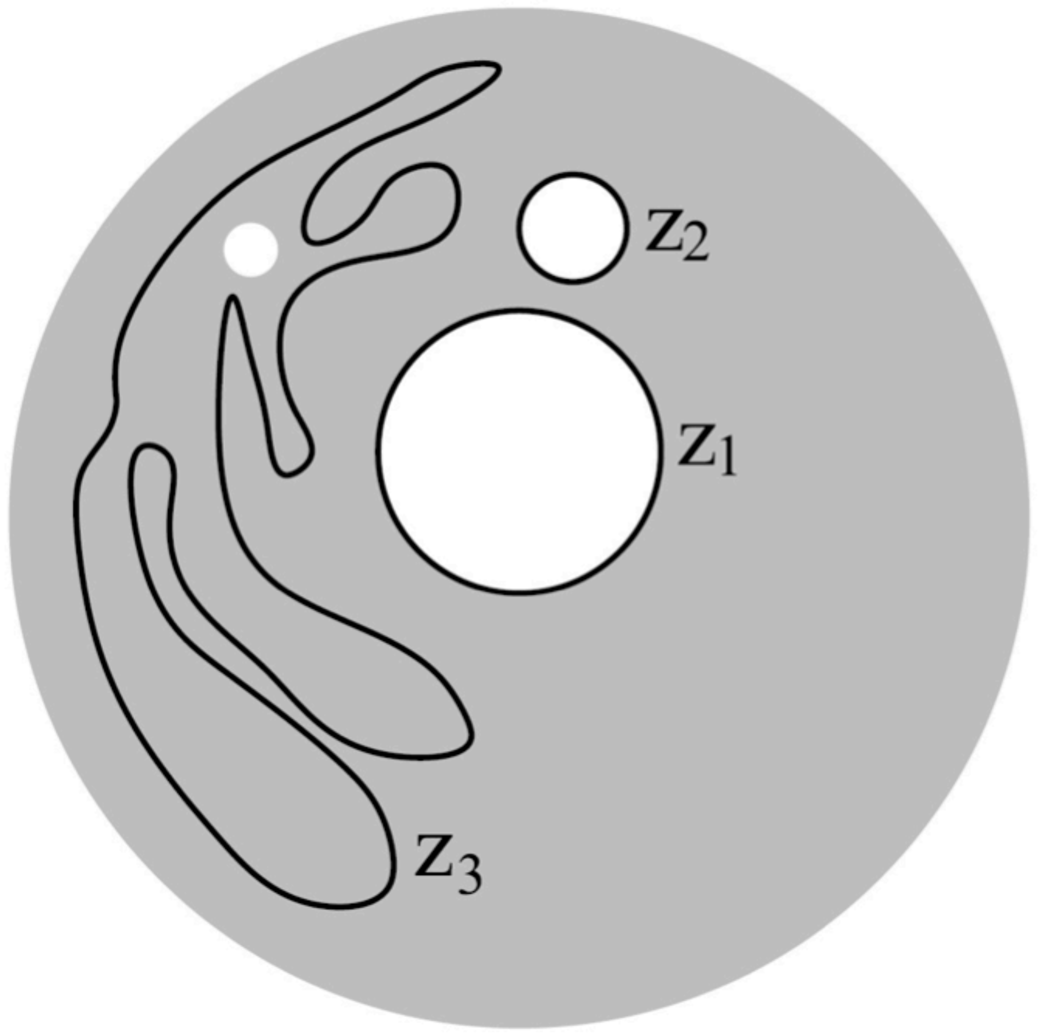
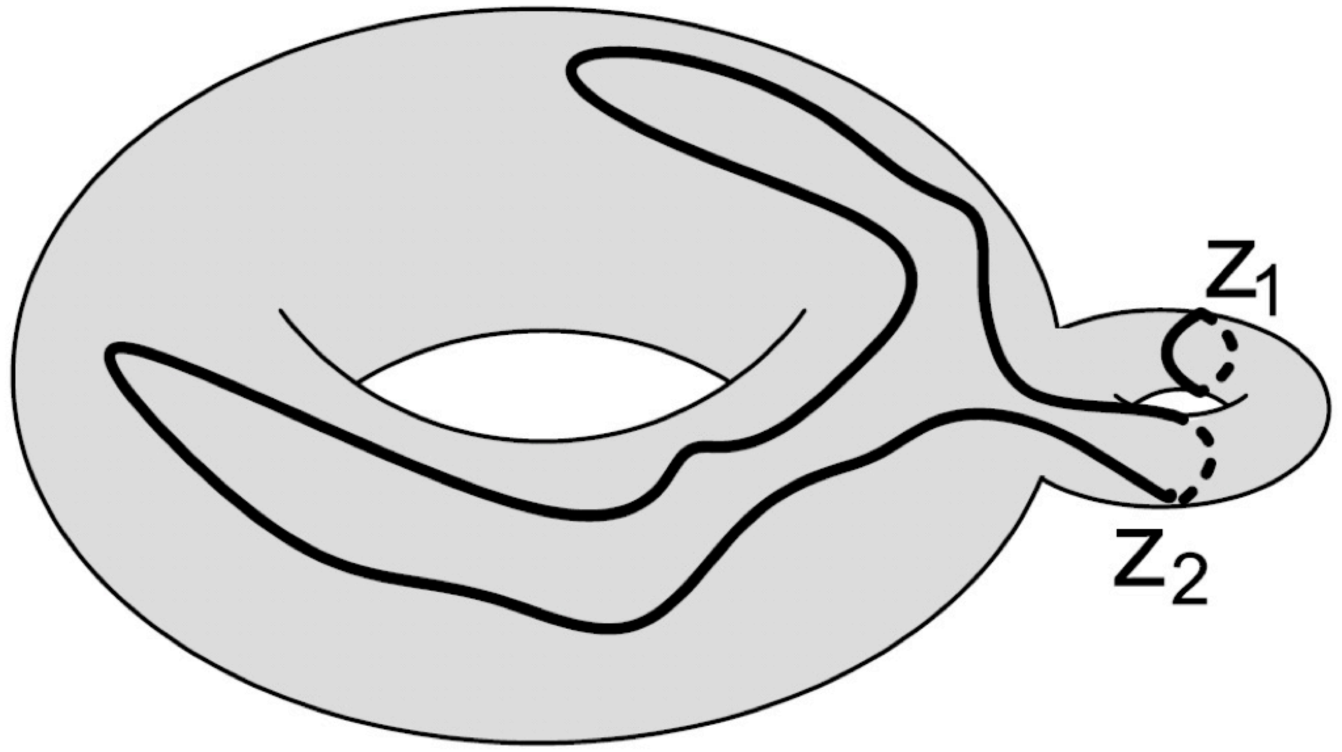
Topological Data Analysis

