Topological simplification problems

Ulrich Bauer

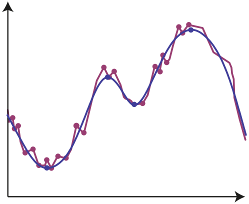
IST Austria

July 2, 2012

ATMCS 5, Edinburgh

Joint work with:
Carsten Lange, Max Wardetzky
Nina Amenta, Dominique Attali, Olivier Devillers, Marc Glisse, André Lieutier

Topological denoising of a function



Noise (even small) can create lots of critical points

Topological denoising by simplification of critical set:

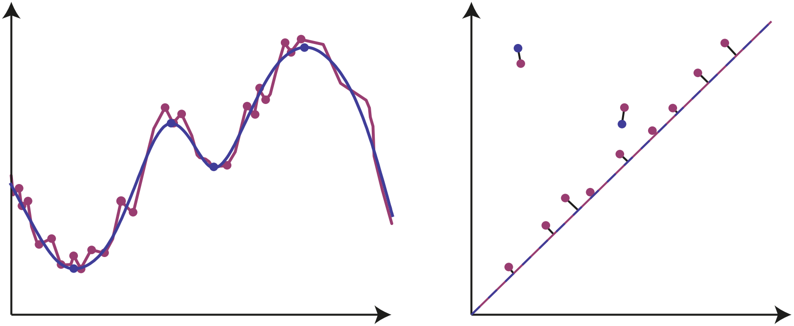
- removing critical points caused by noise

Problem

Given a function f and $\delta > 0$, find a function f_δ that:

- *minimizes number of critical points*
- *stays close to input function:* $\|f_\delta - f\|_\infty \leq \delta$

Persistence diagrams [Cohen-Steiner et al., 2005]

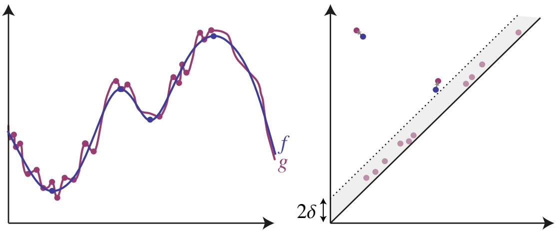


Stability of persistence diagrams

Theorem (Cohen-Steiner et al., 2005)

Let $\|f - g\|_\infty \leq \delta$.

- *The persistence pairs of* f *that have persistence* $> 2\delta$ *can be mapped injectively to the persistence pairs of* g.
- *Corresponding points* p_f, p_g *in the persistence diagrams have distance* $\|p_f - p_g\|_\infty \leq \delta$.



A bound on number of critical points

Corollary

Let f be a discrete Morse function and let $\delta > 0$.

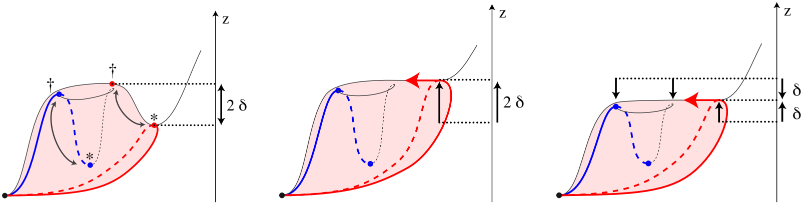
Then for every function f_δ with $\|f_\delta - f\|_\infty \leq \delta$ we have:

$$\begin{aligned} &\#\text{ critical points of } f_\delta \\ \geq &\#\text{ critical points of } f \text{ with persistence} > 2\delta. \end{aligned}$$

Side-effects of elimination

Idea for simplifying critical points [Edelsbrunner et al. 2006, Attali et al. 2009]:

- remove all persistence pairs of f with persistence $\leq 2\delta$
- leave all other persistence pairs unmodified



input   not optimal!   optimal

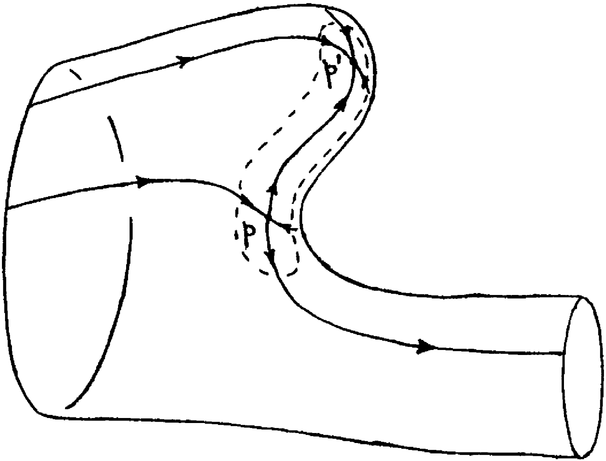
For an optimal solution, we must allow the critical values to change!

