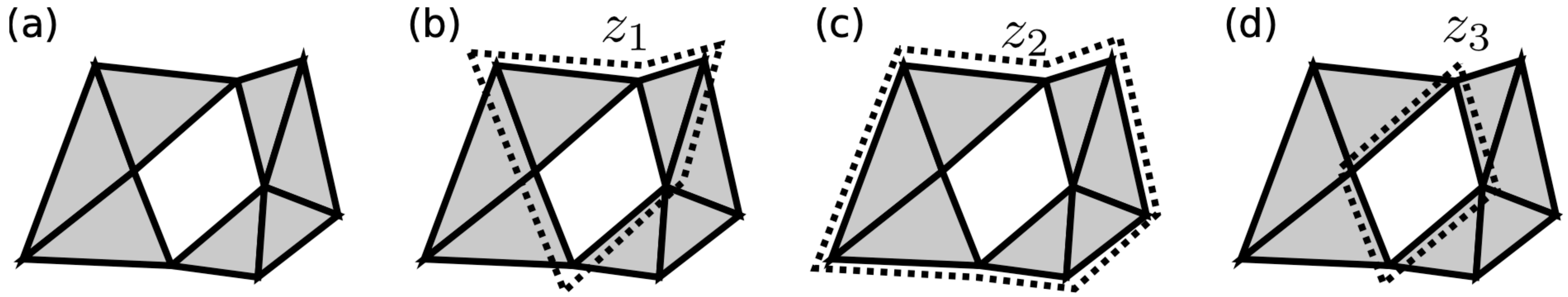
Topological Data Analysis

Lecture 10

Homology representatives

Oleg Kachan

Homology representatives

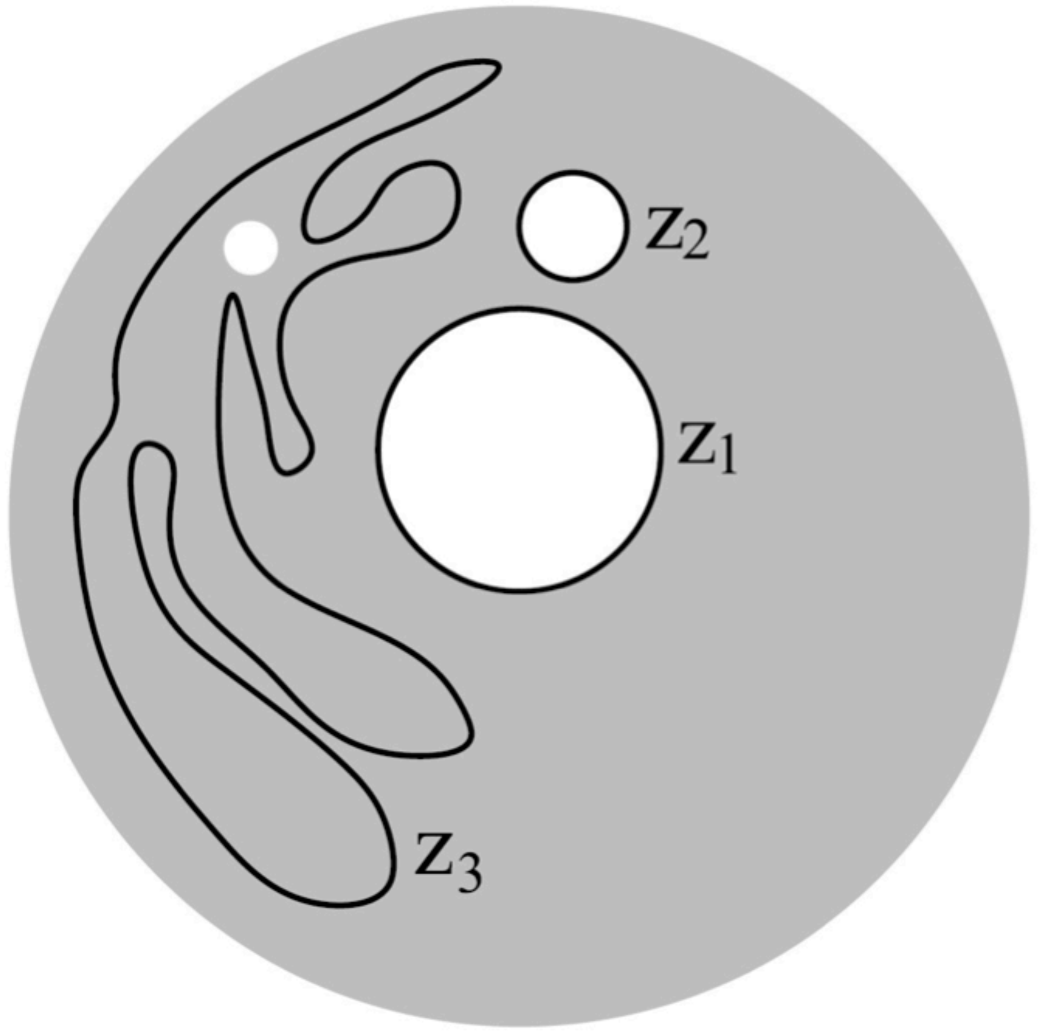


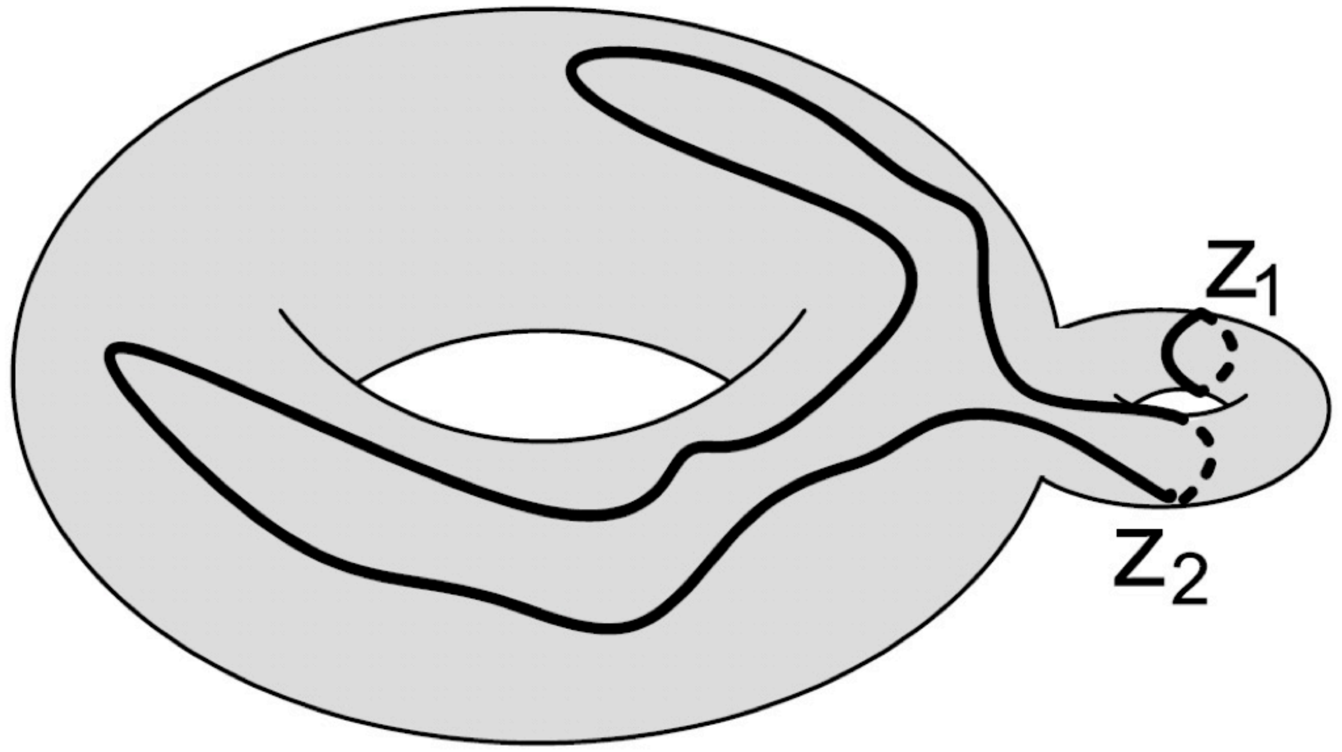
Homologous cycles

$$z \sim z' \iff z - z' \in B_k$$

$$B_k = \{c \in C_k \mid \partial_{k+1} d = c, \text{ for some } d \in C_{k+1}\}$$

Homology representatives

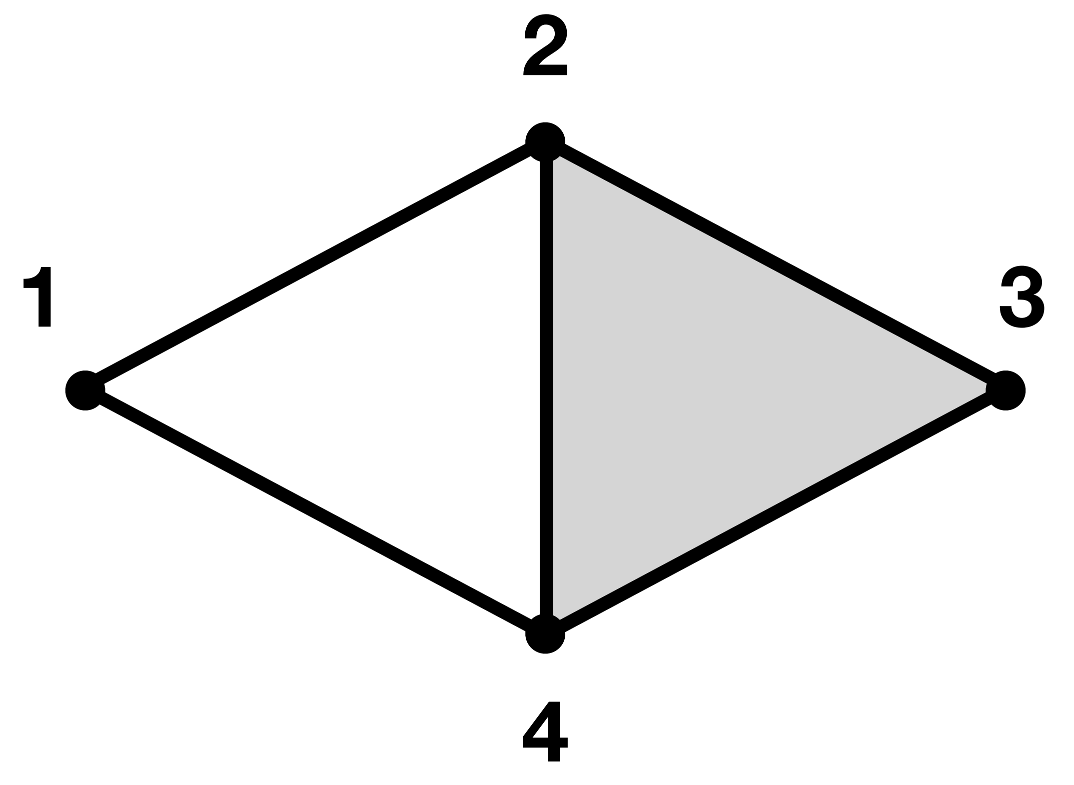




Optimal cycle

$$z^* = \arg\min_{z} \ell(z_0) \quad s.t.\ z \sim z_0$$

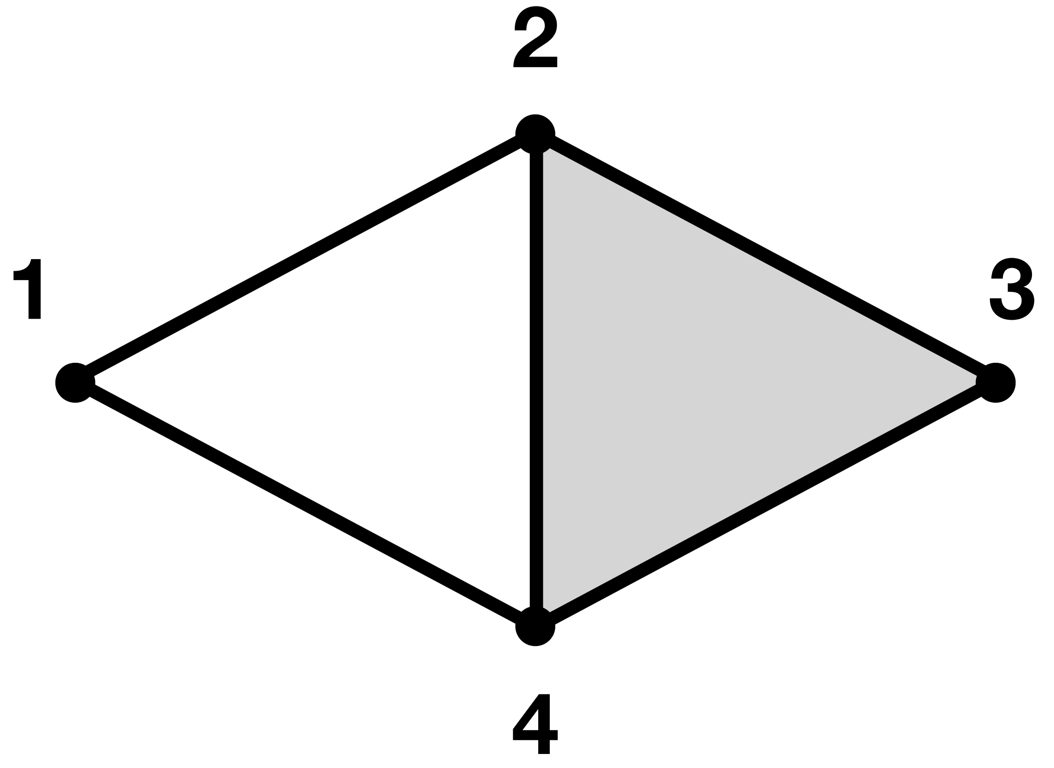
Boundary matrix



$\mathbf{B} =$

<table>
<tr><th></th><th>1</th><th>2</th><th>3</th><th>4</th><th>12</th><th>14</th><th>23</th><th>24</th><th>34</th><th>234</th></tr>
<tr><td>0</td><td colspan="4">B_0</td><td colspan="5"></td><td></td></tr>
<tr><td>1</td><td colspan="4" rowspan="4"></td><td colspan="5" rowspan="4">B_1</td><td rowspan="4"></td></tr>
<tr><td>2</td></tr>
<tr><td>3</td></tr>
<tr><td>4</td></tr>
<tr><td>12</td><td colspan="4" rowspan="5"></td><td colspan="5" rowspan="5"></td><td rowspan="5">B_2</td></tr>
<tr><td>14</td></tr>
<tr><td>23</td></tr>
<tr><td>24</td></tr>
<tr><td>34</td></tr>
</table>