Lecture 11
Homology representatives

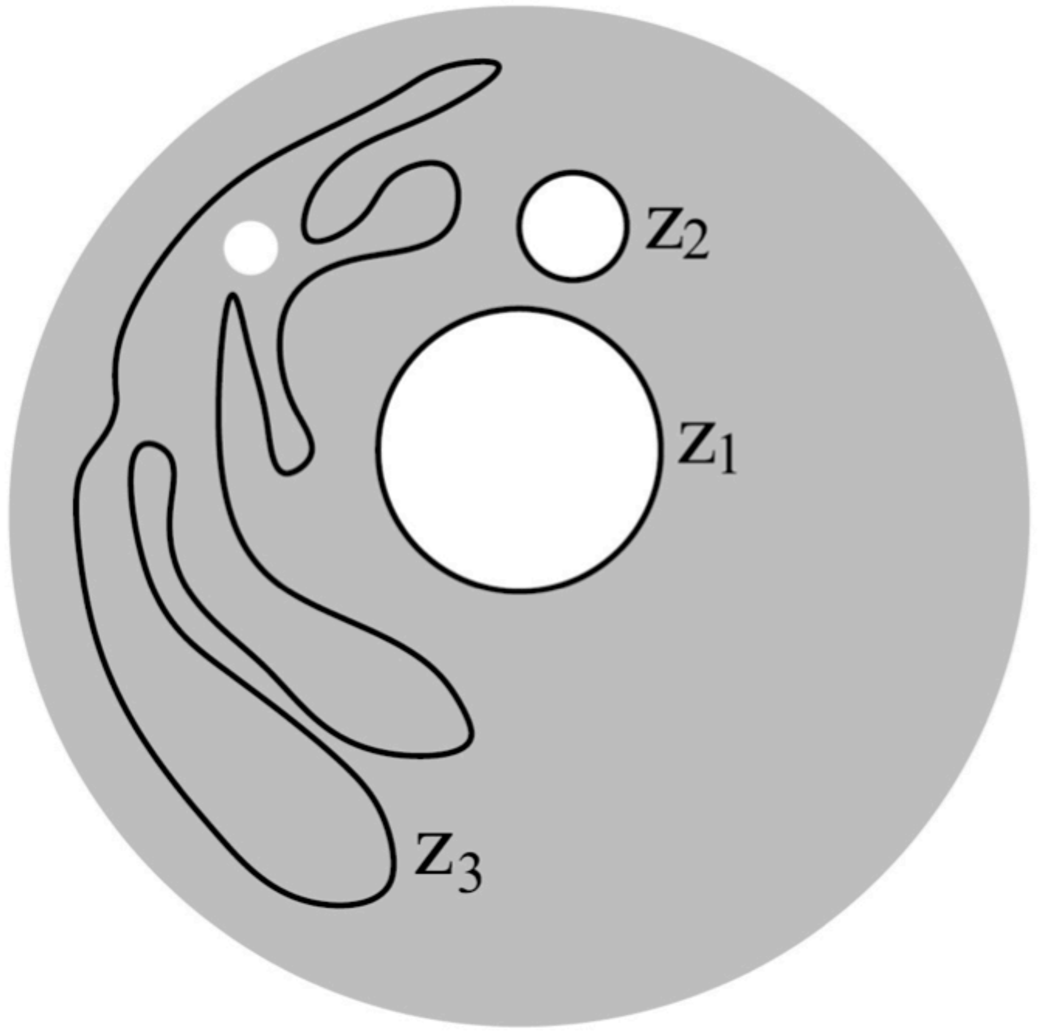
Oleg Kachan

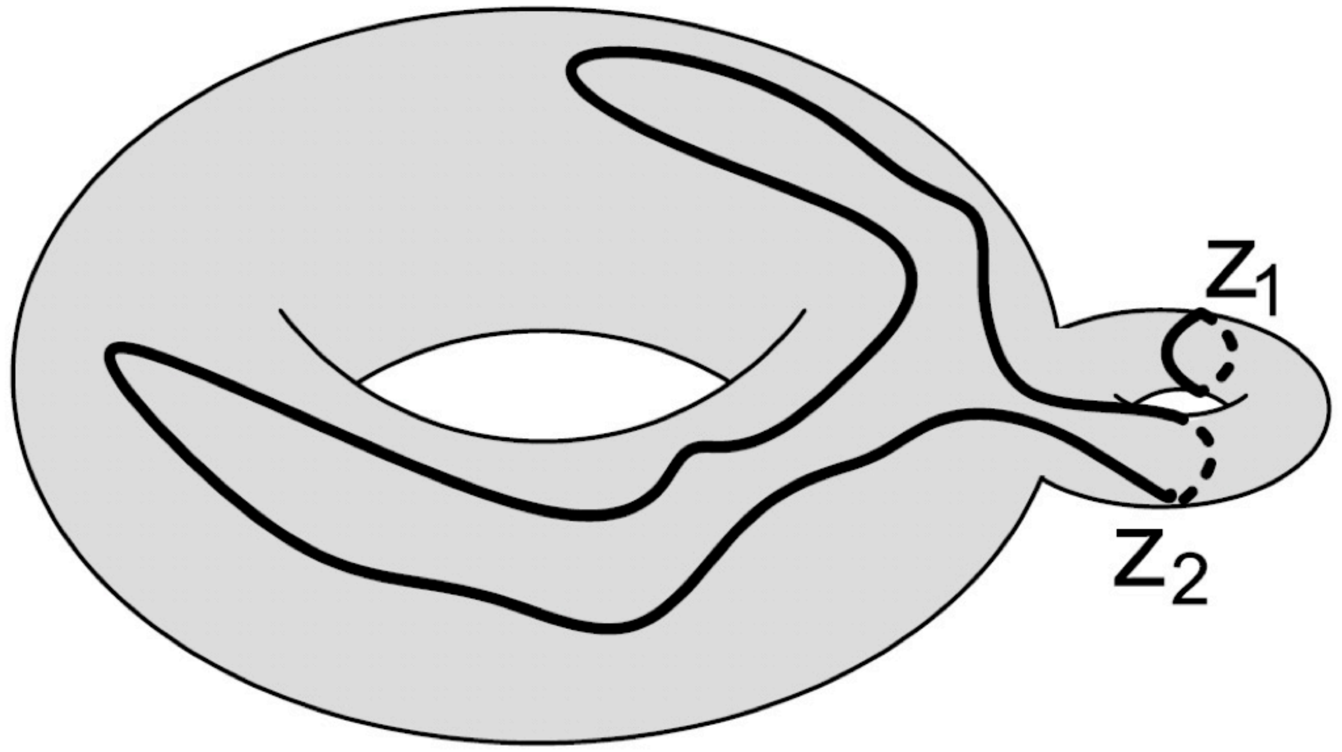
Homology representatives





Homology representatives

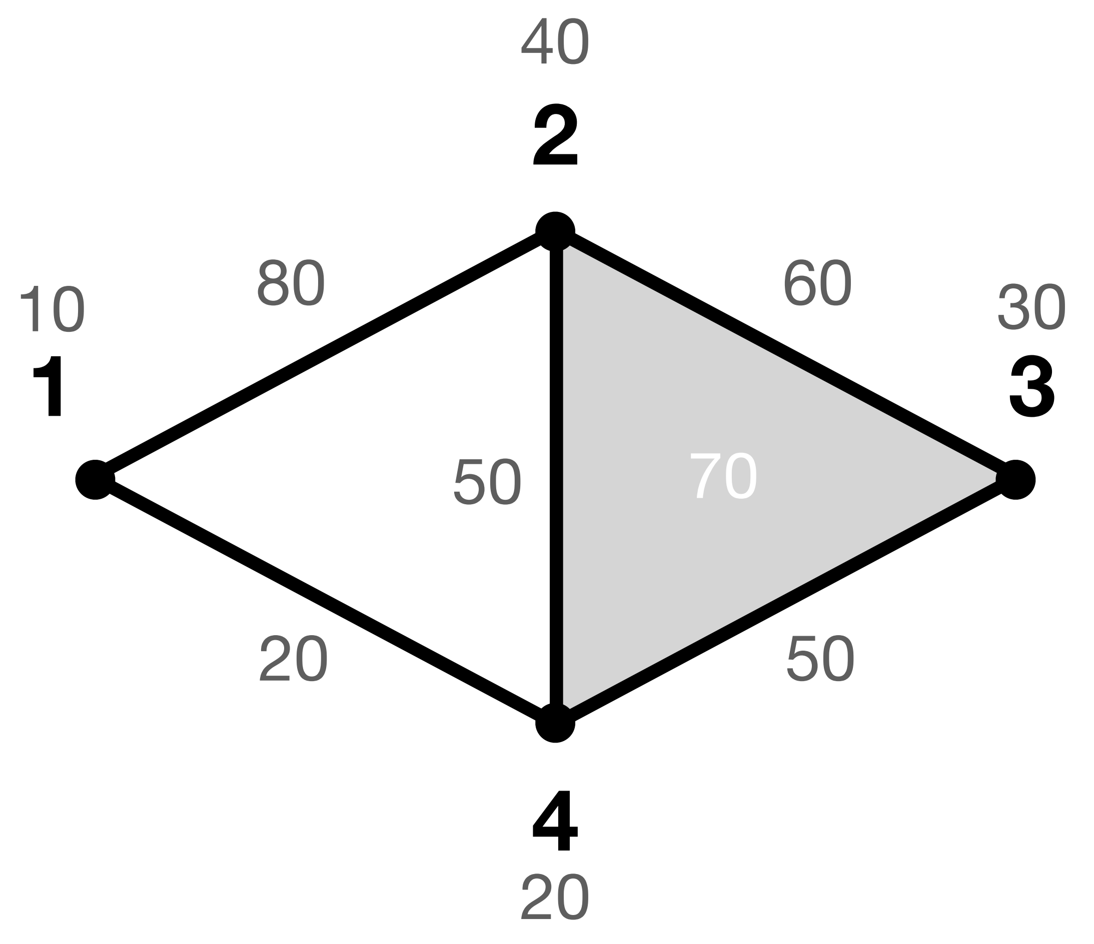




Optimal cycle

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

Boundary matrix reduction



