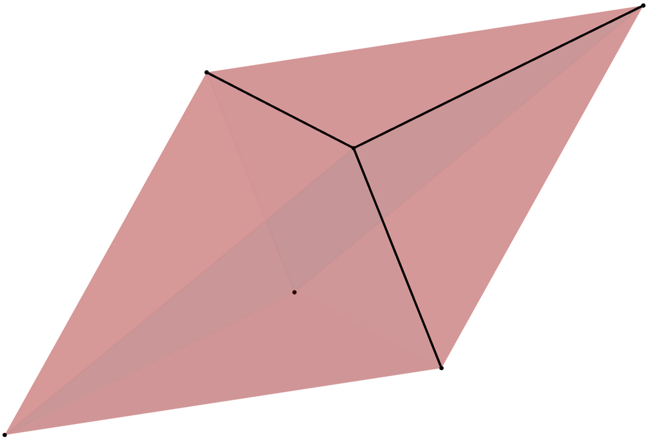
Example Edge Collapse



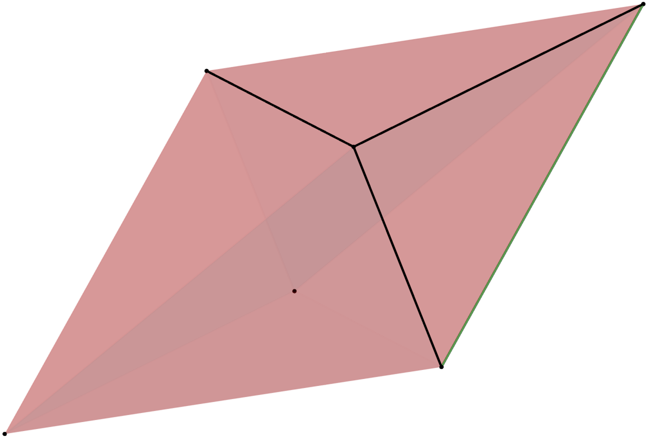
Example Edge Collapse



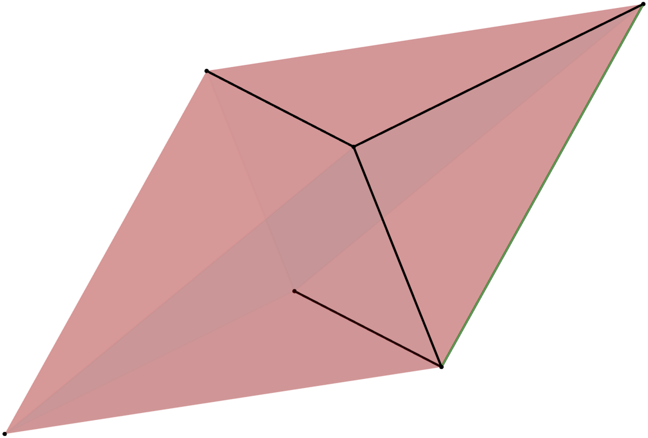
Example Edge Collapse



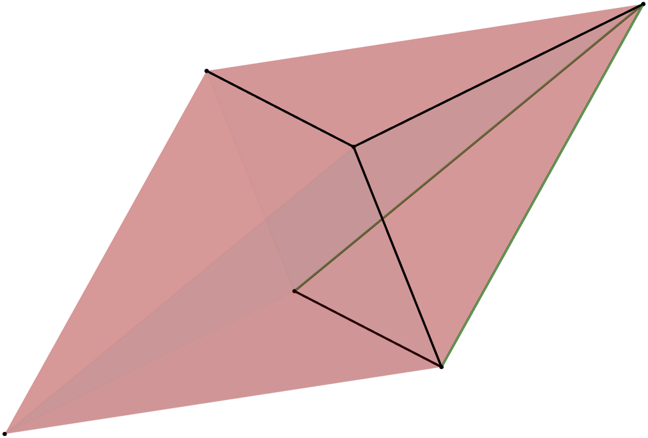
Example Edge Collapse



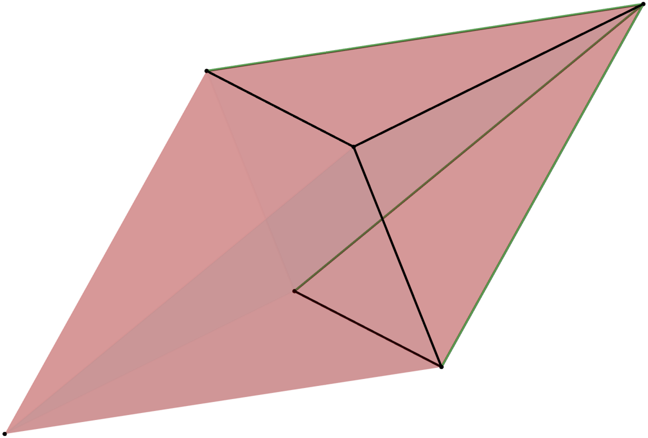
Example Edge Collapse



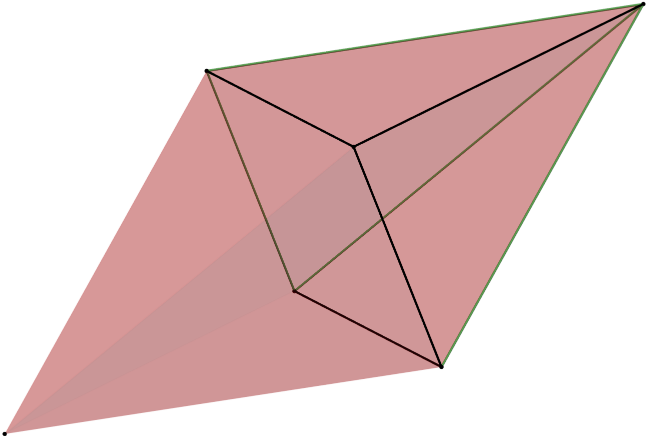
Example Edge Collapse



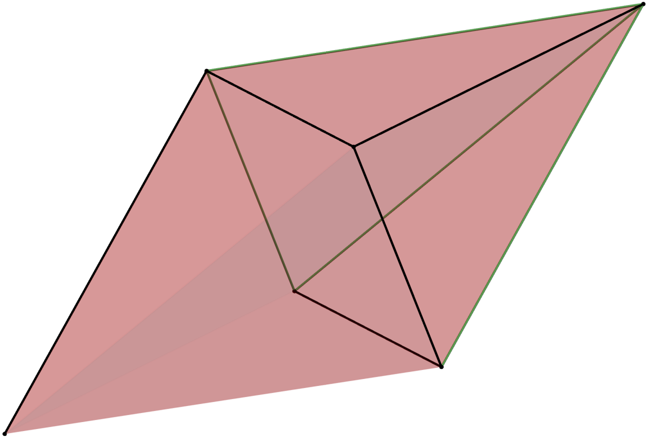
Example Edge Collapse



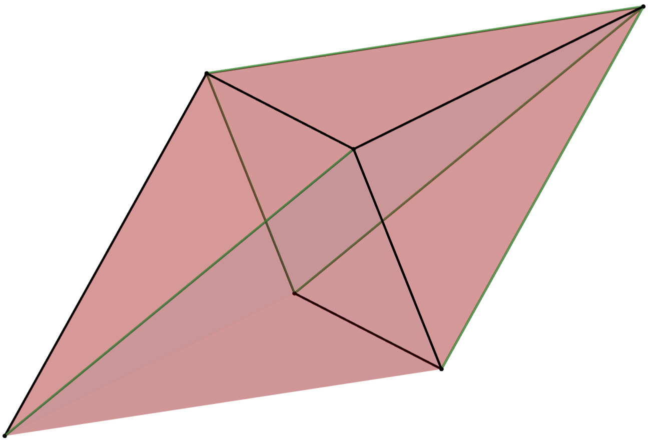
Example Edge Collapse



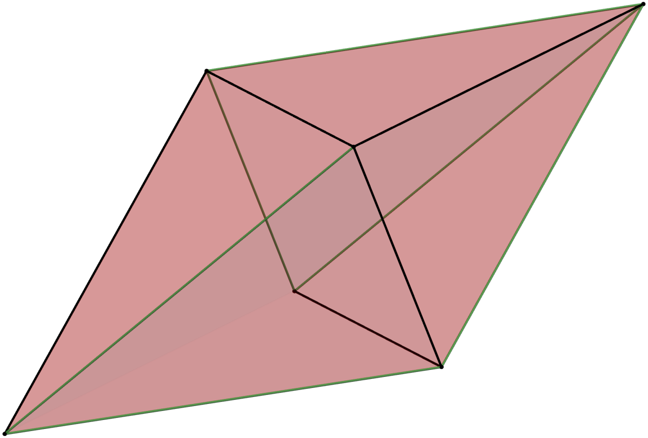
Example Edge Collapse



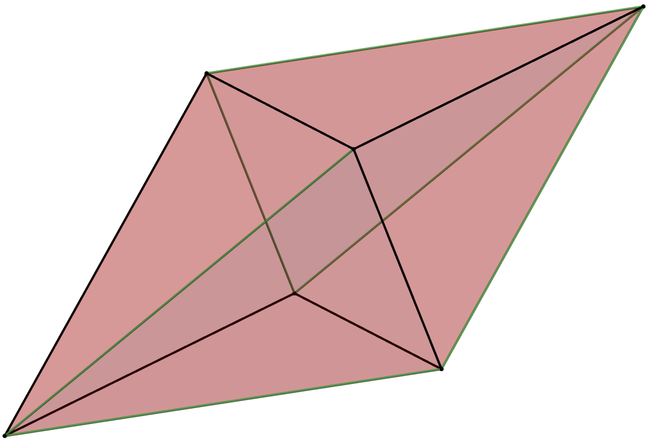
Example Edge Collapse



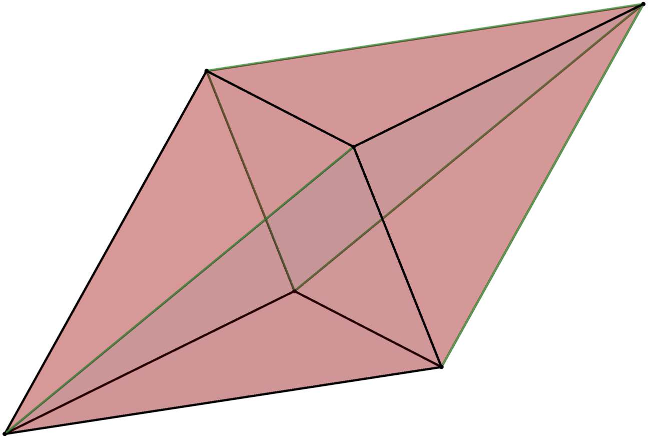
Example Edge Collapse



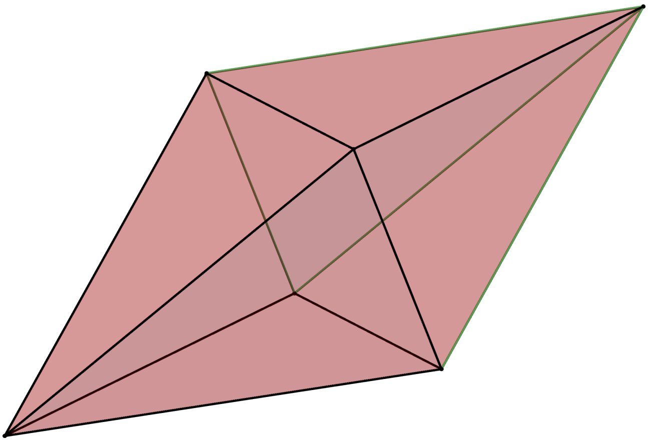
Example Edge Collapse



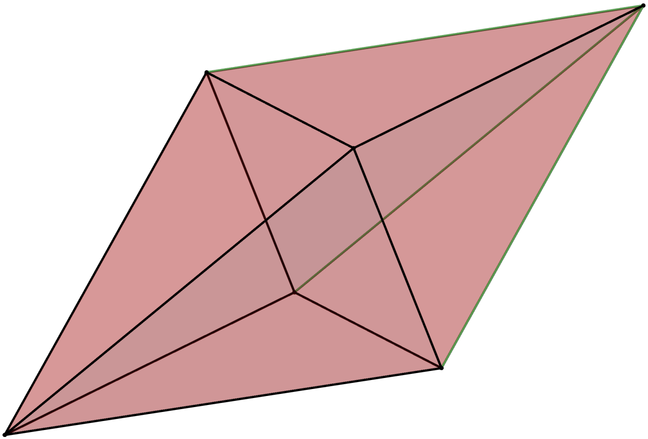