Boundary matrix

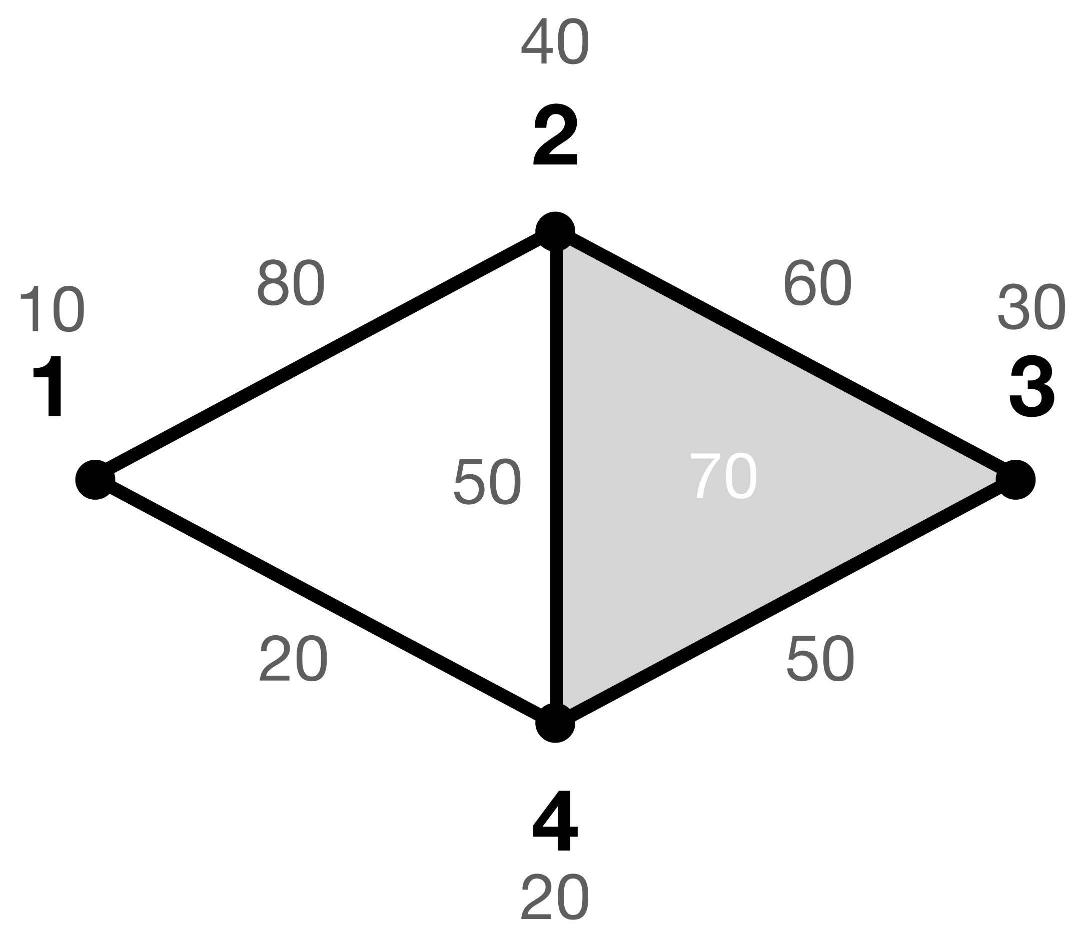


$\mathbf{B} =$

	1	2	3	4	12	14	23	24	34	234
0	1	1	1	1						
1					1	1				
2					1		1	1		
3							1		1	
4						1		1	1	
12										
14										
23										1
24										1
34										1

One may skip adding $\mathbf{B}_0$ to the full matrix or zero out it

Filtration function



A function $f : K \to \mathbb{R}$ is called a filtration function iff, either

$$f(\tau) \leq f(\sigma) \iff \tau \subseteq \sigma$$ (sublevel filtration)

$$f(\tau) \geq f(\sigma) \iff \tau \supseteq \sigma$$ (superlevel filtration)

$$K_t = \{\sigma \in K \mid f(\sigma) \leq t\}$$

sublevel set $t \in (-\infty, +\infty)$

$$K^t = \{\sigma \in K \mid f(\sigma) \geq t\}$$

superlevel set $t \in (+\infty, -\infty)$

A filtration is a sequence of sublevel (superlevel) sets s.t.

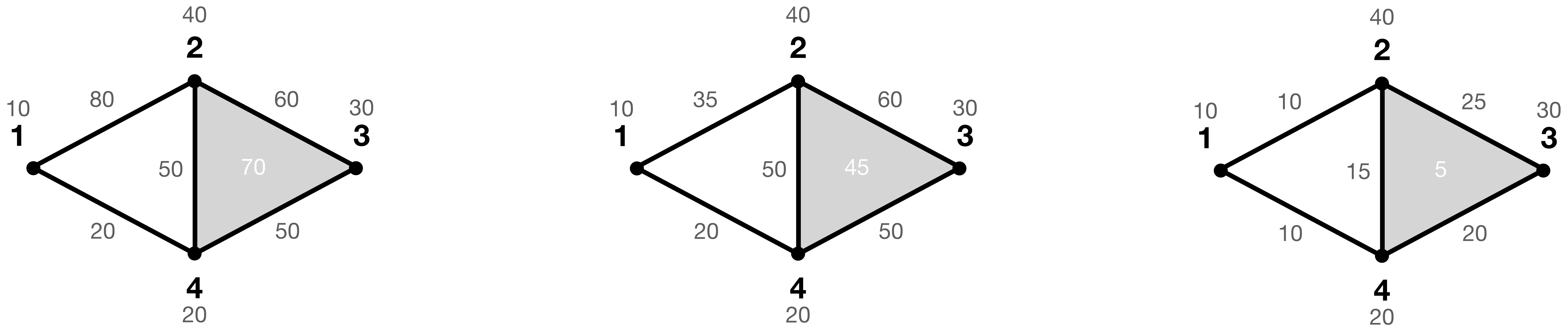
$$\varnothing \subset K_1 \subset K_2 \subset \ldots \subset K$$

sublevel filtration

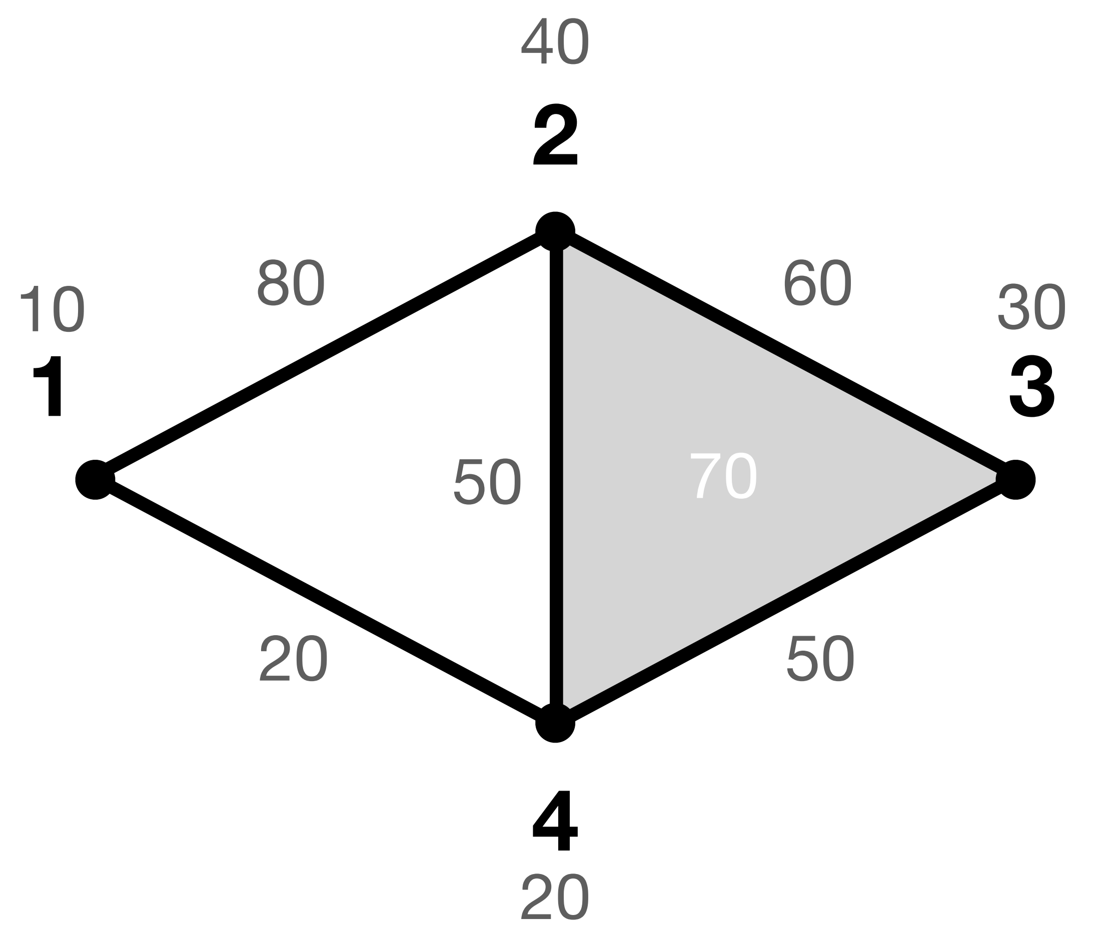
$$\varnothing \subset K^1 \subset K^2 \subset \ldots \subset K$$

superlevel filtration

Filtration function



Reordering w.r.t. filtration function



B =

	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
0										
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

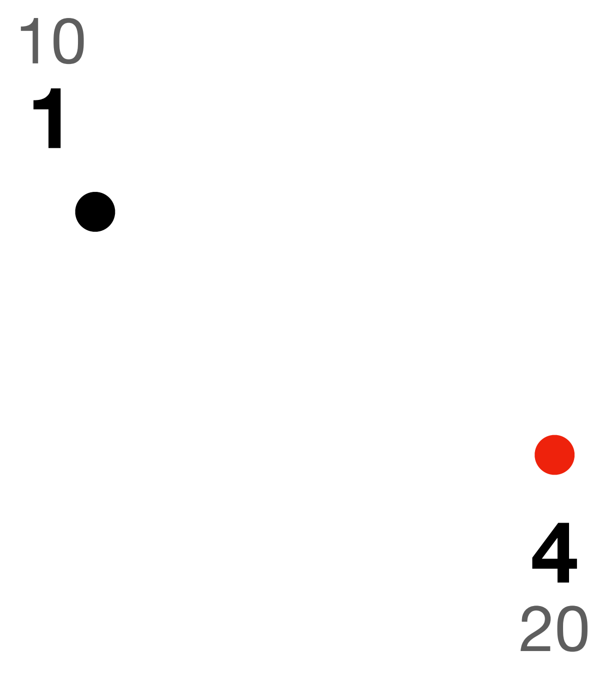
Filtration function provides order on simplices, therefore on columns and rows of the filtration matrix. Ties are broken, first by simplex dimension, second by lexicographic order given by order on vertices.