$\mathbf{B} =$

	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										

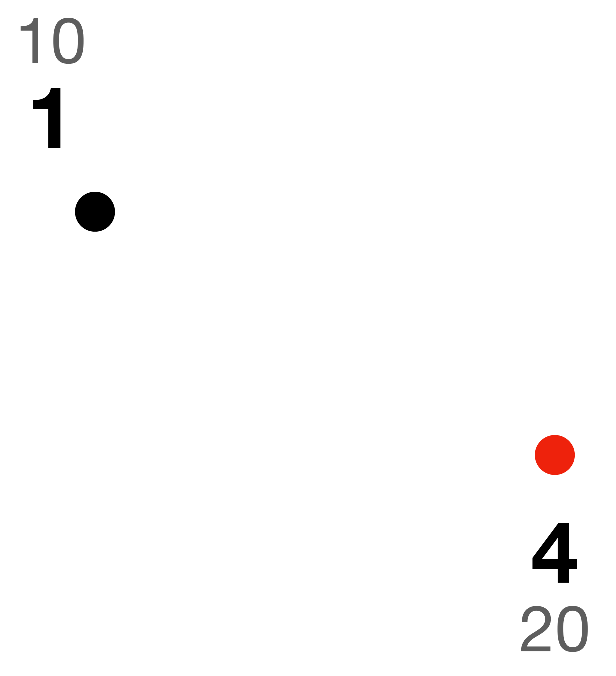
Filtration function provides order on simplices, therefore on columns and rows of the filtration matrix. Ties are broken, first by simplex dimension, second by lexicographic order given by order on vertices.