Interlude: discrete Morse theory

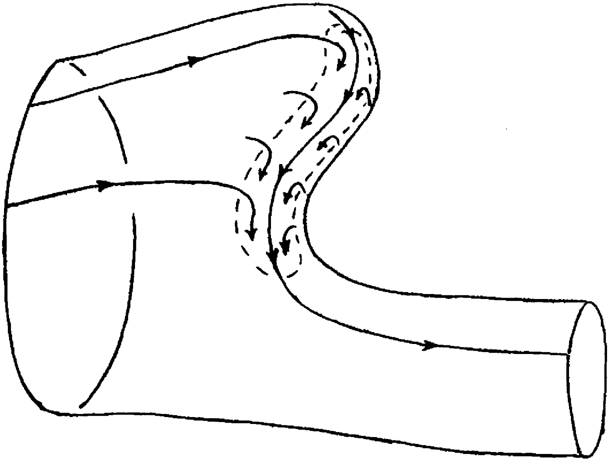
Since Vidit Nanda's talk today was canceled... What you need to know about discrete Morse theory for this talk:

- it's a discrete version of Morse theory for cell complexes
- there are discrete notions of gradient vector fields and critical points
- Morse functions are generic, nondegenerate functions (isolated critical points)
- discrete vector field: set of inequality constraints on function values
- homotopy type of sublevel sets changes only at critical values

Canceling critical points of a gradient field



Canceling critical points of a gradient field



Persistence pairs and Morse cancellations

Theorem (B., Lange, Wardetzky, 2011)

Let f be an excellent discrete Morse function on a surface (distinct critical values) with gradient field V. Any persistence pair (σ, τ) can be canceled in V after all persistence pairs $(\tilde{\sigma}, \tilde{\tau})$ with

$$f(\sigma) < f(\tilde{\sigma}) < f(\tilde{\tau}) < f(\tau)$$

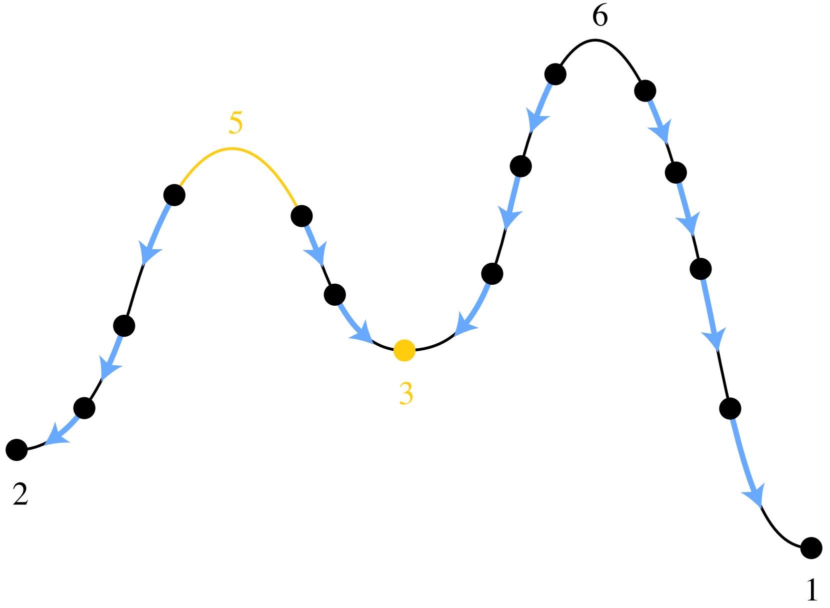
have been canceled.