Boundary matrix reduction

	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

t=10

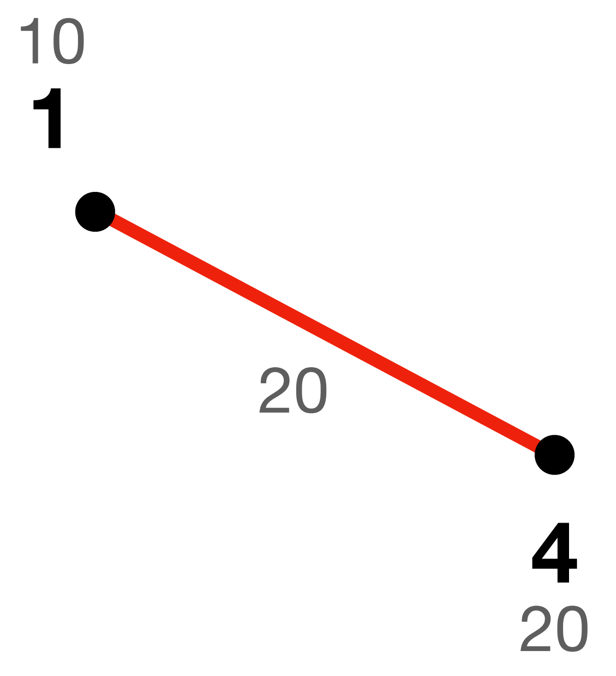
Boundary matrix reduction



	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

t=20

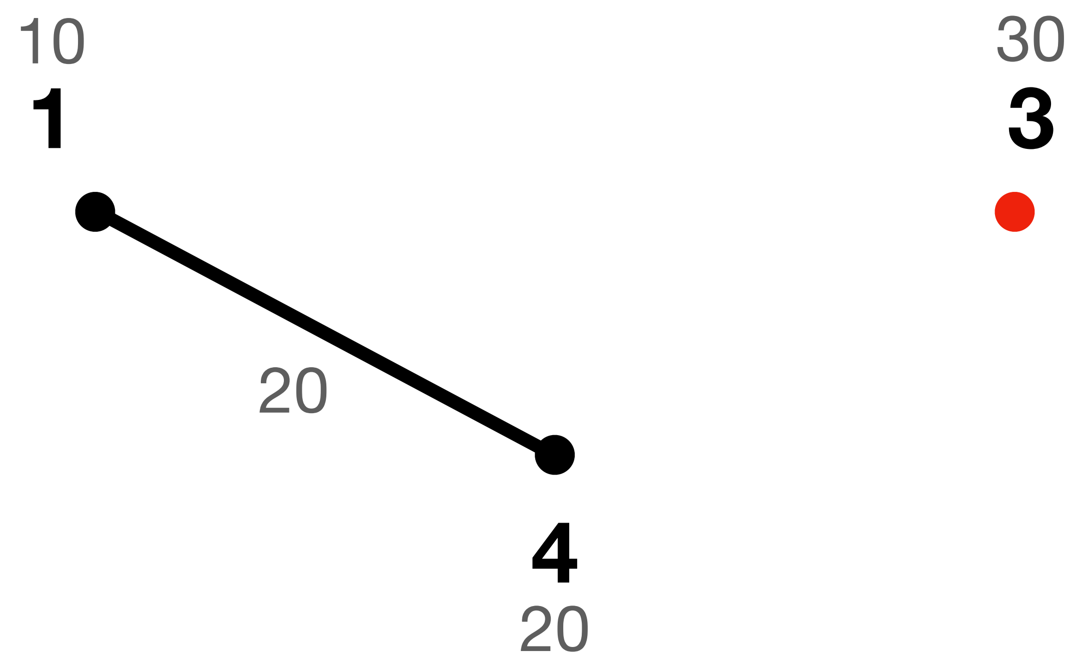
Boundary matrix reduction



	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			**1**			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