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

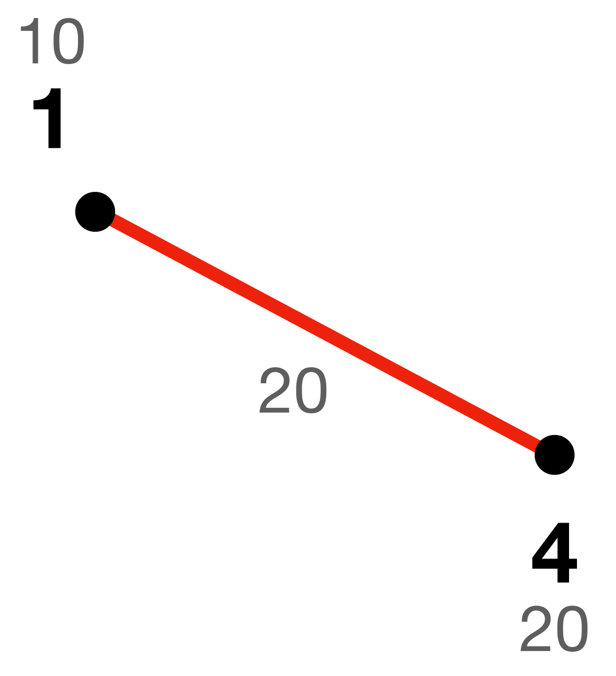
Boundary matrix reduction



	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

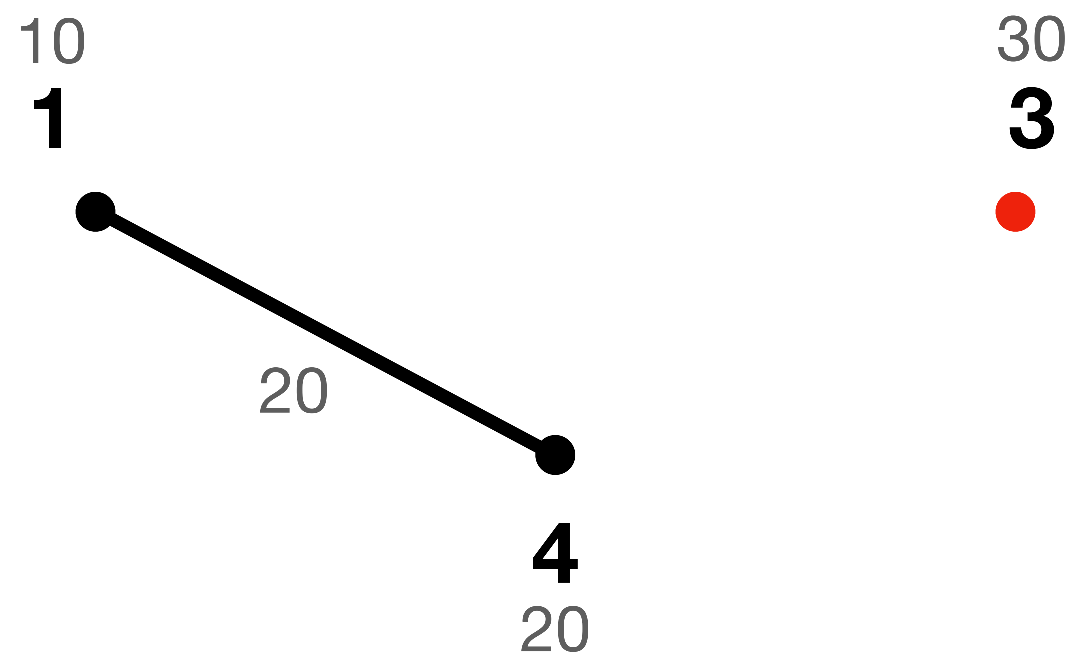
Boundary matrix reduction



	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