Example Edge Collapse



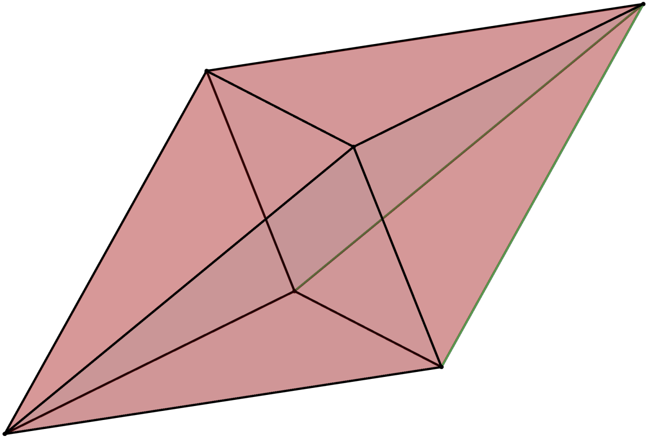
Example Edge Collapse



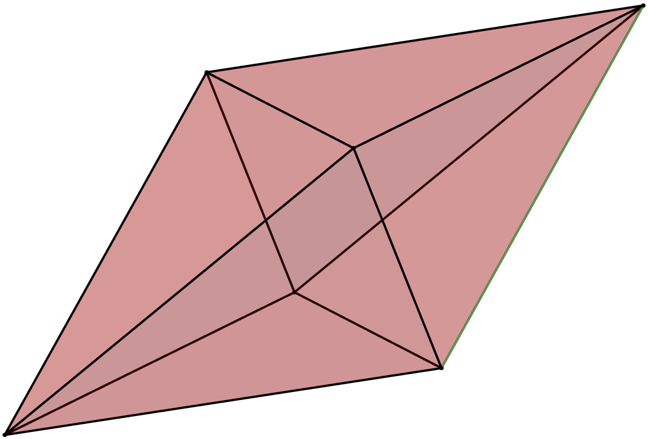
Example Edge Collapse



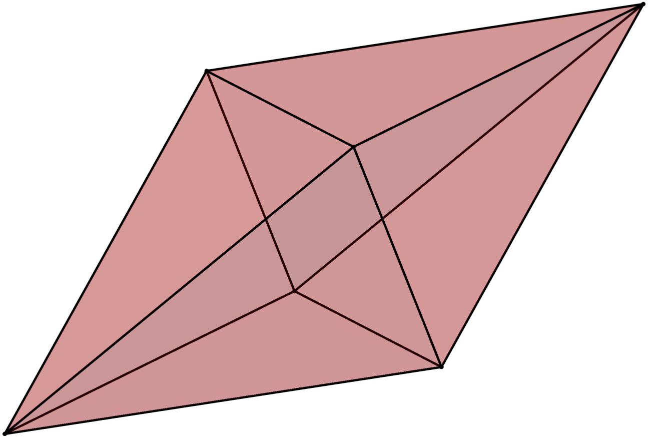
Example Edge Collapse



Example Edge Collapse



Example Edge Collapse



Computational Experiments

Data	Pnt	Thrsld	EdgeCollapser +PD				
			Edge(I)/Edge(C)	Size/Dim	*dim*	Pre-Time	Tot-Time
netw-sc	379	5.5	8.4K/417	1K/6	∞	0.62	0.73
senate	103	0.415	2.7K/234	663/4	∞	0.21	0.24
eleg	297	0.3	9.8K/562	1.8K/6	∞	1.6	1.7
HIV	1088	1050	182K/6.9K	86.9M/?	6	491	2789
torus	2000	1.5	428K/14K	44K/3	∞	288	289

Data	Pnt	Threshold	Ripser		Ripser		Ripser	
			dim	Time	*dim*	Time	*dim*	Time
netw-sc	379	5.5	4	25.3	5	231.2	6	∞
senate	103	0.415	3	0.52	4	5.9	5	52.3
"	"	"	6	406.8	7	∞		
eleg	297	0.3	3	8.9	4	217	5	∞
HIV	1088	1050	2	31.35	3	∞		
torus	2000	1.5	2	193	3	∞		