t=20

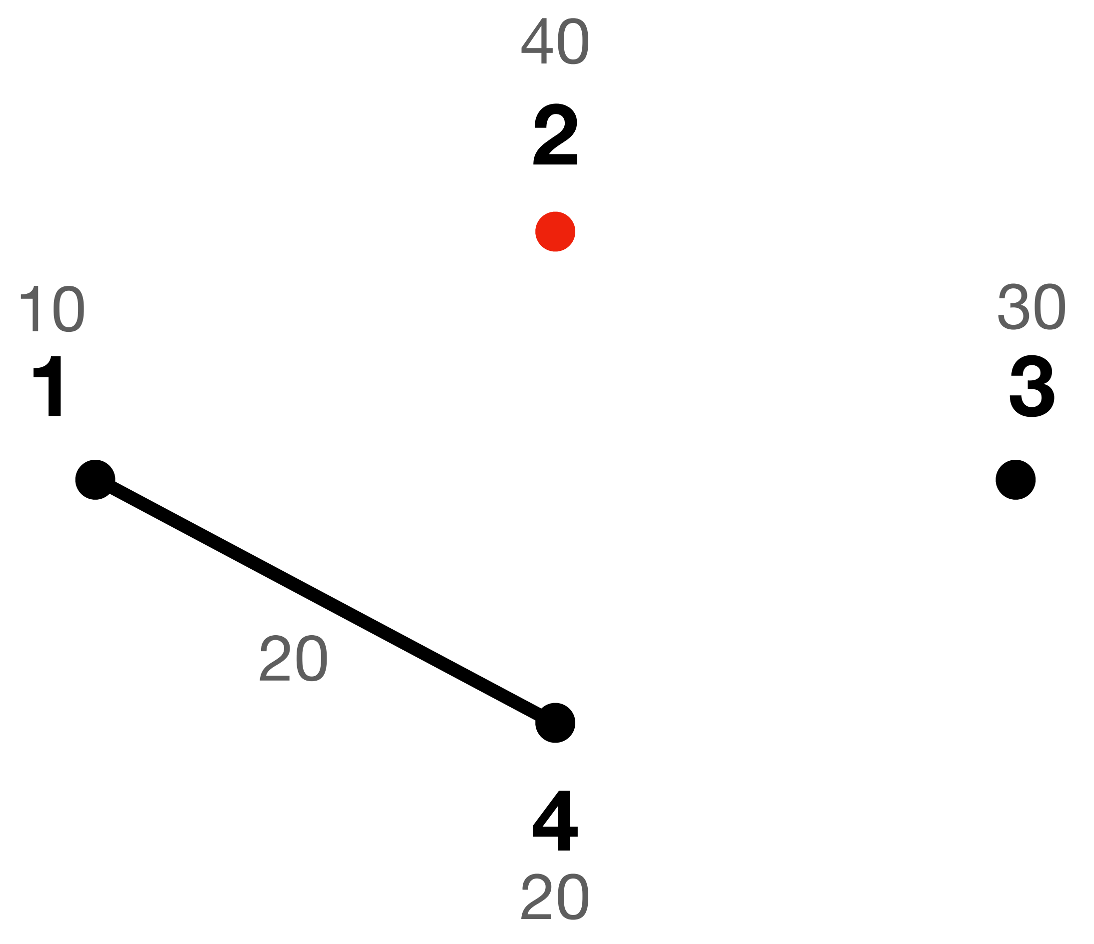
Boundary matrix reduction



	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

t=30

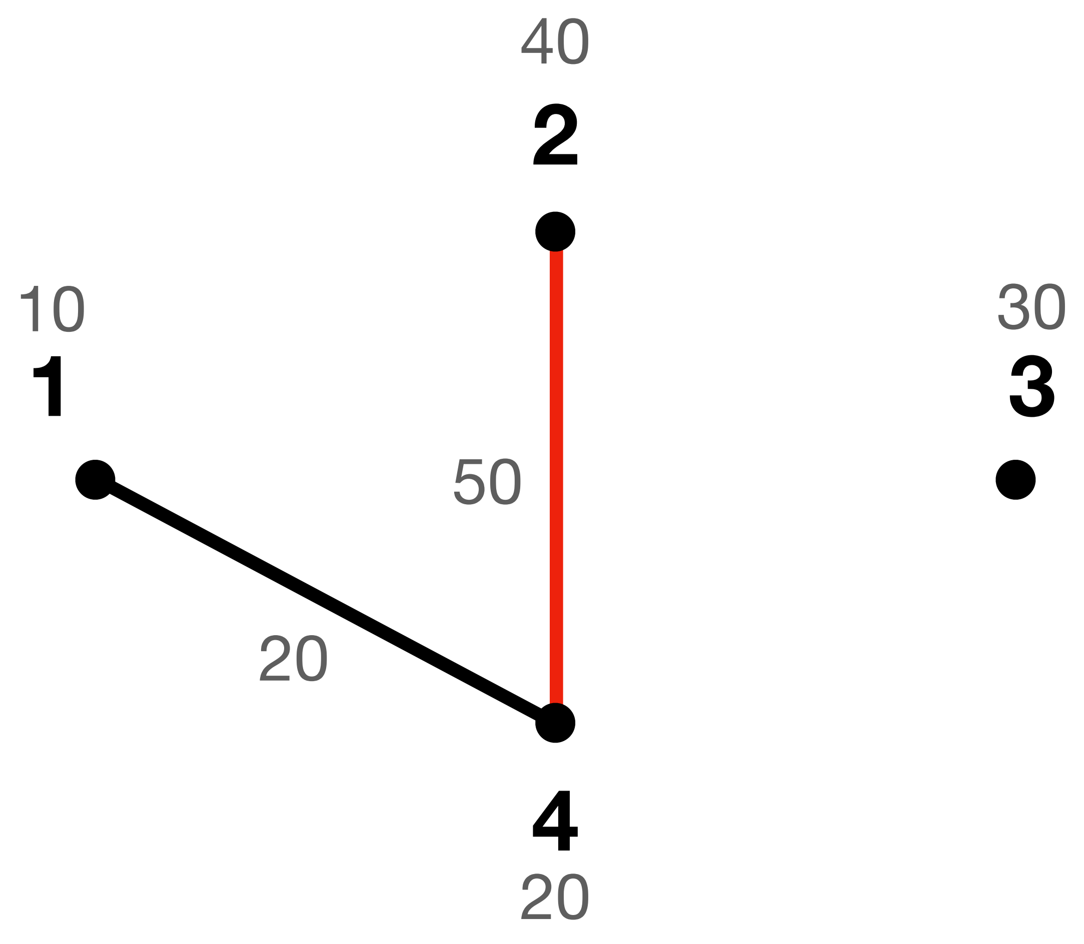
Boundary matrix reduction



	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

t=40

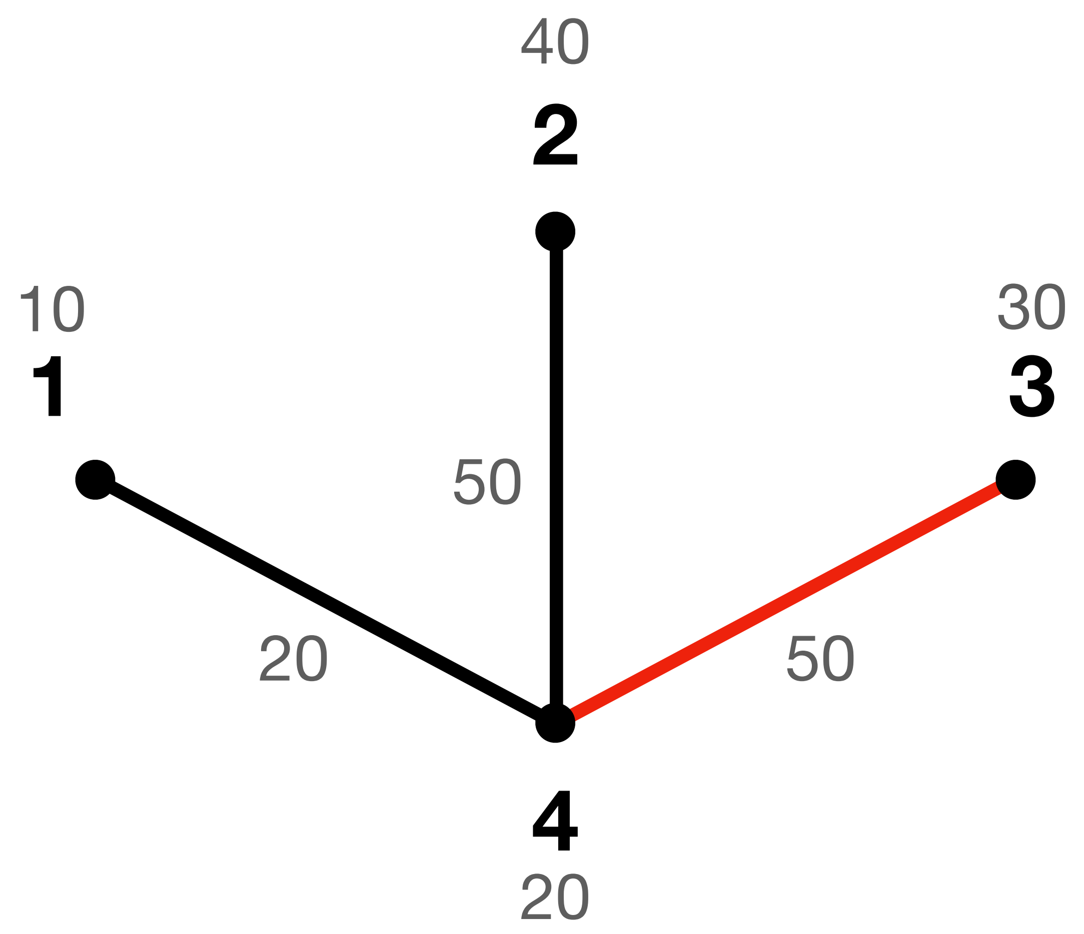
Boundary matrix reduction



	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

t=50

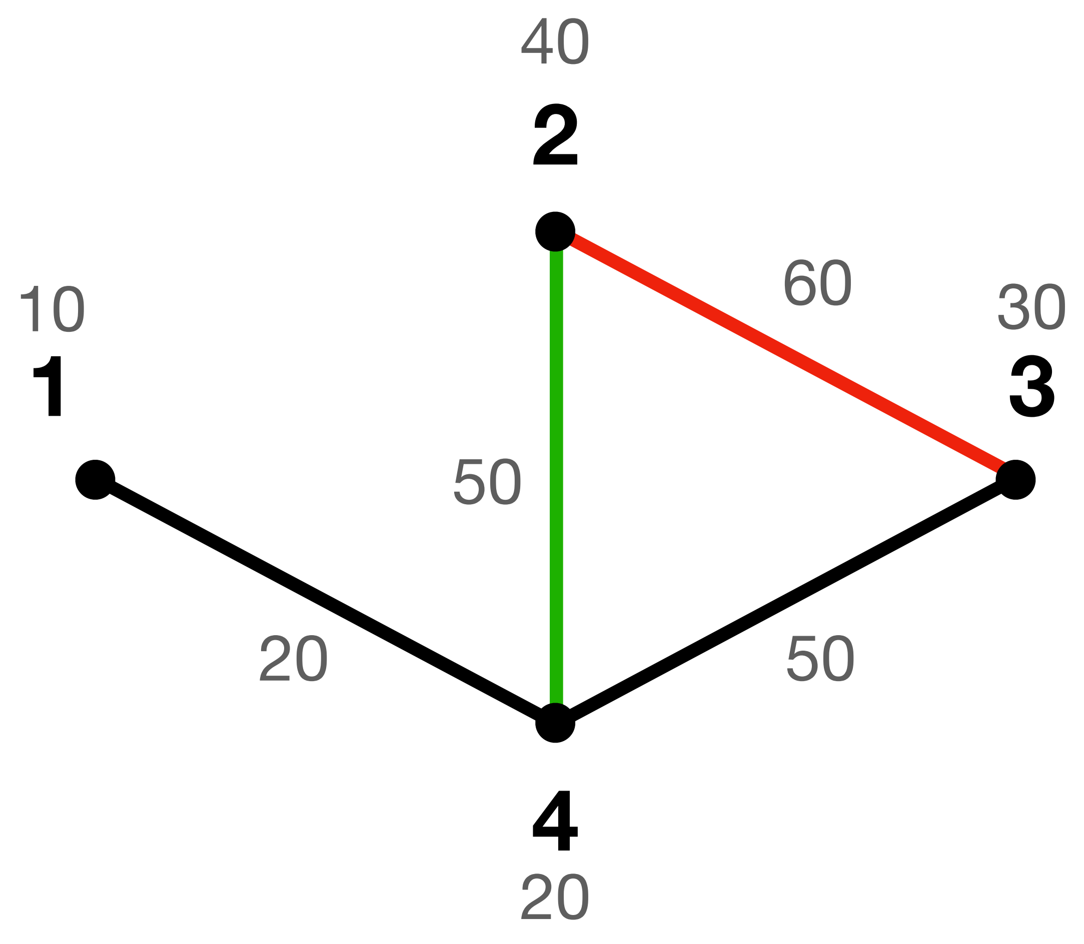
Boundary matrix reduction



	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

t=50

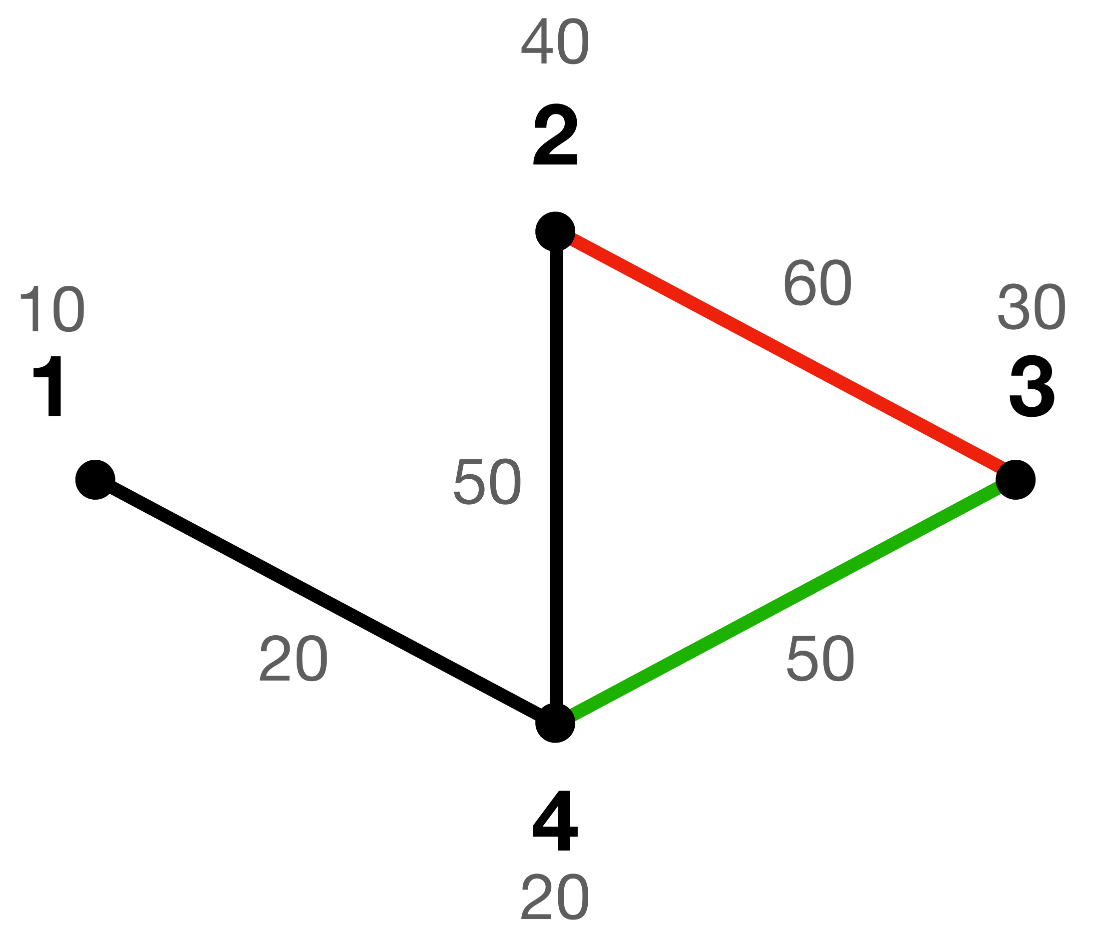
Boundary matrix reduction



	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

t=60

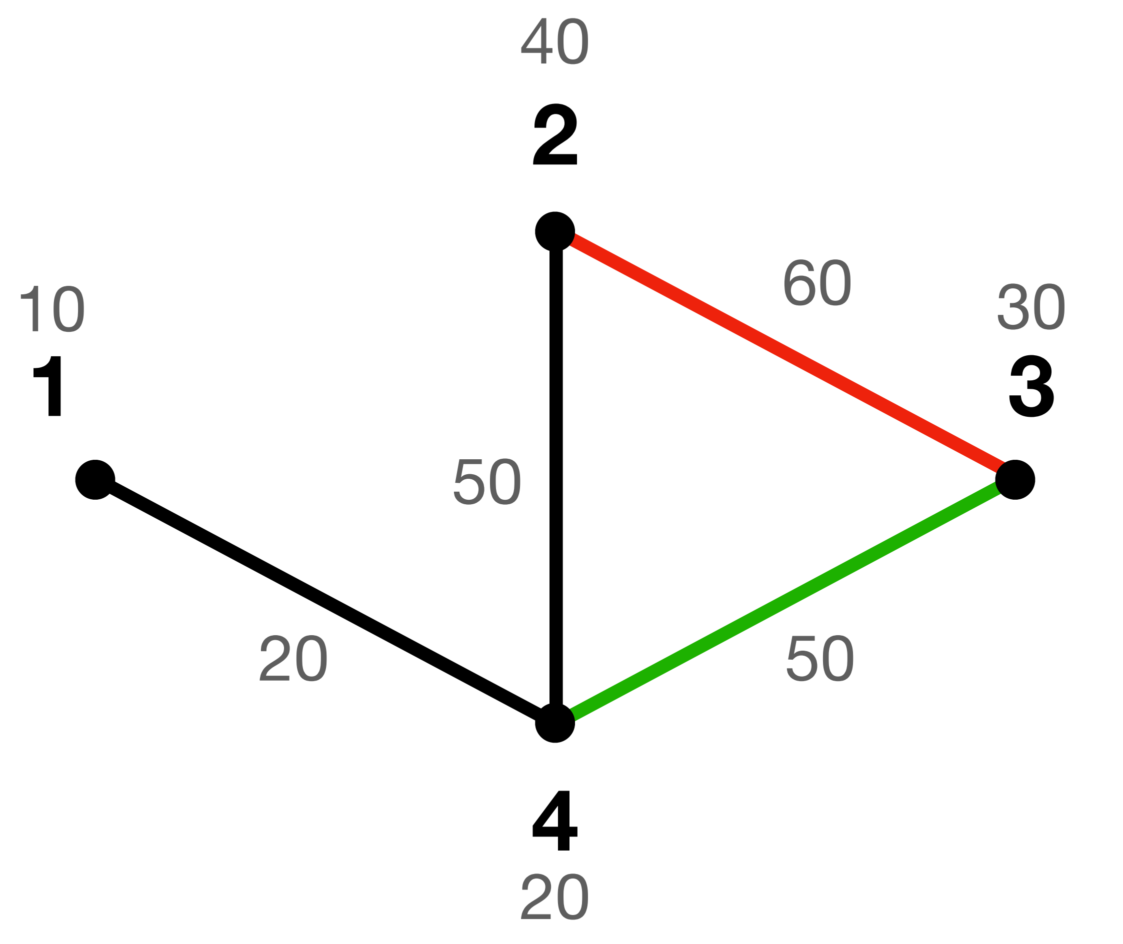
Boundary matrix reduction



								[23+24]		
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			**1**			1	1	1		
14										
3							**1**	1		
2						**1**				1
24									1	
34									1	
23									1	
12										

t=60

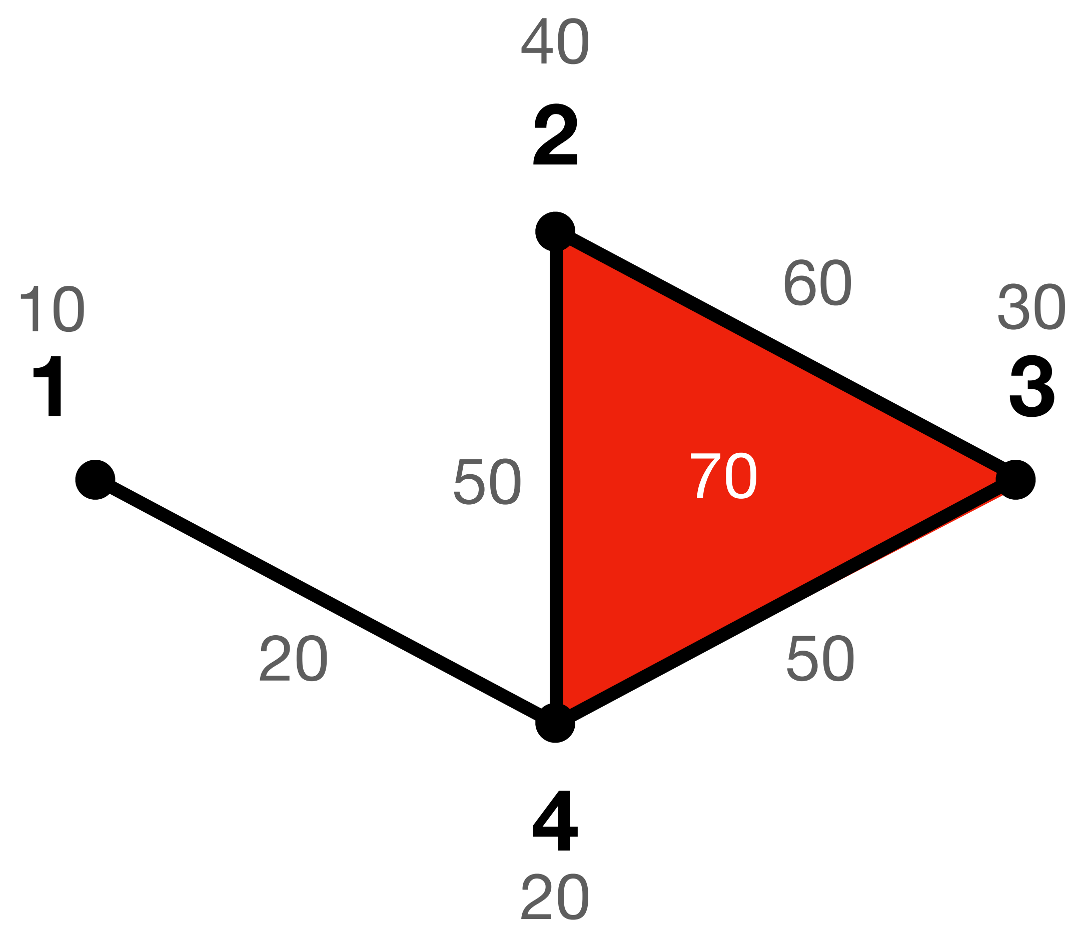
Boundary matrix reduction



								[23+24+34]		
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1			
2						1				1
24									1	
34									1	
23									1	
12										