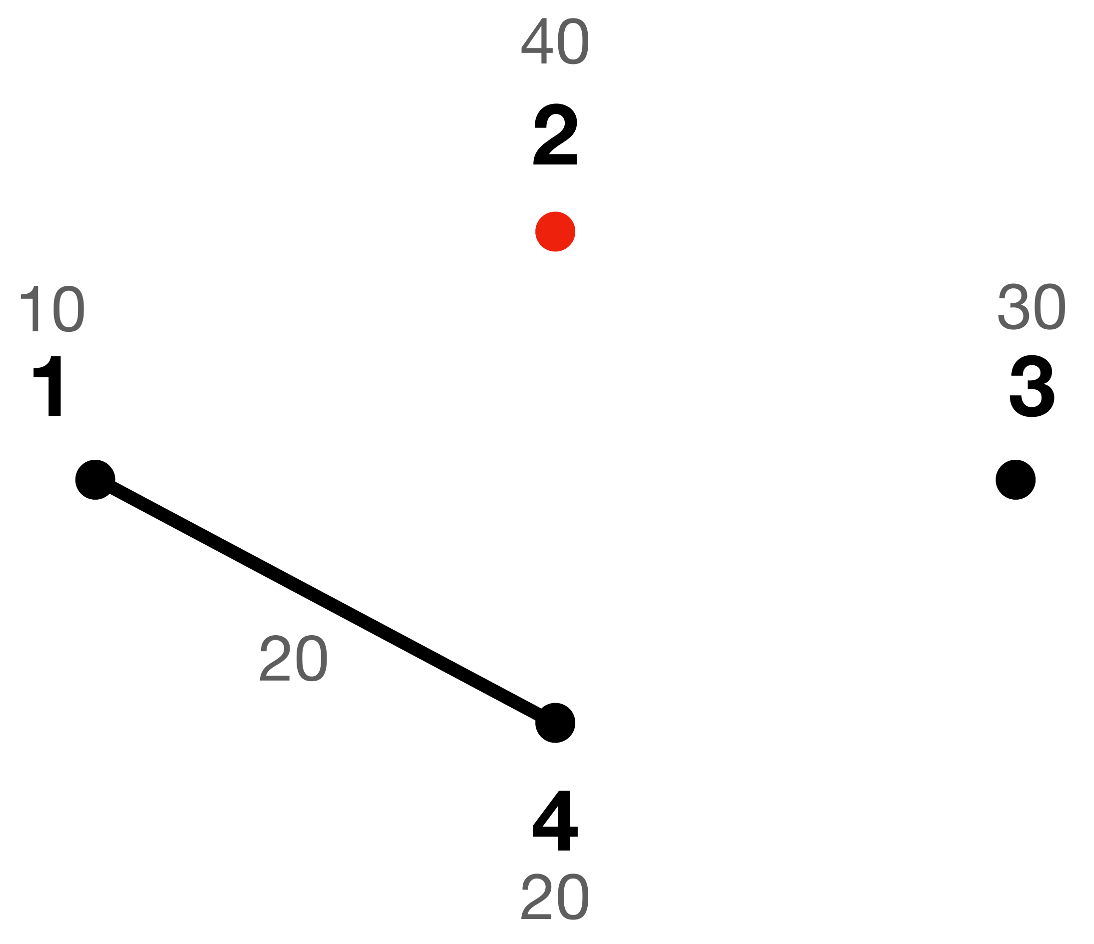
Boundary matrix reduction



	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

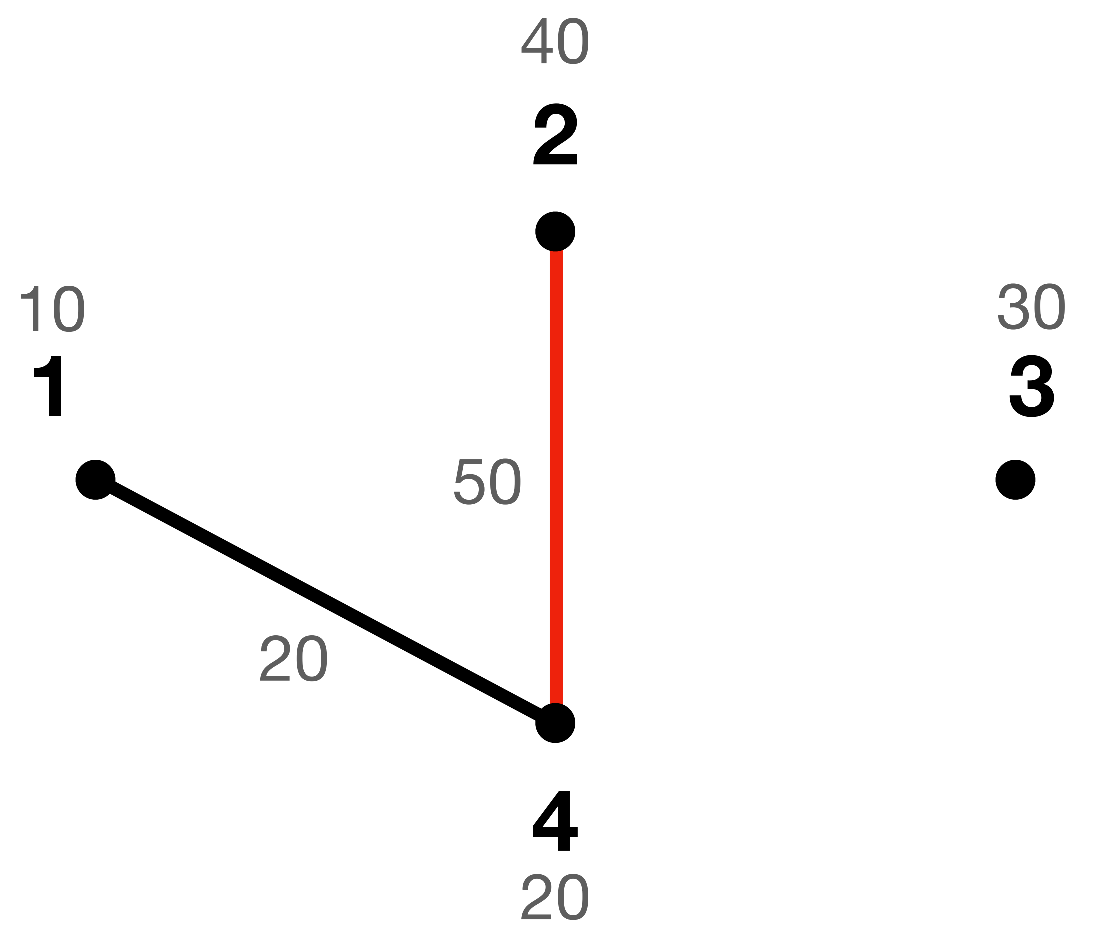
Boundary matrix reduction



	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

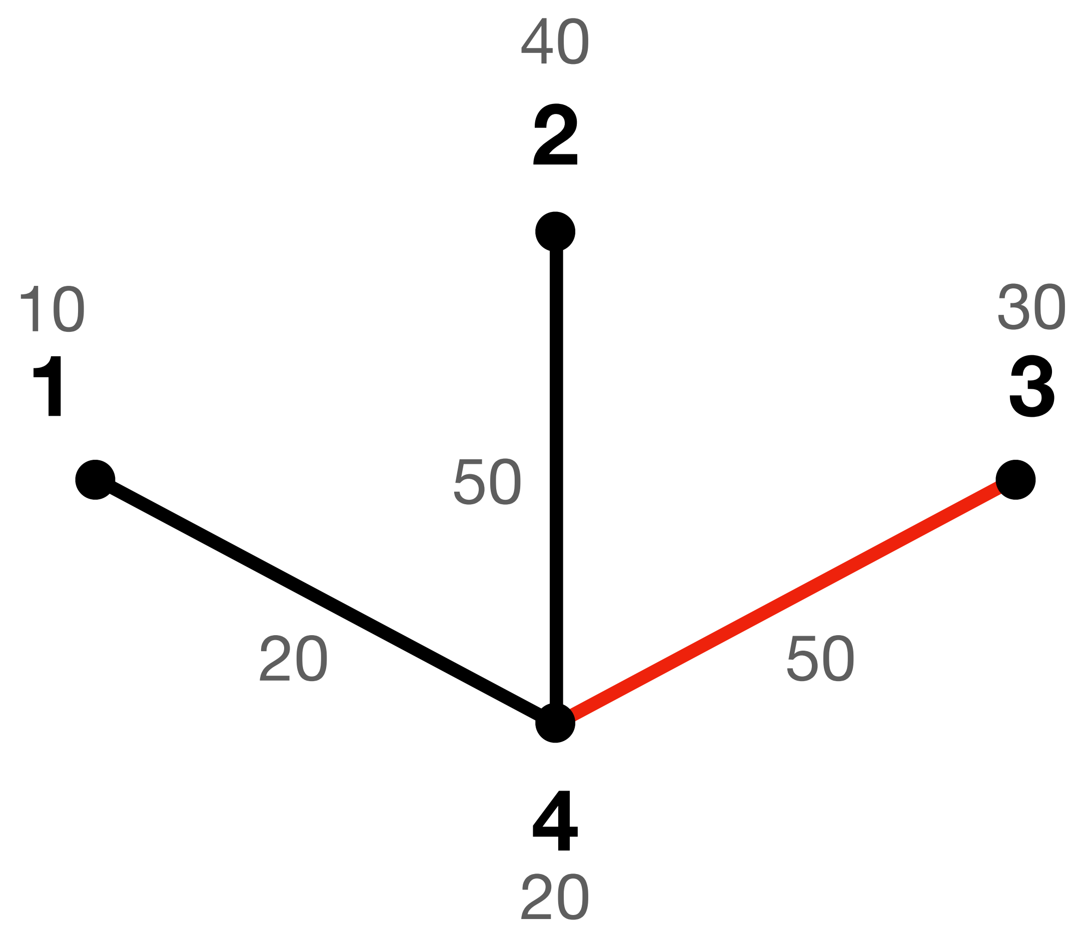
Boundary matrix reduction