Corollary

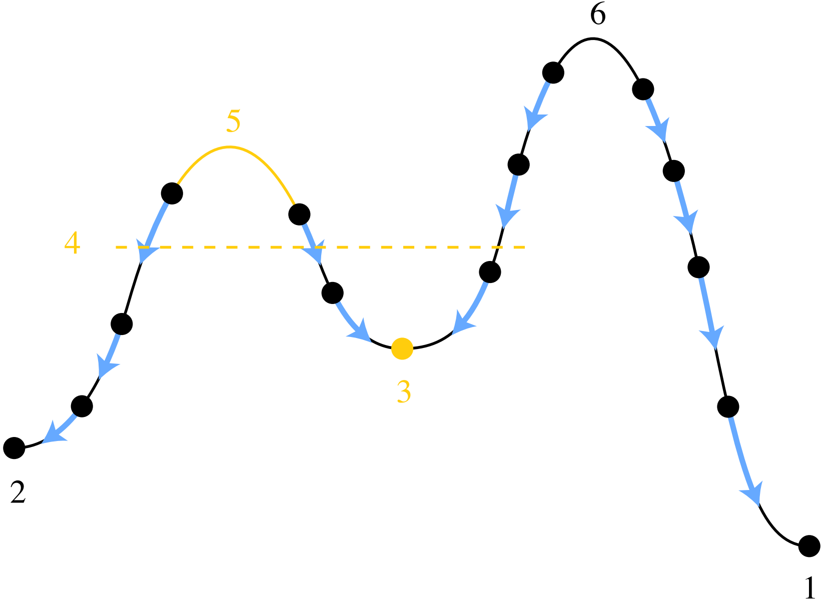
On a surface, it is possible to cancel just the persistence pairs with persistence $\leq 2\delta$ (without canceling the other pairs).

Does not hold in higher dimensions!