t=60

Boundary matrix reduction

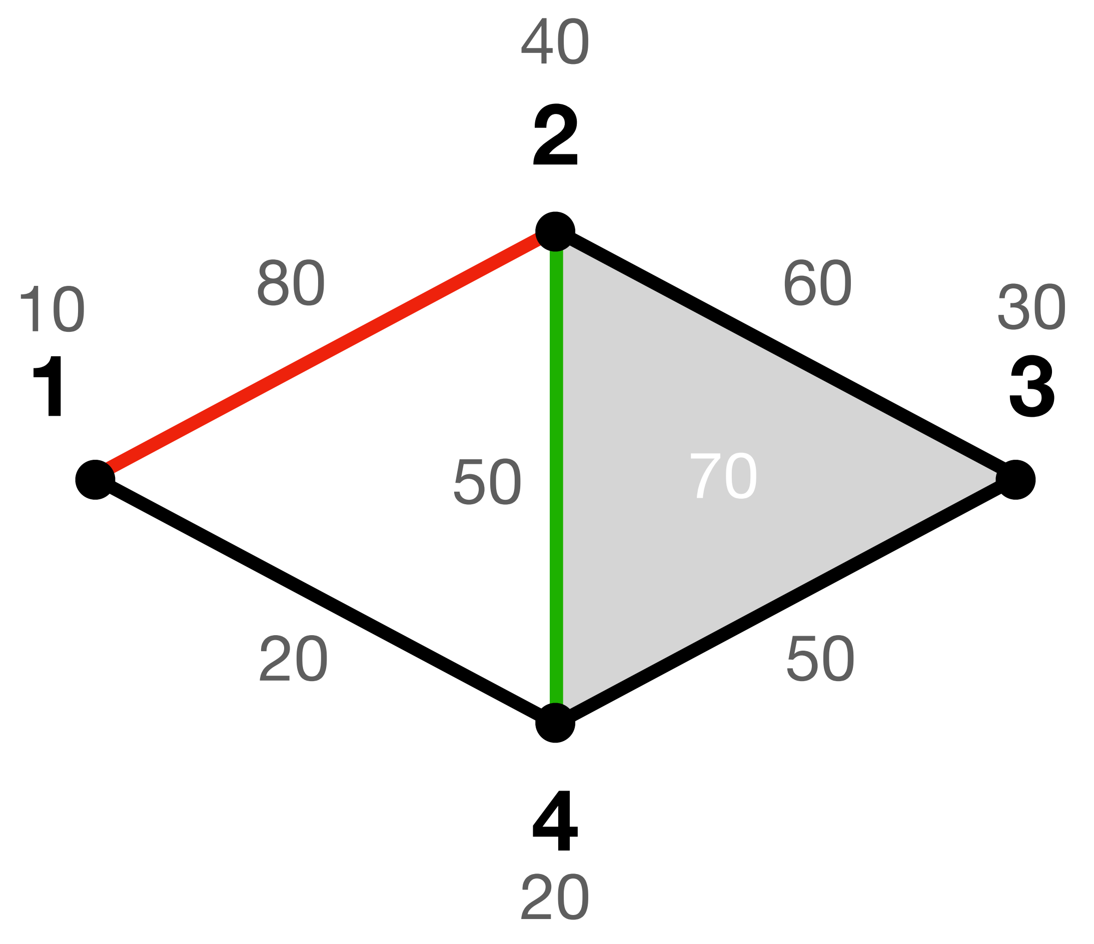


[23+24+34]

	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1			
2						1				1
24									1	
34									1	
23									1	
12										

t=70

Boundary matrix reduction

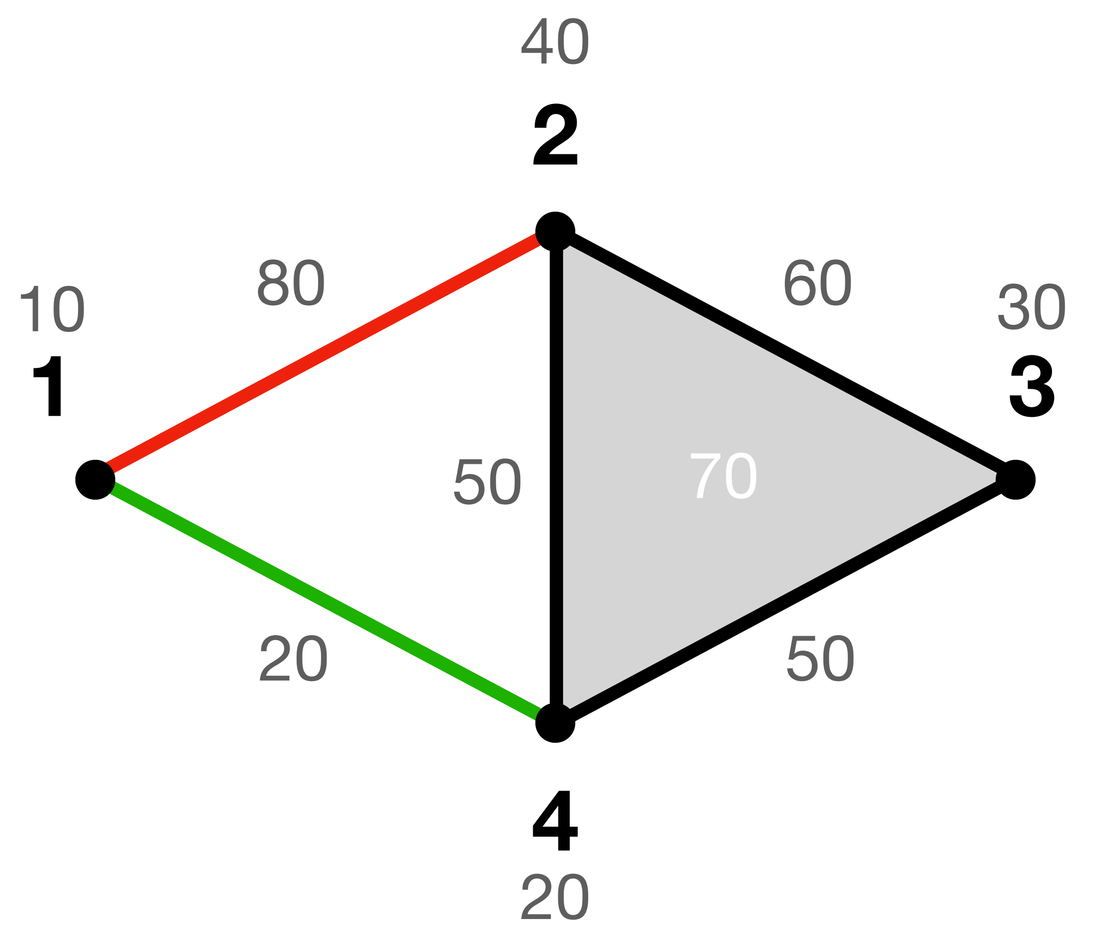


[23+24+34]

	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1			
2						1				1
24									1	
34									1	
23									1	
12										

t=80

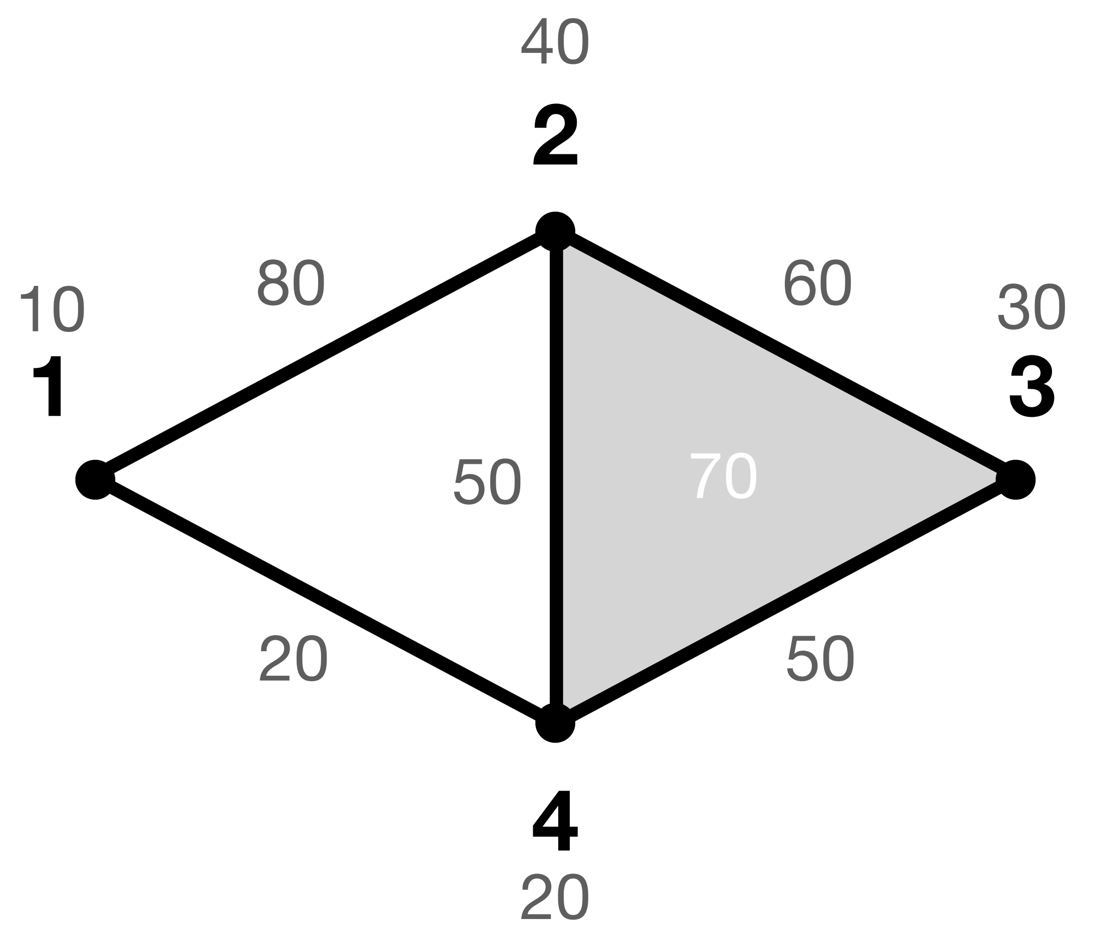
Boundary matrix reduction



								[23+24+34]		[12+24]
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			1
14										
3							1			
2						1				
24									1	
34									1	
23									1	
12										

t=80

Boundary matrix reduction

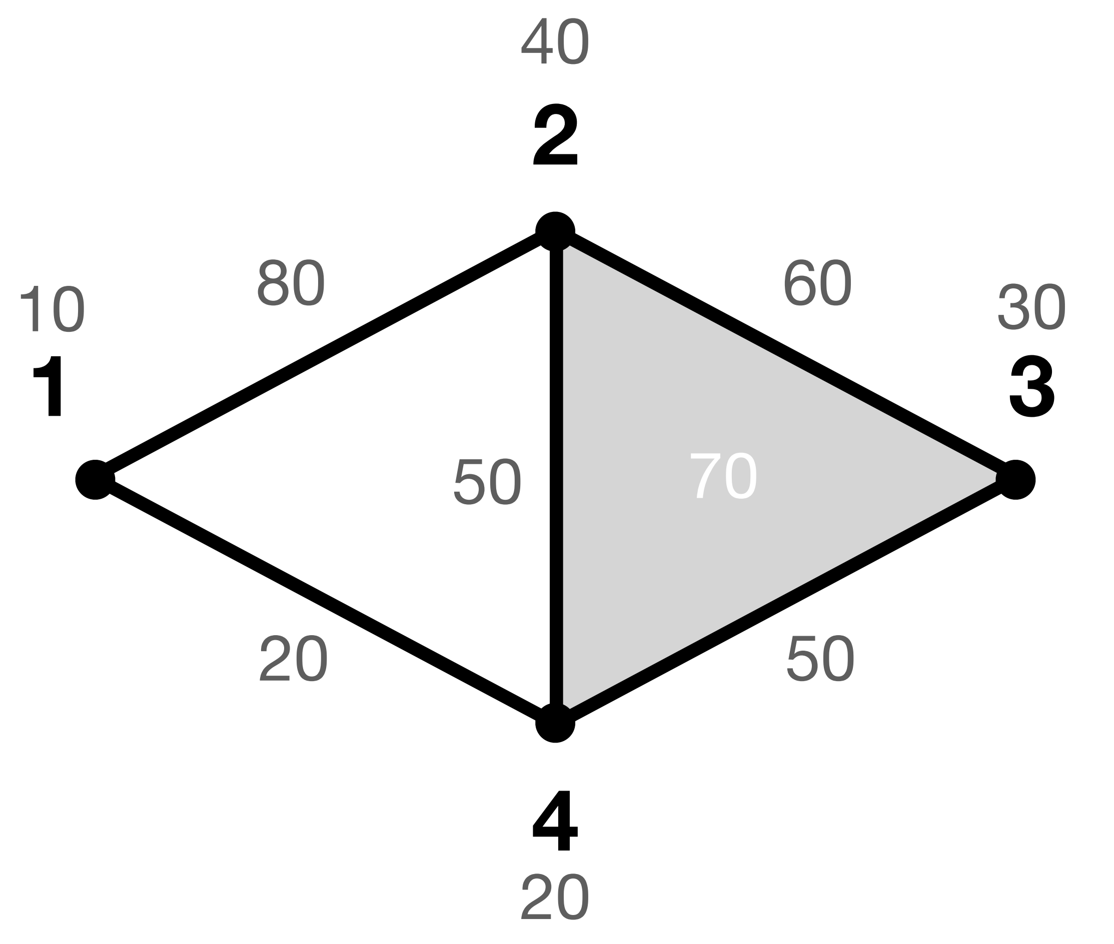


$\mathbf{R} =$

									[23+24+34]	[12+14+24]
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							
4			**1**			1	1			
14										
3							**1**			
2						**1**				
24									1	
34									1	
23									**1**	
12										