	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

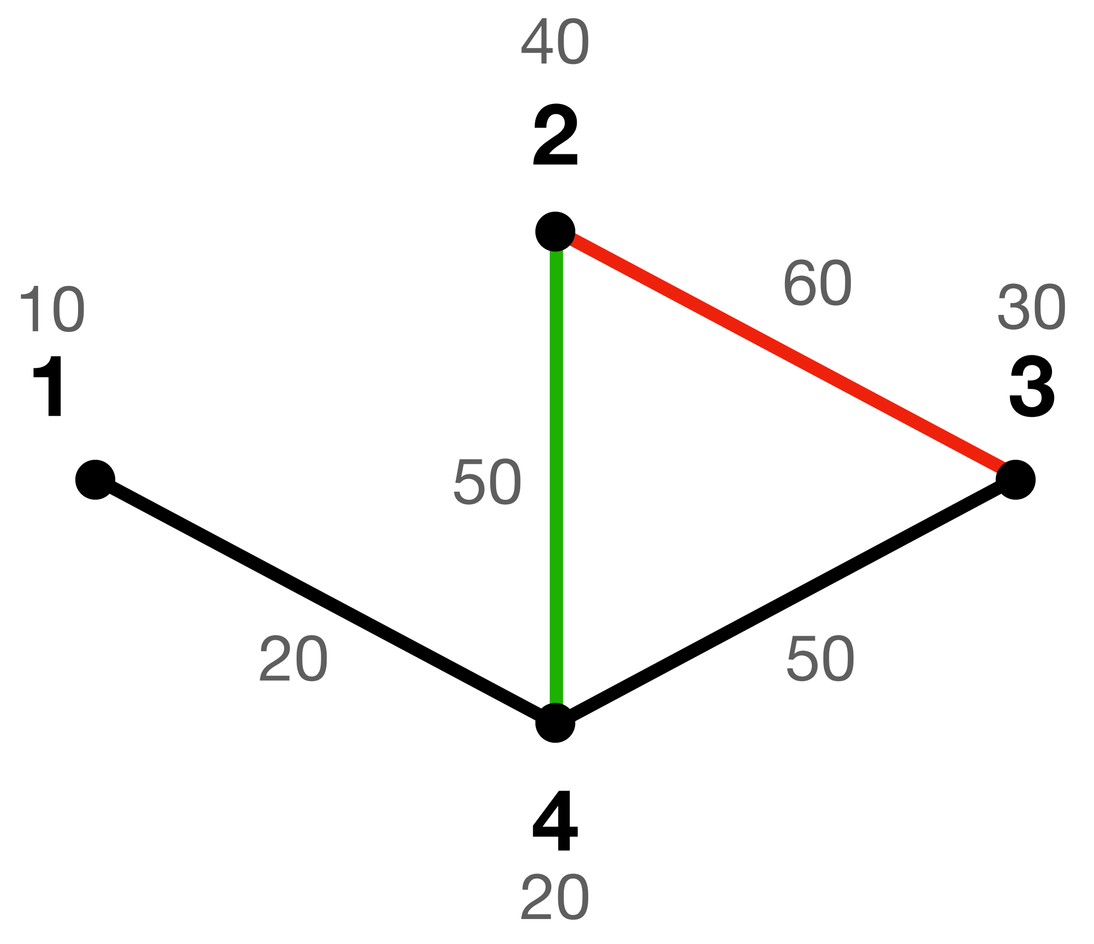
Boundary matrix reduction



	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

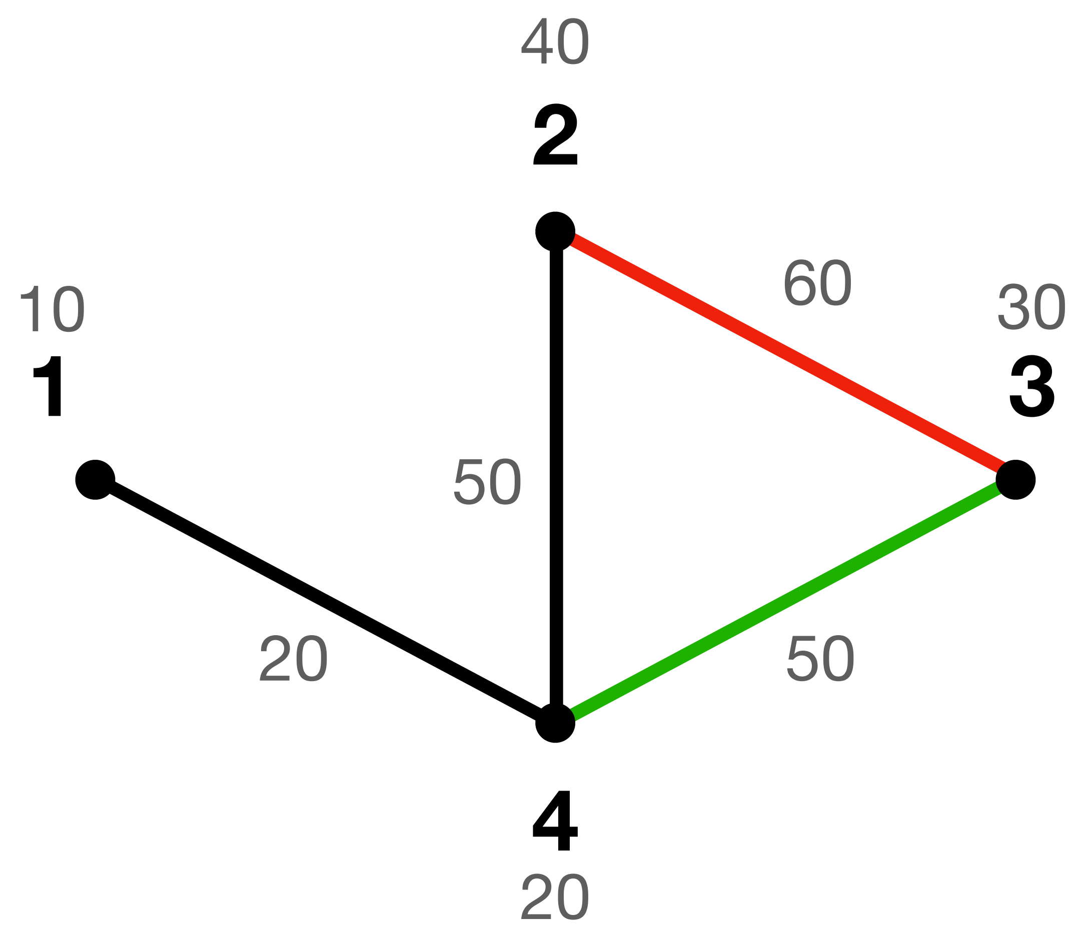
Boundary matrix reduction



	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