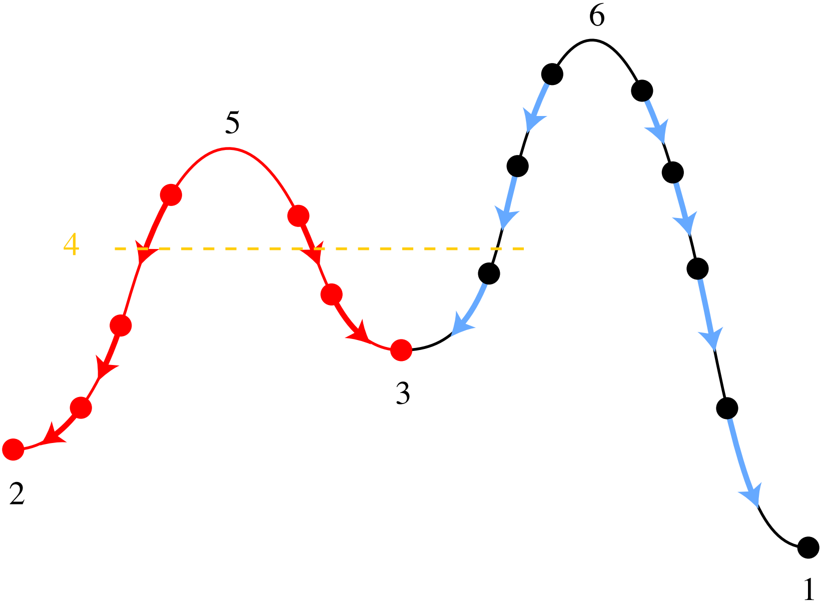
Canceling critical points of a function



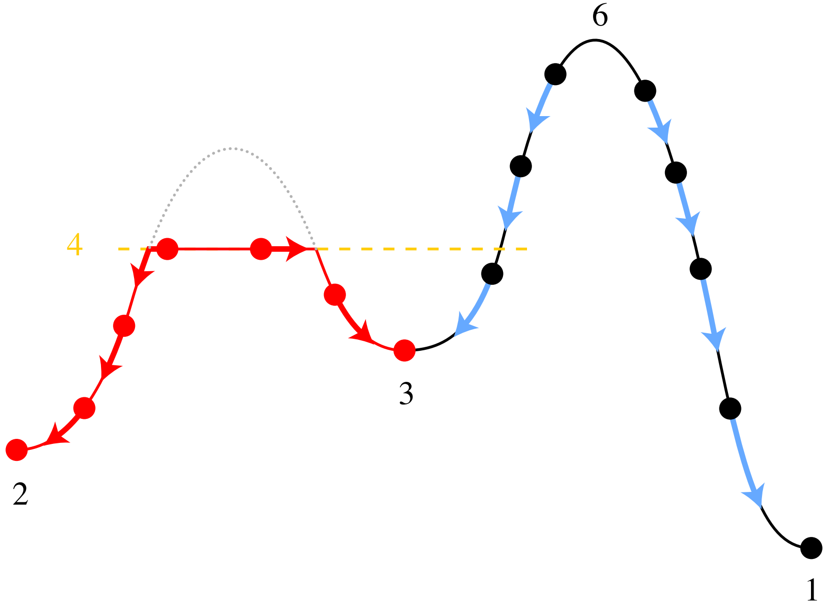
Canceling critical points of a function



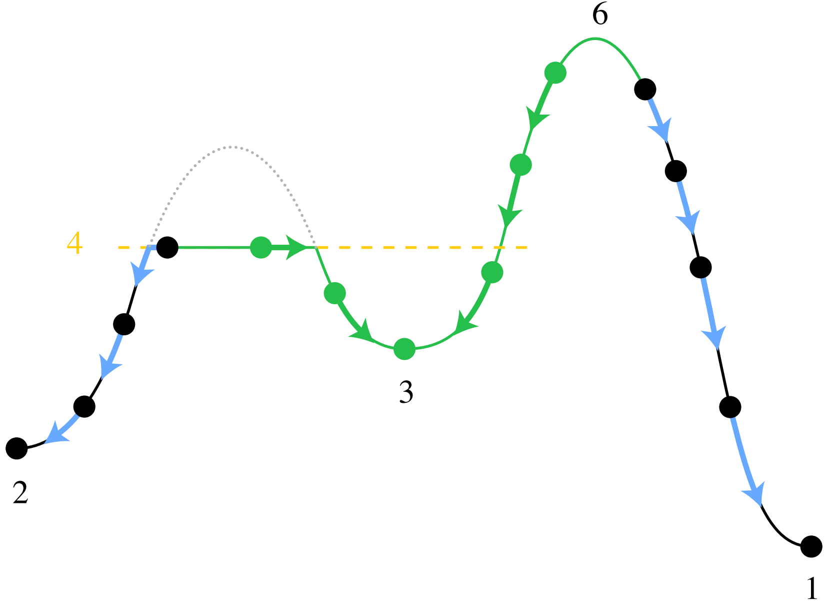
Canceling critical points of a function



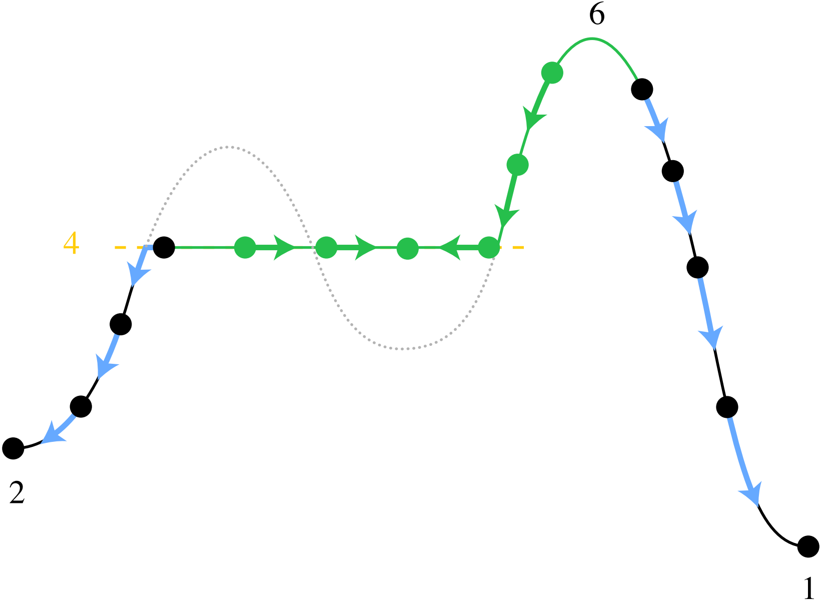
Canceling critical points of a function



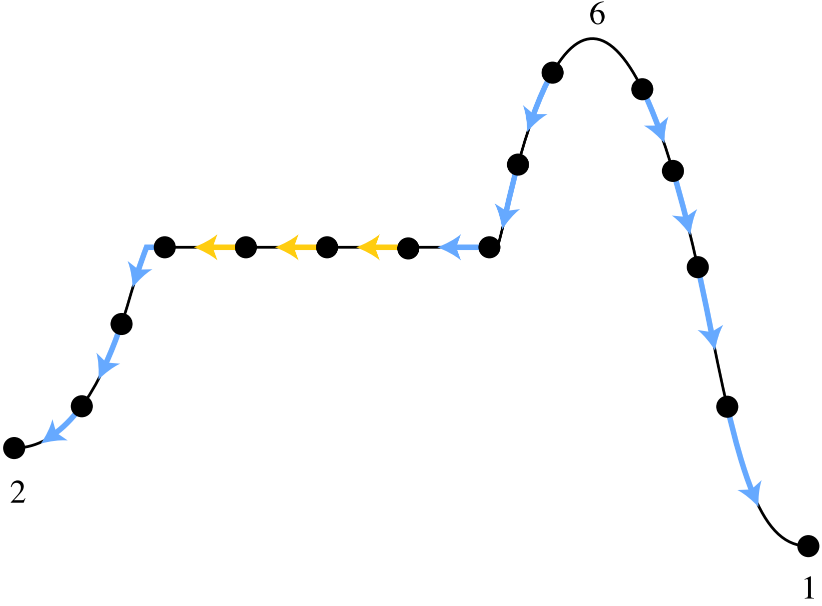
Canceling critical points of a function



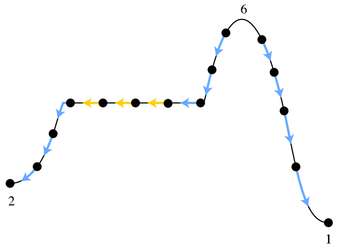
Canceling critical points of a function



Canceling critical points of a function



Degenerate functions



After cancelation, function is no longer Morse

- *pseudo-Morse:* replace strict inequalities by weak ones
  - f is *consistent* with V: if σ is facet of τ,
    $(\sigma, \tau) \notin V \Rightarrow f(\sigma) \leq f(\tau)$
    $(\sigma, \tau) \in V \Rightarrow f(\sigma) \geq f(\tau)$
  - closure of the set of discrete Morse functions
- Gradient vector field is no longer unique in general

Symbolic perturbation

- Use infinitesimal perturbations to resolve degeneracies

Let f pseudo-Morse, g excellent Morse (distinct critical values), both consistent with gradient vector field V.

- For any $\epsilon > 0$,
  $f_\epsilon = f + \epsilon g$ is an excellent Morse function consistent with V

Assume additionally that the order induced by g extends the order induced by f: $g(\sigma) < g(\tau) \Rightarrow f(\sigma) \leq f(\tau)$

- f_ϵ induces the same order as g: $g(\sigma) < g(\tau) \Leftrightarrow f_\epsilon(\sigma) < f_\epsilon(\tau)$
- the persistence pairs of f_ϵ are the persistence pairs of g

Most important statements allow passing to the limit $\epsilon \to 0$!

Optimal topological simplification

Theorem (B., Lange, Wardetzky, 2011)

Let f be a pseudo-Morse function on a surface and let $\delta > 0$. Let f_δ be obtained from f by canceling all persistence pairs with persistence $\leq 2\delta$. Then

$$\|f_\delta - f\|_\infty \leq \delta.$$

I.e., f_δ achieves the lower bound on the number of critical points.