Matrix is called reduced if
all lowest nonzero elements are in unique rows

Boundary matrix reduction

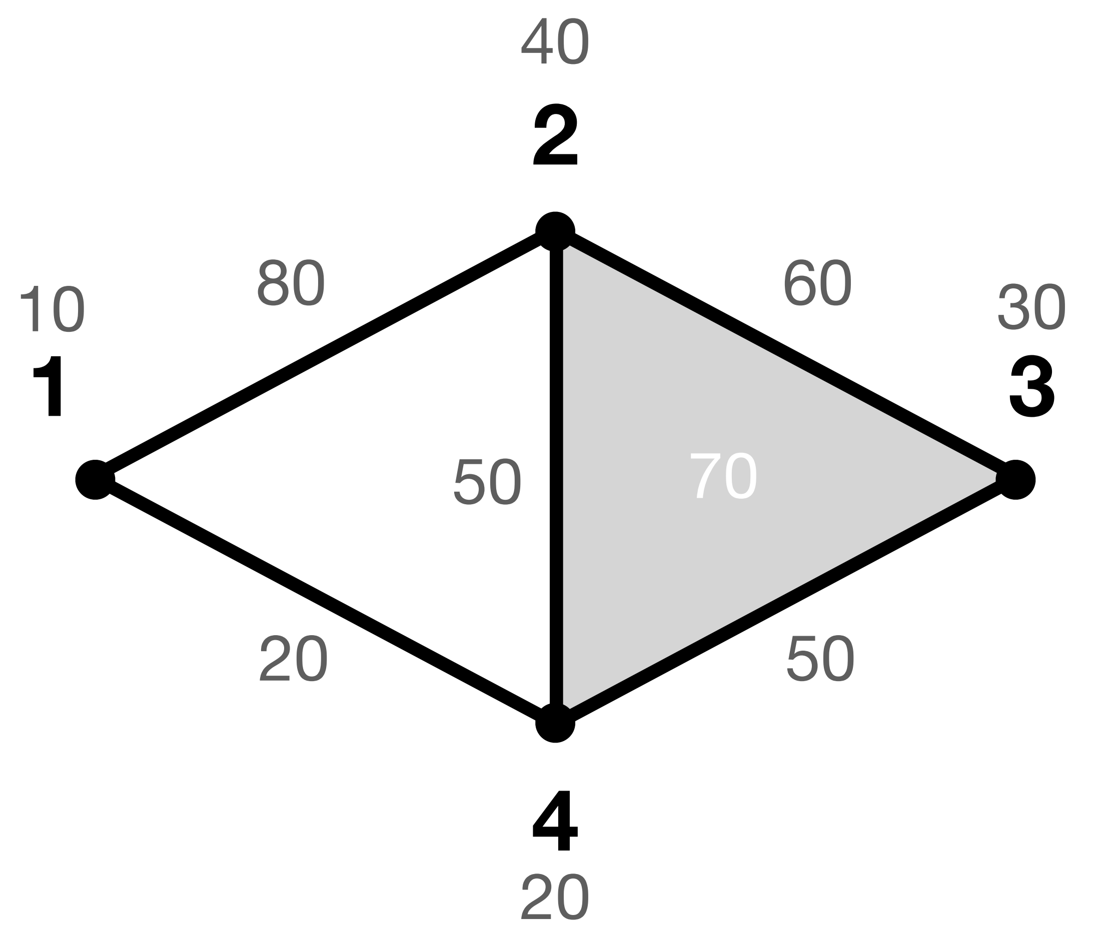


$\mathbf{R} =$

									[23+24+34]	[12+14+24]
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							
4			**1**			1	1			
14										
3							**1**			
2						**1**				
24									1	
34									1	
23									**1**	
12										

Matrix is called reduced if
all lowest nonzero elements are in unique rows

Extracting information



$\mathbf{R} =$

								[23+24+34]		[12+14+24]
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							
4			**1**			1	1			
14										
3							**1**			
2						**1**				
24									1	
34									1	
23									**1**	
12										

Essential simplices correspond to unpaired empty columns

Persistence pairing

(4, 14) 0

(2, 24) 0

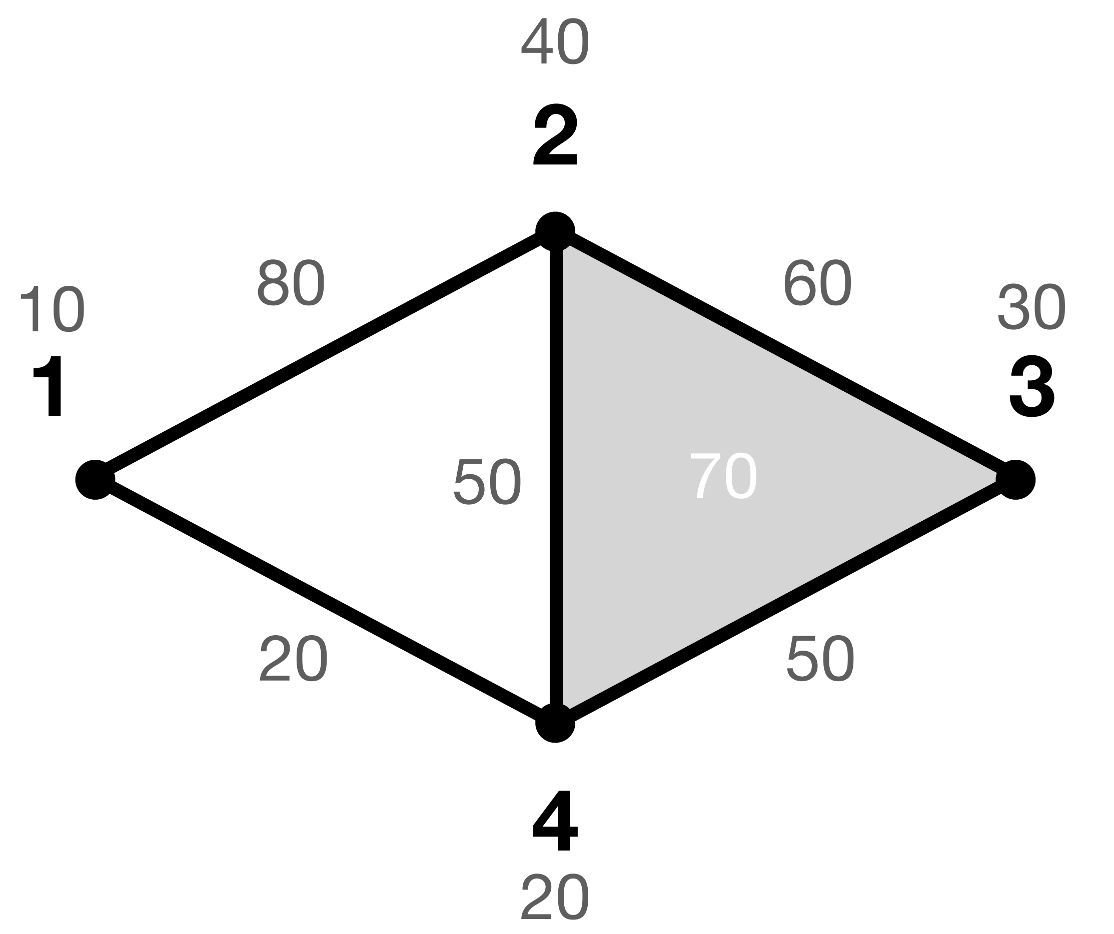
(3, 34) 0

(23, 234) 1

(1, <u>∅</u>) 0

(12, <u>∅</u>) 1

Extracting information



$\mathbf{R} =$

								[23+24+34]		[12+14+24]
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							
4			**1**			1	1			
14										
3							**1**			
2						**1**				
24									1	
34									1	
23									**1**	
12										

Persistence pairing

(4, 14) 0
(2, 24) 0
(3, 34) 0
(23, 234) 1

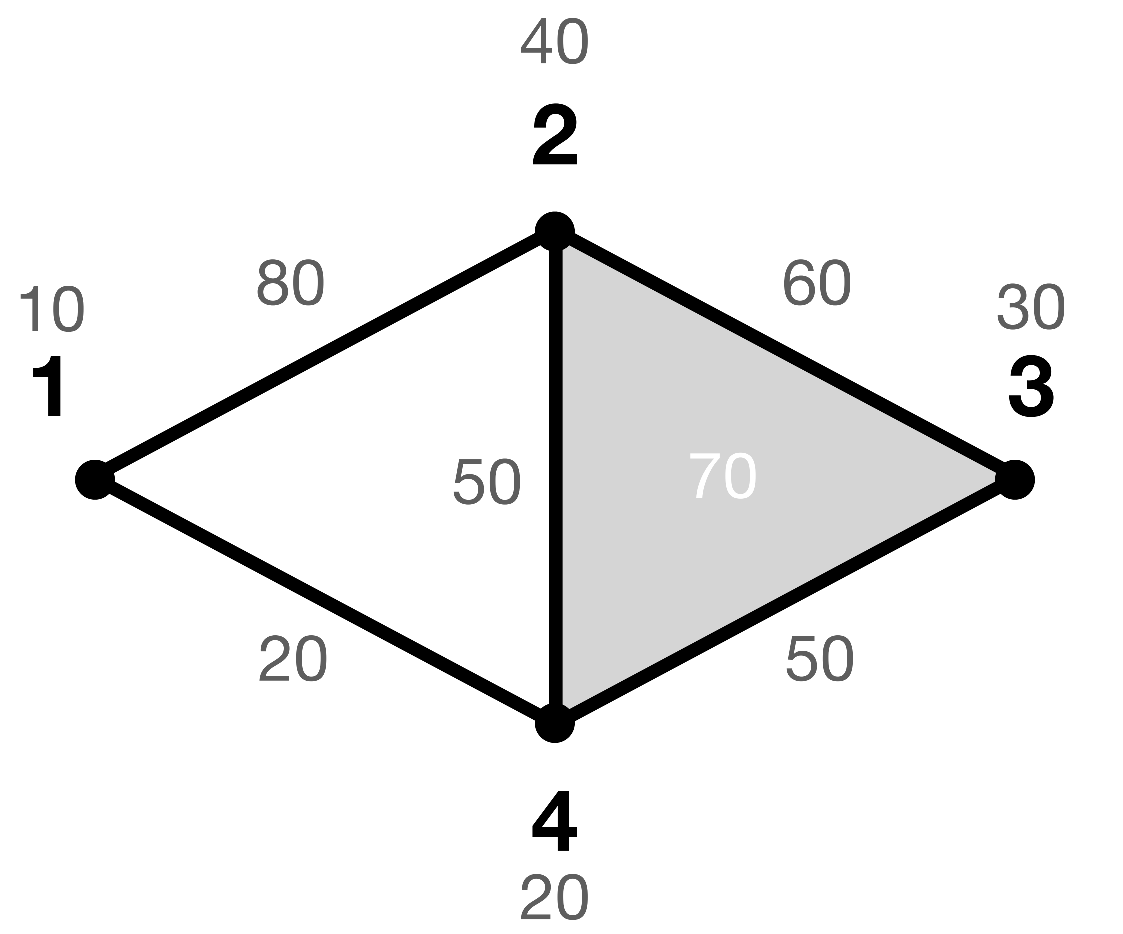
(1, Ø) 0
(12, Ø) 1

Persistence diagram

(20, 20) 0
(40, 50) 0
(30, 50) 0
(60, 70) 1

(10, Ø) 0
(80, Ø) 1

Extracting information



[23+24+34] [12+14+24]