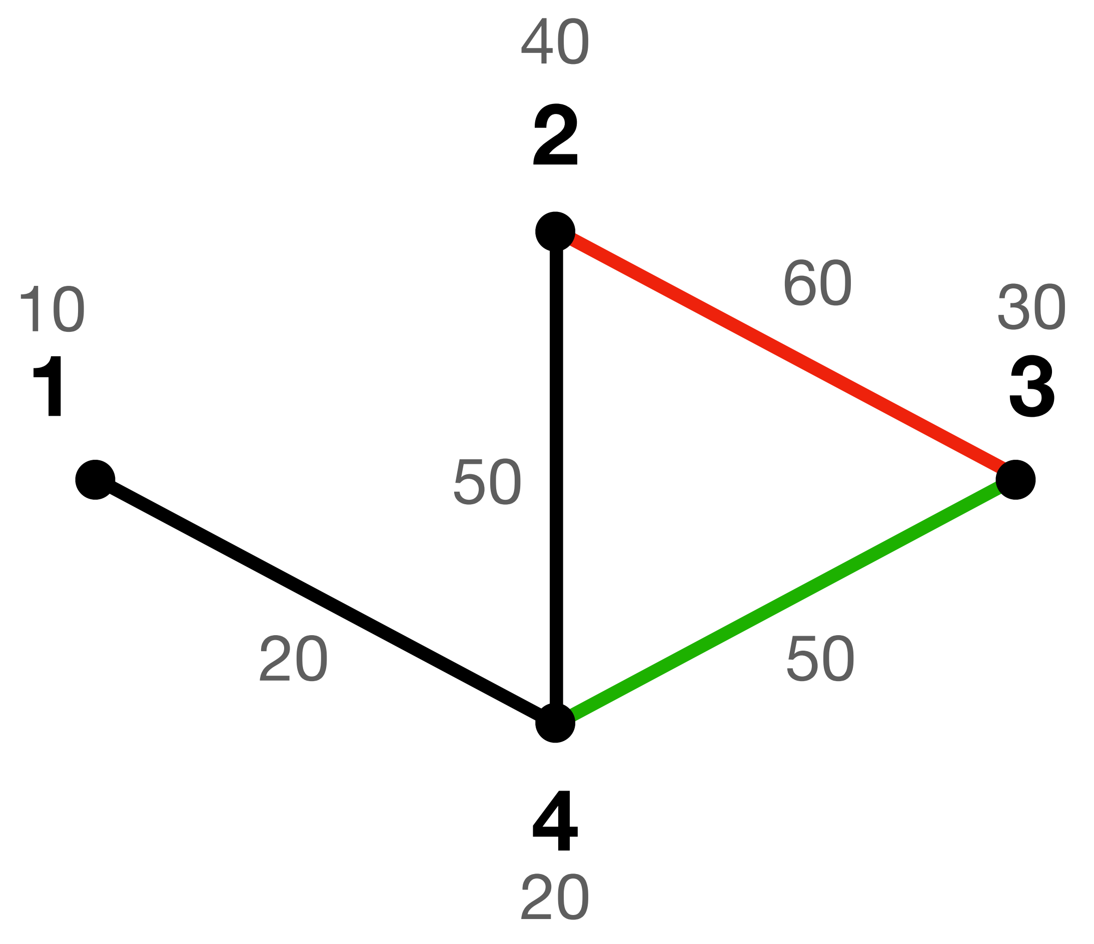
Boundary matrix reduction



								[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

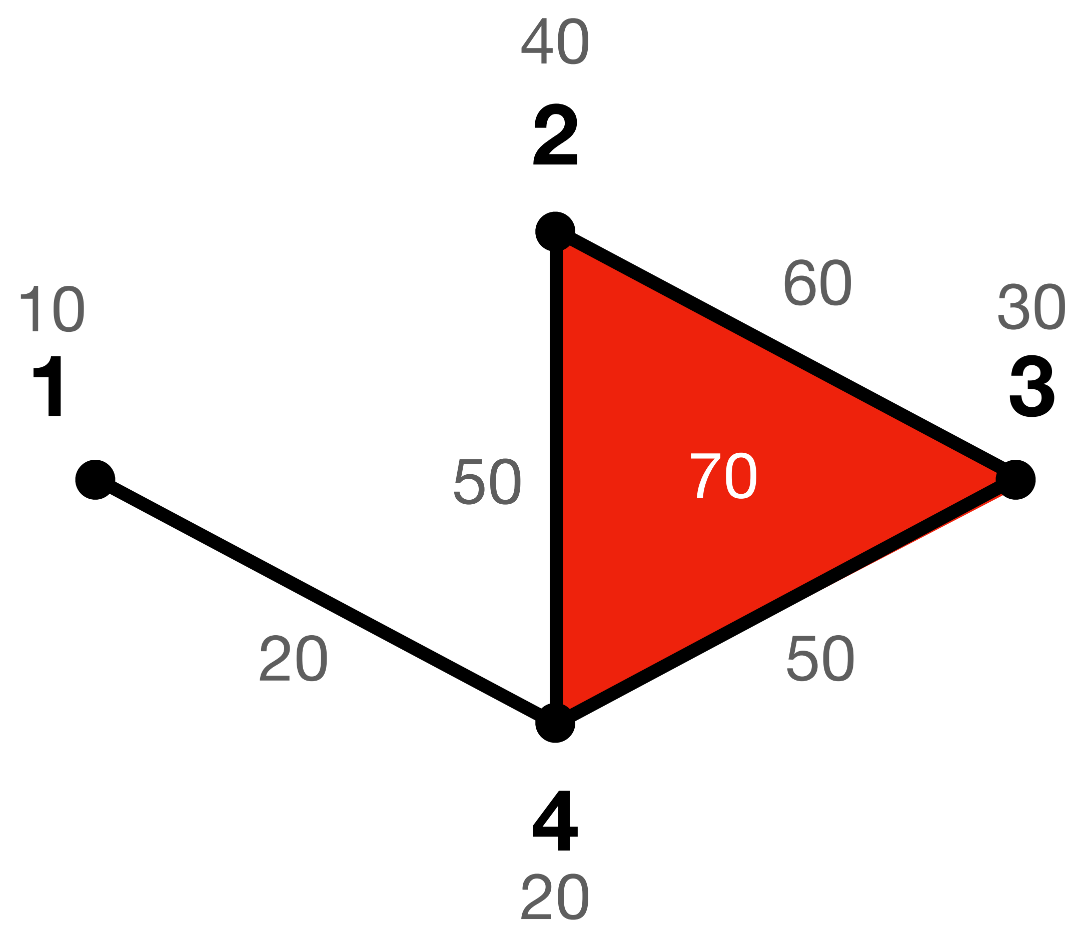


[23+24+34]

	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

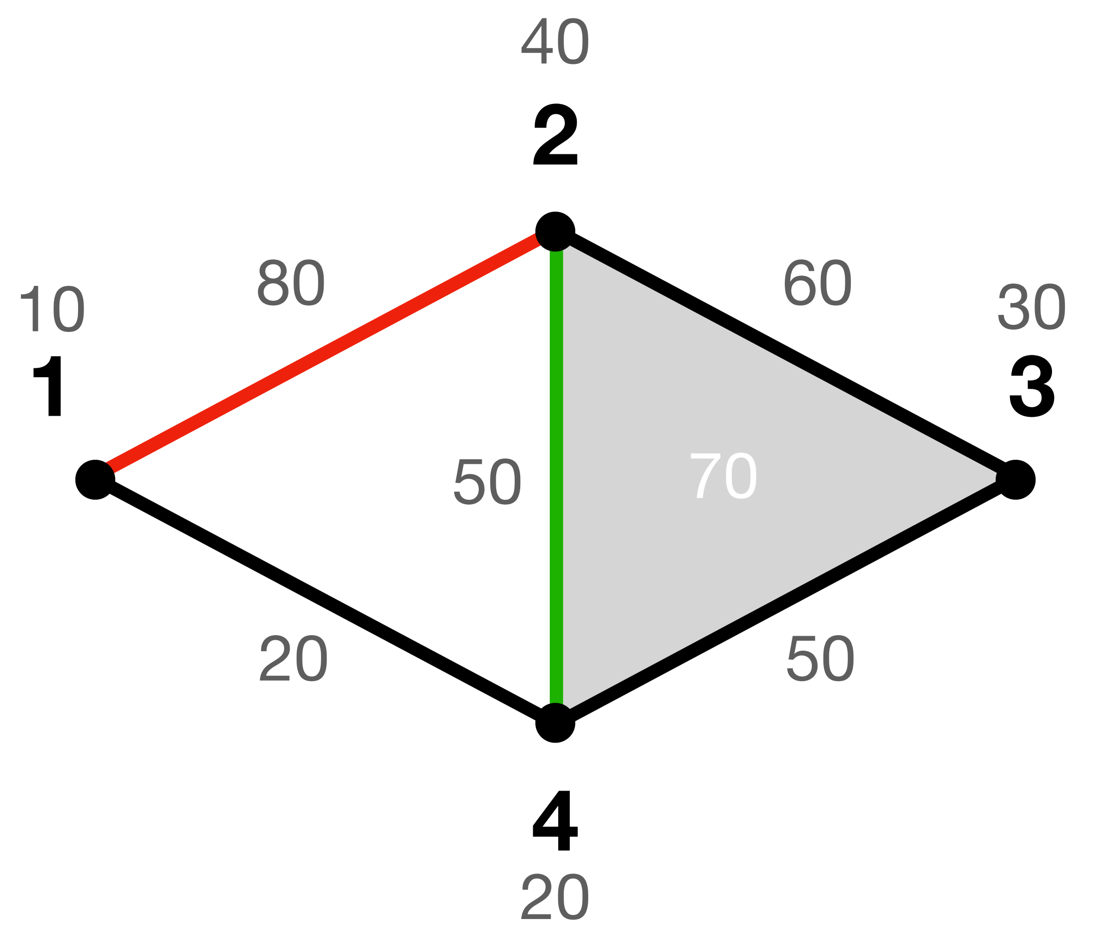


[23+24+34]