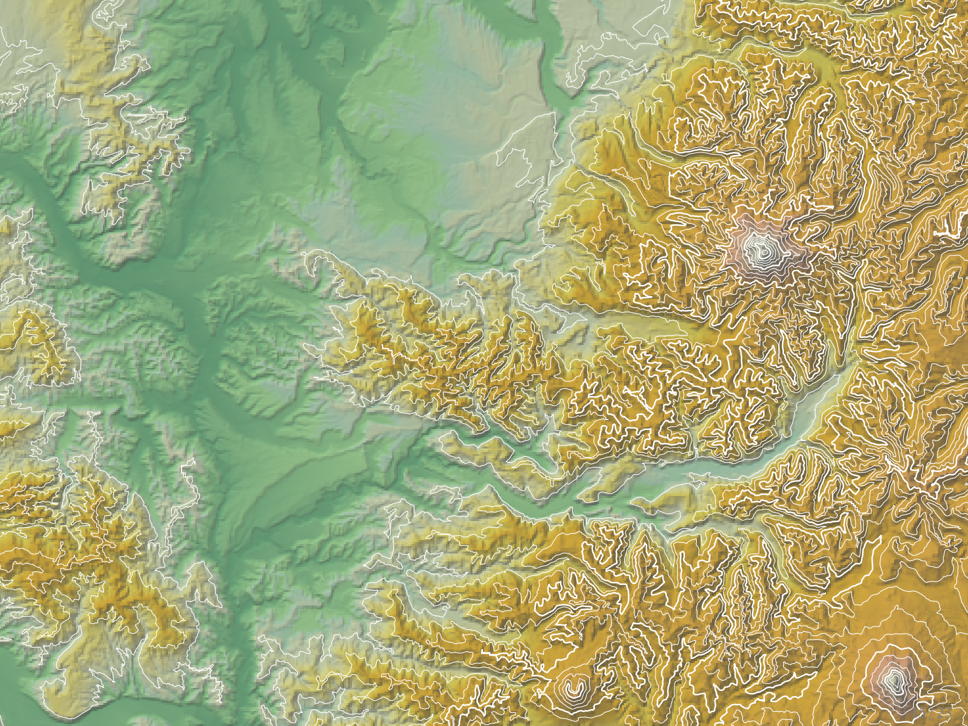
- Does not hold for non-manifold 2-complexes or higher dimensions (in general, simplification is NP-hard)
- Solution can be found in linear time after computation of persistence pairs

Combination with energy methods

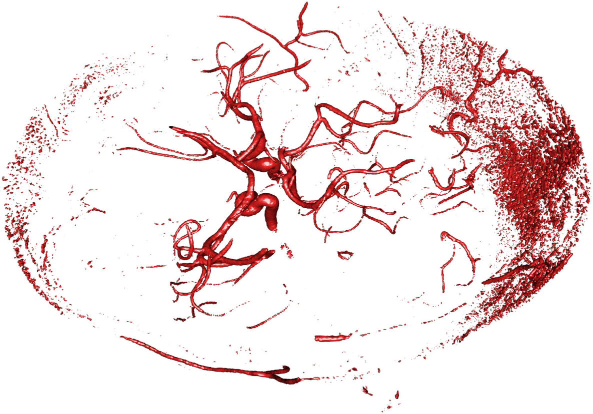
Recall: simplified vector field V_δ imposes inequalities on simplified function consistent with V_δ

$\|f_\delta - f\|_\infty \leq \delta$: another set of linear inequalities

- defines convex polytope of solutions:
  any function g consistent with V_δ and $\|g - f\|_\infty \leq \delta$ is a solution
- find the “best” solution using your favorite energy functional



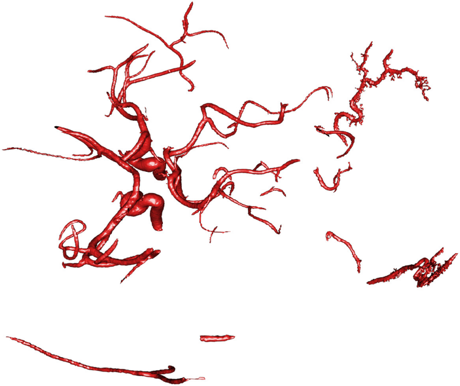
Removing local extrema from 3D data



In 3D:

- simplifying critical points is hard
- simplifying only extrema is easy

Removing local extrema from 3D data



In 3D:

- simplifying critical points is hard
- simplifying only extrema is easy

Sublevel set simplification

Let $F_{\leq t} = f^{-1}(-\infty, t]$ denote the t-sublevel set of f.

Problem

Given a PL function $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ *and* $t \in \mathbb{R}$, $\delta > 0$,
find a PL function g *with* $\|g - f\|_\infty \leq \delta$
minimizing $\beta_*(G_{\leq t})$.