$\mathbf{R} =$

	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							
4			1			1	1			
14										
3							1			
2						1				
24									1	
34									1	
23									1	
12										

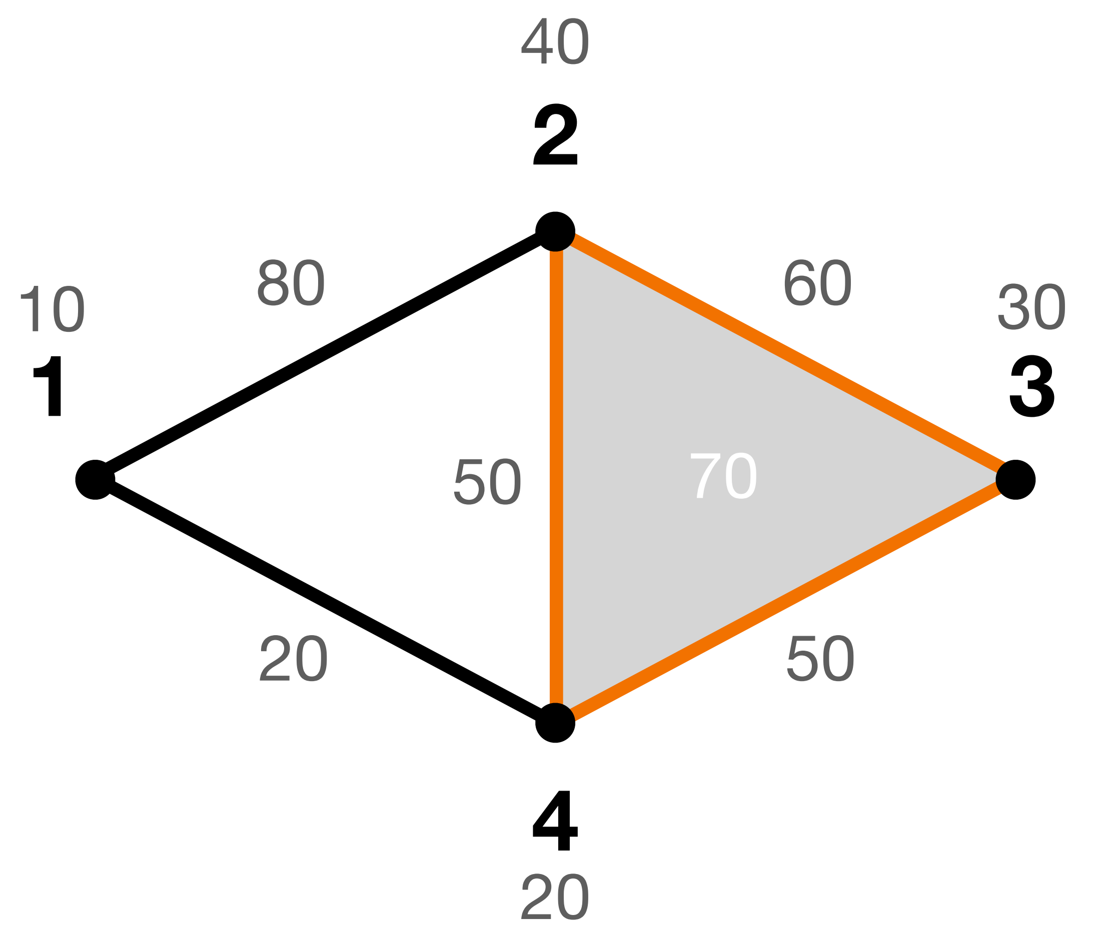
Persistence pairing [representatives]

(4, 14) 0 [4]
(2, 24) 0 [2]
(3, 34) 0 [3]
(23, 234) 1 [23+24+34]
(1, Ø) 0 [1]
(12, Ø) 1 [12+14+24]

Persistence diagram

(20, 20) 0
(40, 50) 0
(30, 50) 0
(60, 70) 1
(10, Ø) 0
(80, Ø) 1

Representatives

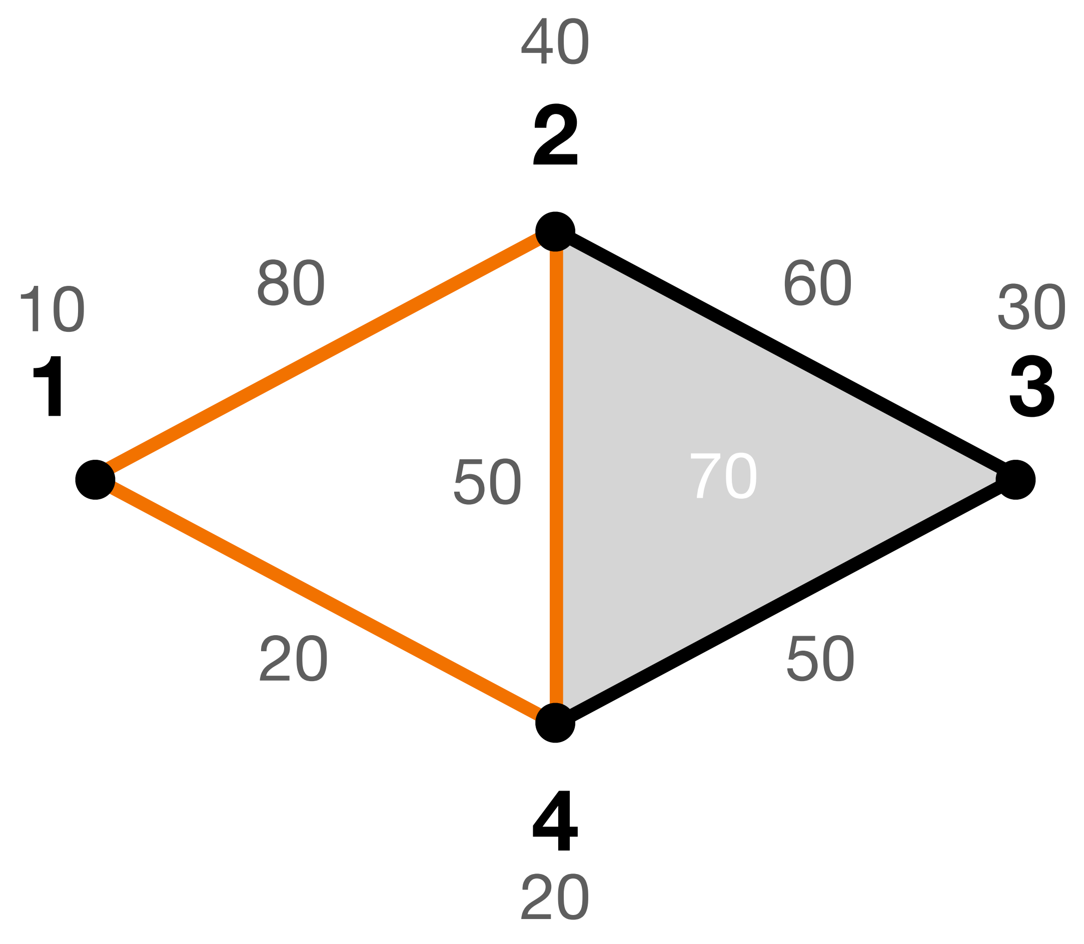


$\mathbf{R} =$

								[23+24+34]		[12+14+24]
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							
4			1			1	1			
14										
3							1			
2						1				
24									1	
34									1	
23									1	
12										

Representatives are given by the linear combination of columns corresponding to the reduced columns

Representatives



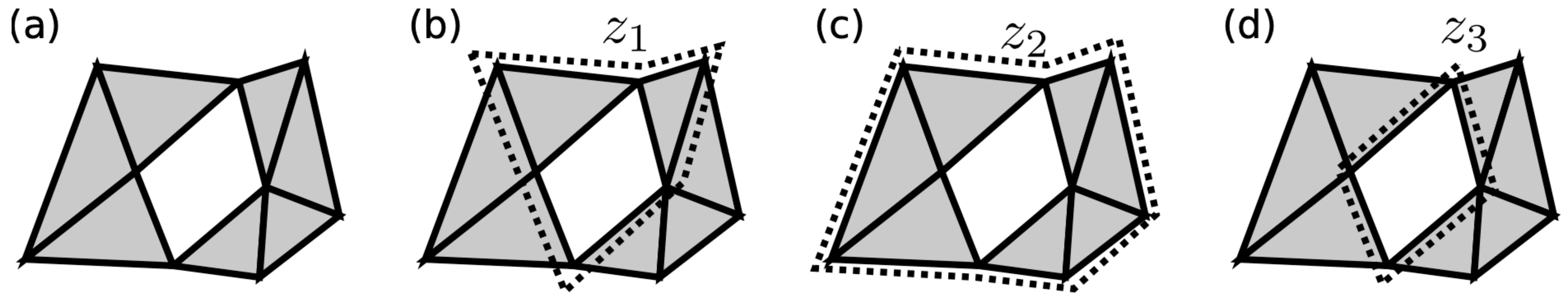
$\mathbf{R} =$

								[23+24+34]		[12+14+24]
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							
4			**1**			1	1			
14										
3							**1**			
2						**1**				
24									1	
34									1	
23									**1**	
12										

Representatives are given by the linear combination of columns corresponding to the reduced columns

Representatives

Optimal representatives



Optimal cycle

$$\text{minimize } \|z\|_0 \text{ subject to } z \sim z_1.$$

$$\text{minimize } \|z\|_0 \text{ subject to}$$
$$z = z_1 + \partial w,$$
$$w \in C_2(X).$$

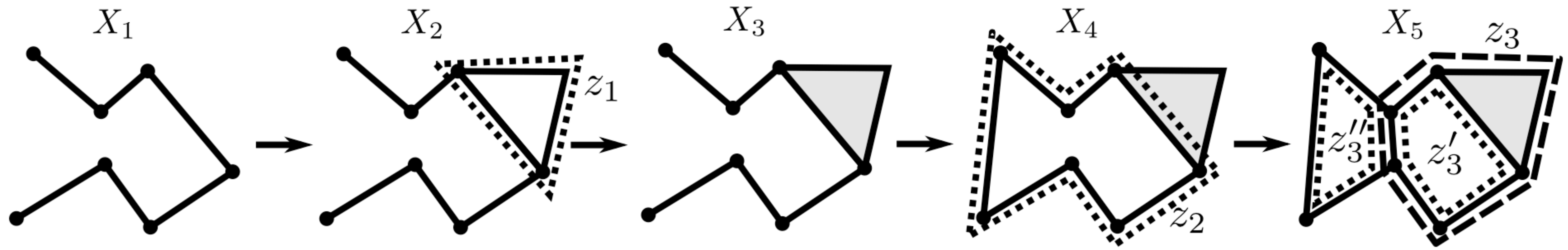
$$z = z_0 + a_1 \partial w_1 + a_2 \partial w_2 + \ldots a_n \partial w_n$$

Minimal Cycle Representatives in Persistent Homology using Linear Programming
Li et al. *Frontiers in Artificial Intelligence (2021)*

Volume-optimal Cycle: Tightest Representative Cycle of a Generator in Persistent Homology
Obayashi I. *SIAM Journal on Applied Algebra and Geometry (2018)*

Representatives

Optimal representatives w.r.t. a filtration — cycles



$$\begin{aligned} \text{minimize } \|z\|_0 \;\; &\text{subject to} \\ z &= z_3 + \partial w + k z_2, \\ w &\in C_2(X_5), \\ k &\in \Bbbk. \end{aligned}$$

Algorithm 1 Computing an optimal cycle on a filtration.