Discussion

- Efficient computation (of persistent homology) is a balance between making as few as possible computation steps and being frugal with memory (RAM)

Discussion

- Efficient computation (of persistent homology) is a balance between making as few as possible computation steps and being frugal with memory (RAM)
  - Reduction can be especially useful for high-dim. homology and to avoid exhausting the RAM.

Discussion

- Efficient computation (of persistent homology) is a balance between making as few as possible computation steps and being frugal with memory (RAM)
  - Reduction can be especially useful for high-dim. homology and to avoid exhausting the RAM.
  - If the homological information is mostly concentrated in lower dimensions the reduction might be less effective

Discussion

- Efficient computation (of persistent homology) is a balance between making as few as possible computation steps and being frugal with memory (RAM)
  - Reduction can be especially useful for high-dim. homology and to avoid exhausting the RAM.
  - If the homological information is mostly concentrated in lower dimensions the reduction might be less effective
- Complexity of edge collapse: $\mathcal{O}(nn_c k^2)$

Discussion

- Efficient computation (of persistent homology) is a balance between making as few as possible computation steps and being frugal with memory (RAM)
  - Reduction can be especially useful for high-dim. homology and to avoid exhausting the RAM.
  - If the homological information is mostly concentrated in lower dimensions the reduction might be less effective
- Complexity of edge collapse: $\mathcal{O}(n n_c k^2)$
- Questions and other remarks?