Let $K = F_{\leq t+\delta}$, $L = F_{\leq t-\delta}$.

- For any g, we have $L \subset G_{\leq t} \subset K$.
- For any X with $L \subset X \subset K$, there is g with $G_{\leq t} = X$.

Thus we are looking for X with $L \subset X \subset K$ minimizing $\beta_*(X)$.

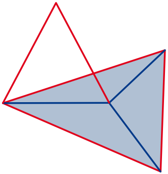
Homological factorization

Problem

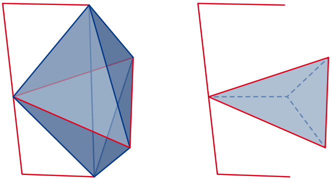
Given a simplicial pair (K, L), find X with $L \subset X \subset K$ such that $H_*(L \hookrightarrow X)$ is surjective and $H_*(X \hookrightarrow K)$ is injective.

Such an X is called a *homological factorization* of (K, L).

- If $L \subset X \subset K$ then $\beta_*(X) \geq \operatorname{rank} H_*(L \hookrightarrow K)$
- in $\mathbb{R}^3$: equality iff X is a homological factorization



homological factorizations do not always exists



a homological factorization may exist, but not as a subcomplex of K

Homological factorizability in $\mathbb{R}^3$ is NP-complete

Theorem (Attali, Lieutier; 2010)

Deciding whether (K, L) has a homological factorization as a subcomplex is NP-complete.

Theorem (A., A., B., D., G., L.; 2012)

This holds even for complexes K embeddable in $\mathbb{R}^3$.

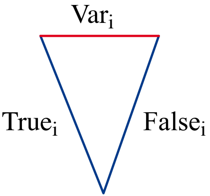
Corollary

Sublevel set simplification in $\mathbb{R}^3$ is NP-hard.

Idea of proof: reduction from 3-SAT

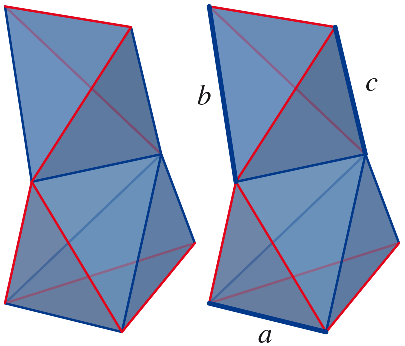
- given a 3-SAT instance, construct a simplicial pair (K, L) with trivial persistent homology group $H_*(L \hookrightarrow K)$
- X is homological factorization $\Leftrightarrow$ X is acyclic, $L \subset X \subset K$

Reduction from 3-SAT: the variable gadget



- red: contained in L, blue: $K \setminus L$
- X can contain only one of the edges $\text{True}_i, \text{False}_i$
- edges $\text{True}_i, \text{False}_i$ correspond to truth assignment of variable x_i

Reduction from 3-SAT: the clause gadget



- red: contained in L, blue: $K \setminus L$
- X must contain one of the edges a, b, c

- For every clause (e.g., $(x_1 \vee \neg x_2 \vee x_4)$): identify a, b, c with edges of variable gadgets corresponding to the literals $(\text{True}_1, \text{False}_2, \text{True}_4)$
- We can transform a homological factorization of (K, L) into a satisfying assignment and vice versa

Simplification of level sets

Theorem

Level set simplification in $\mathbb{R}^3$ is NP-hard.

Idea of proof:

Assume $K = F_{\leq t+\delta}$, $L = F_{\leq t-\delta}$. Let g be a t-level set simplification of f (a function g minimizing $\beta_*(G_{=t})$).
Then $G_{\leq t}$ is a homological factorization of (K, L), if one exists.

- $\beta_*(G_{=t}) = \beta_*(G_{<t}) + \beta_*(G_{\leq t})$
- g may be assumed to have regular value t. Hence $\beta_*(G_{=t}) = 2\beta_*(G_{\leq t})$
- If a homological factorization exists, then the lower bound on $\beta_*(G_{\leq t})$ can be achieved
- since g minimizes $\beta_*(G_{=t}) = 2\beta_*(G_{\leq t})$, it achieves this bound

Thanks for your attention!