Compute $D_q(\mathbb{X})$ and persistence cycles $z_1, \ldots, z_n$
Have $(b_i, d_i) \in D_q(\mathbb{X})$ be chosen by a user
Solve the following optimization problem:

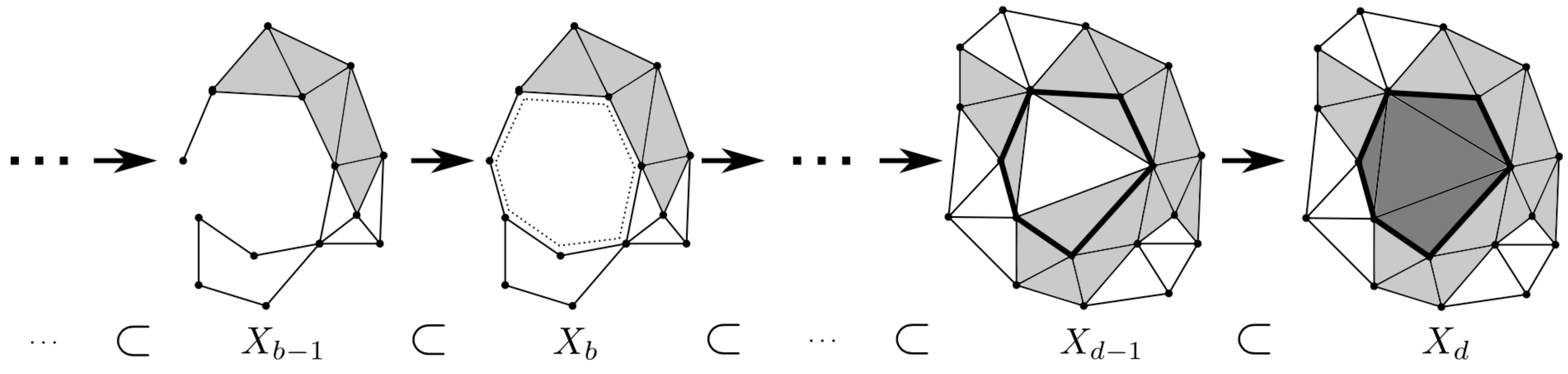
$$\begin{aligned} \text{minimize } \|z\|_1 \;\; &\text{subject to} \\ z &= z_i + \partial w + \sum_{j \in T_i} \alpha_j z_j, \\ w &\in C_{q+1}(X_{b_i}), \\ \alpha_j &\in \Bbbk, \\ \text{where } T_i &= \{j \mid b_j < b_i < d_j\} \end{aligned}$$

Minimal Cycle Representatives in Persistent Homology using Linear Programming
Li et al. *Frontiers in Artificial Intelligence (2021)*

Volume-optimal Cycle: Tightest Representative Cycle of a Generator in Persistent Homology
Obayashi I. *SIAM Journal on Applied Algebra and Geometry (2018)*

Representatives

Optimal representatives w.r.t. a filtration — volumes



Algorithm 2 Algorithm for a volume-optimal cycle.

procedure VOLUME-OPTIMAL-CYCLE($\mathbb{X}, r$)

Compute the persistence diagram $D_q(\mathbb{X})$

Have a user choose a birth-death pair $(b_i, d_i) \in D_q(\mathbb{X})$

Solve the following optimization problem:

$$\begin{aligned} &\text{minimize } \|z\|_1 \text{ subject to} \\ &z = \sigma_{d_i} + \sum_{\sigma_k \in \mathcal{F}_{q+1}^{(r)}} \alpha_k \sigma_k, \\ &\tau^*(\partial z) = 0 \text{ for all } \tau \in \mathcal{F}_q^{(r)} \end{aligned} \tag{14}$$

Minimal Cycle Representatives in Persistent Homology using Linear Programming
Li et al. *Frontiers in Artificial Intelligence (2021)*

Volume-optimal Cycle: Tightest Representative Cycle of a Generator in Persistent Homology
Obayashi I. *SIAM Journal on Applied Algebra and Geometry (2018)*

Representatives

Scaffolds

Given a graph *G* a homological scaffold *H(G)* is a subgraph of G induced by the edges present in the representatives of homology classes.

The *frequency homological scaffold $H^F(G)$* as the network composed of all the cycle paths corresponding to generators, where an edge e is weighted by the number of different cycles it belongs to.

$$\omega_e^P = \sum_{[z]_i} \mathbf{1}_{e \in [z]_i}$$

The *persistence homological scaffold $H^P(G)$* is the network composed of all the cycle paths corresponding to generators weighted by their persistence.

$$\omega_e^F = \sum_{[z]_i \mid e \in [z]_i} \pi_{[z]_i}$$

If an edge e belongs to multiple cycles $z_0, z_1, \ldots, z_s$, its weight is defined as the sum of the generators' persistence.

Homological Scaffolds of Brain Functional Networks
Petri et al. *Journal of The Royal Society Interface (2014)*

Homological Scaffold via Minimal Homology Bases
Guerra et al. *Scientific Reports (2021)*

Harmonic representatives

Hodge Laplacian operator

$$C_2 \underset{\partial_2^*}{\overset{\partial_2}{\rightleftarrows}} C_1 \underset{\partial_1^*}{\overset{\partial_1}{\rightleftarrows}} C_0$$

$$\mathbf{L}_k = \mathbf{B}_k^T \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T$$

Hodge decomposition

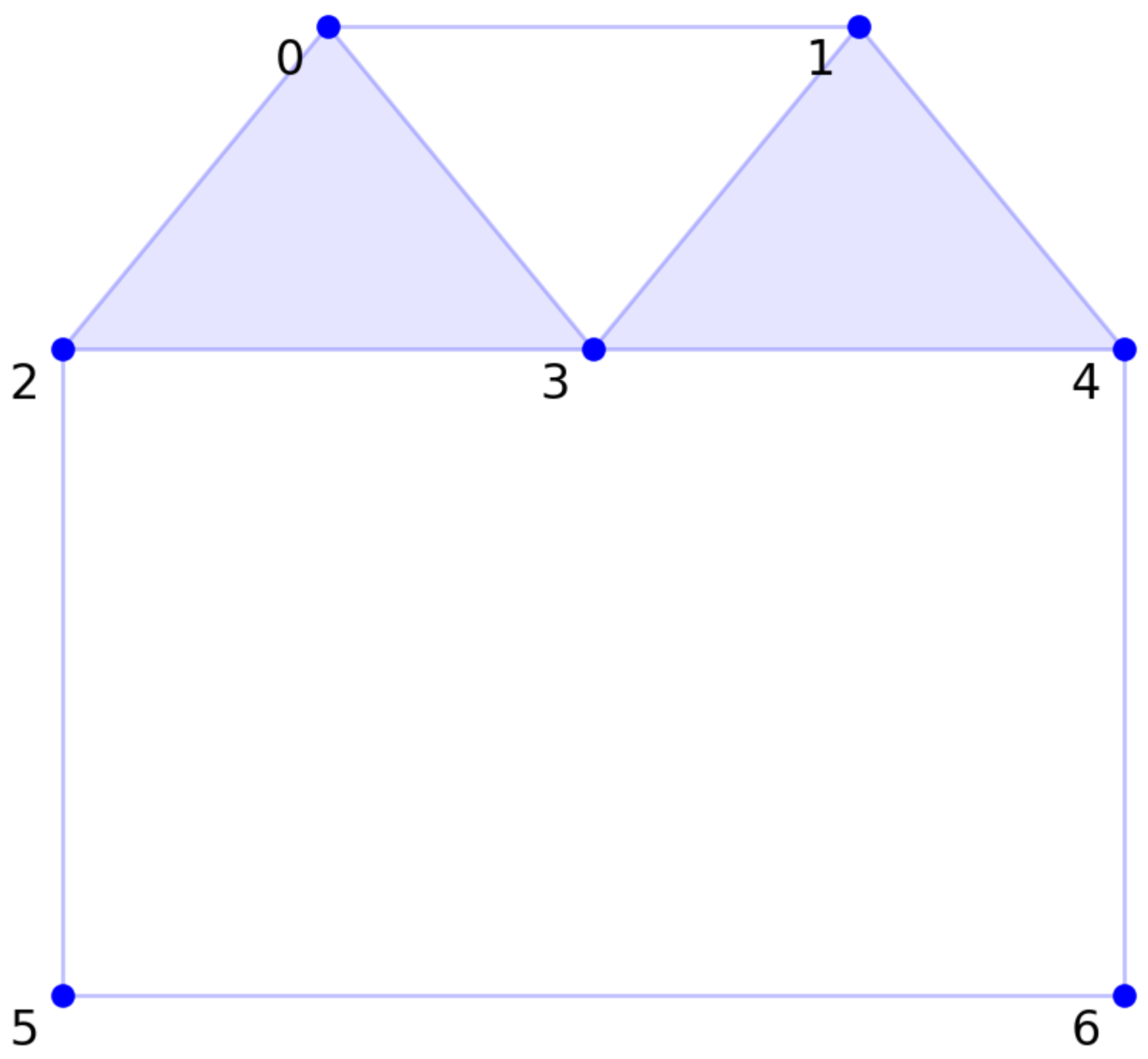
$$C_k = \mathrm{im}\partial_{k+1} \oplus \ker L_k \oplus \partial_k^T$$

Betti numbers via Hodge Laplacian

$$\beta_k = \dim \ker(\mathbf{L}_k)$$

Harmonic cycles

$$z_k^H \in \ker L_k$$



Harmonic representatives

Hodge Laplacian operator

$$C_2 \underset{\partial_2^*}{\overset{\partial_2}{\rightleftarrows}} C_1 \underset{\partial_1^*}{\overset{\partial_1}{\rightleftarrows}} C_0$$

$$\mathbf{L}_k = \mathbf{B}_k^T \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T$$

Hodge decomposition

$$C_k = \mathrm{im}\partial_{k+1} \oplus \ker L_k \oplus \partial_k^T$$

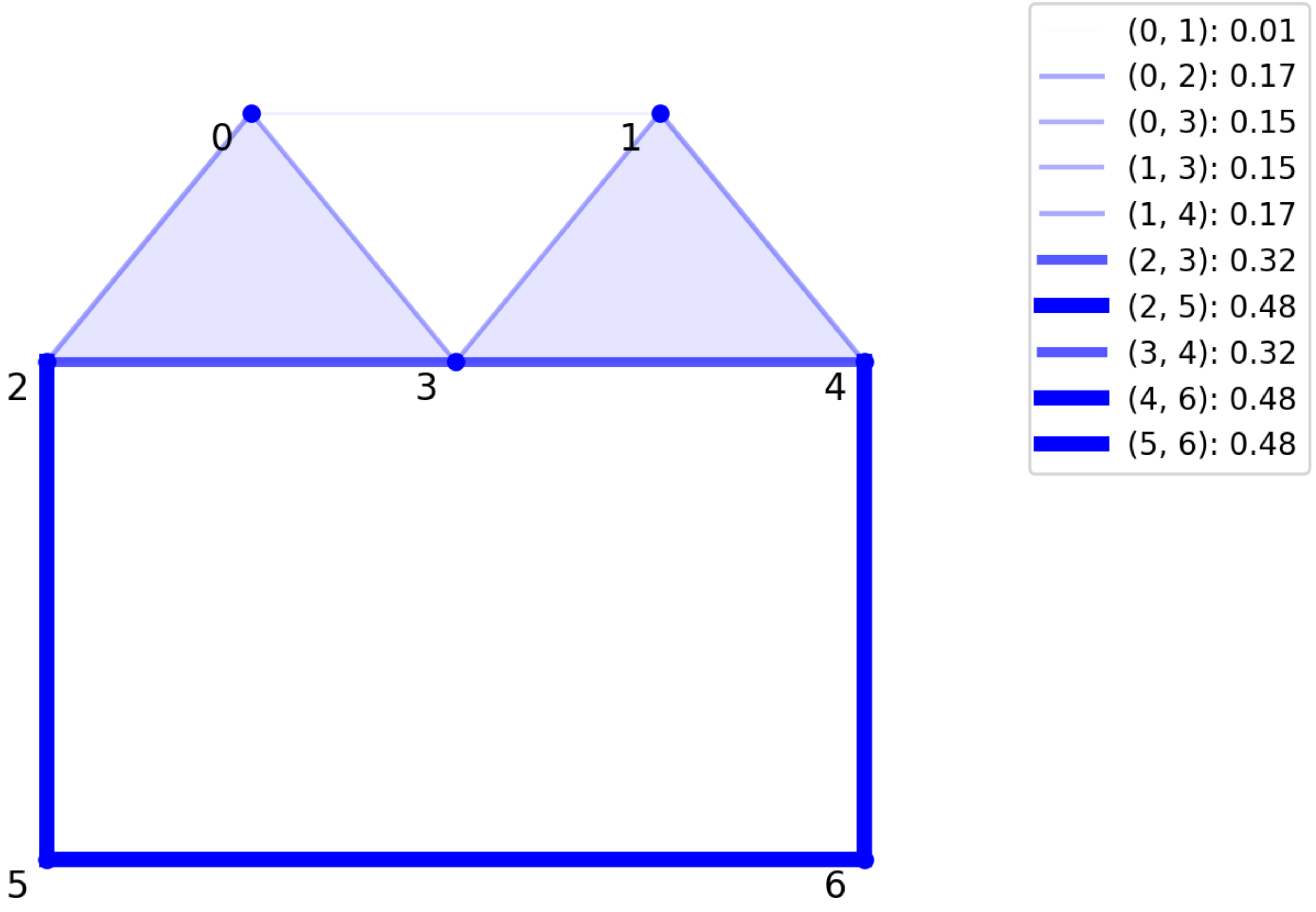
Betti numbers via Hodge Laplacian

$$\beta_k = \dim \ker(\mathbf{L}_k)$$

Harmonic cycles

$$z_k^H \in \ker L_k$$

Eigenvectors corresponding to zero eigenvalues of $\mathbf{L}_k$



Harmonic representatives

Hodge Laplacian operator

$$C_2 \underset{\partial_2^*}{\overset{\partial_2}{\rightleftarrows}} C_1 \underset{\partial_1^*}{\overset{\partial_1}{\rightleftarrows}} C_0$$

$$\mathbf{L}_k = \mathbf{B}_k^T \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T$$

Hodge decomposition

$$C_k = \mathrm{im}\partial_{k+1} \oplus \ker L_k \oplus \partial_k^T$$

Betti numbers via Hodge Laplacian

$$\beta_k = \dim \ker(\mathbf{L}_k)$$

Harmonic cycles

$$z_k^H \in \ker L_k$$

Eigenvectors corresponding to zero eigenvalues of $\mathbf{L}_k$