	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

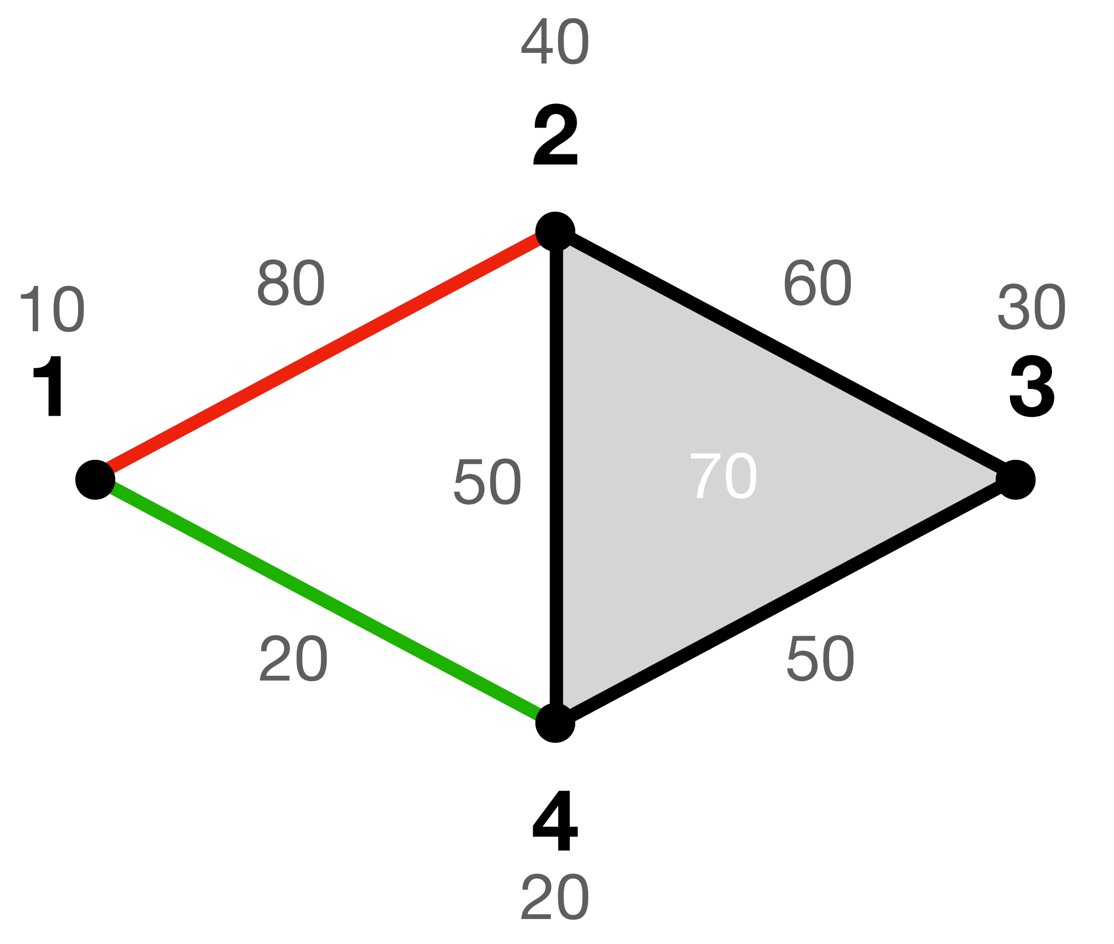
Boundary matrix reduction



								[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

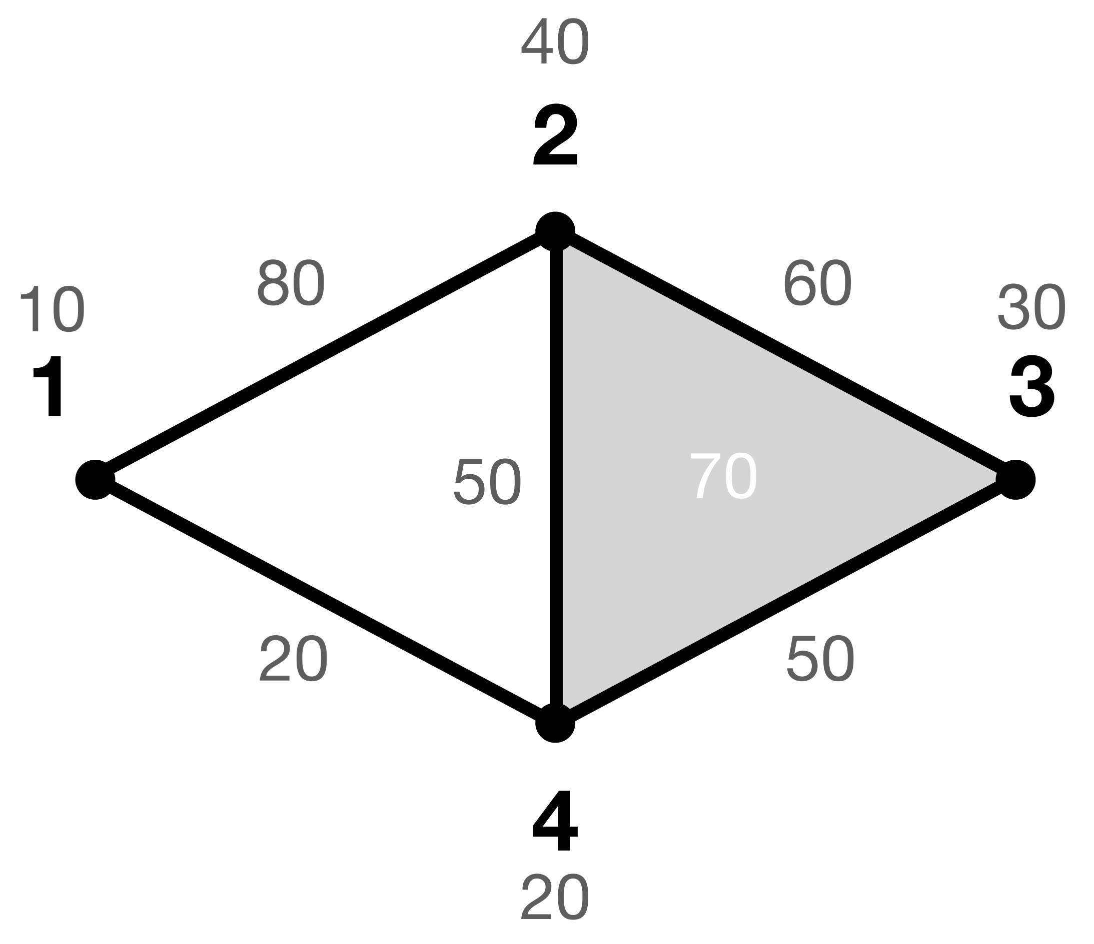
Boundary matrix reduction



								[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

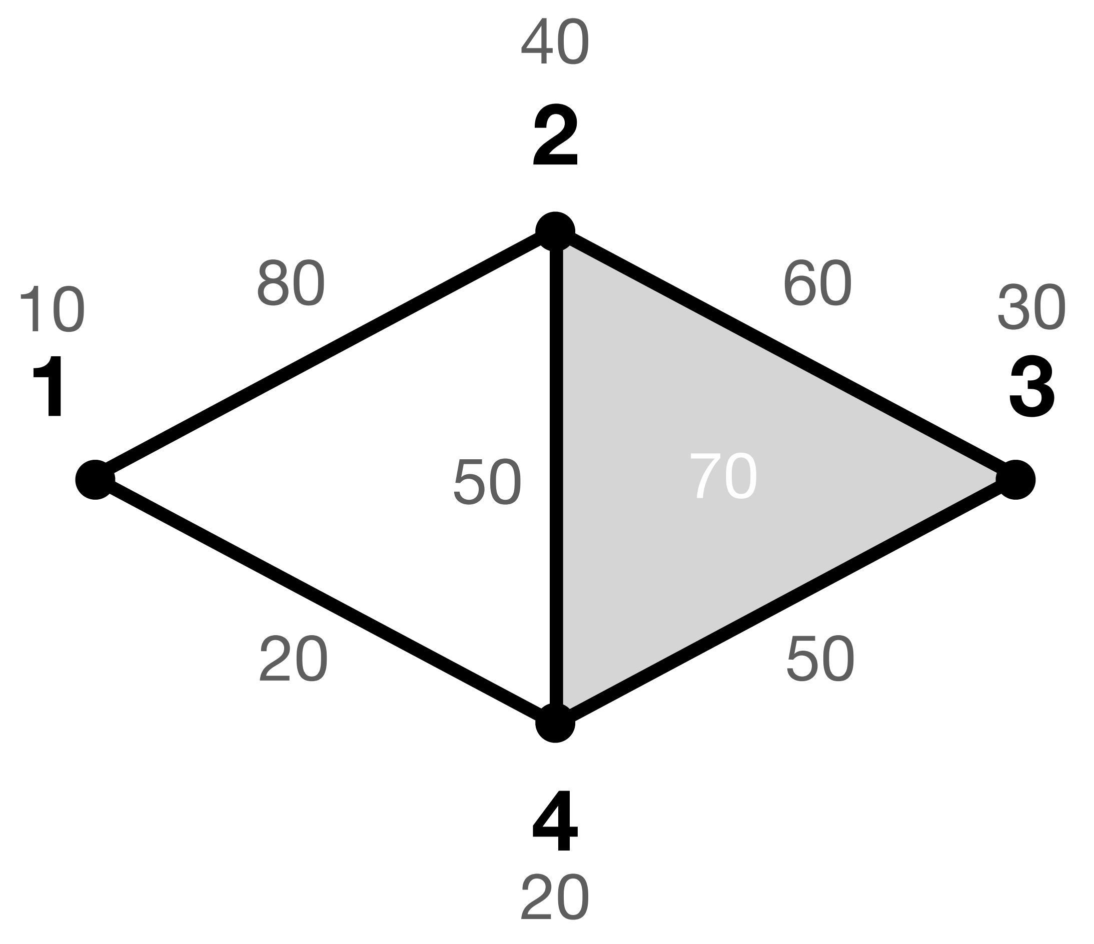


$\mathbf{R} =$

								[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

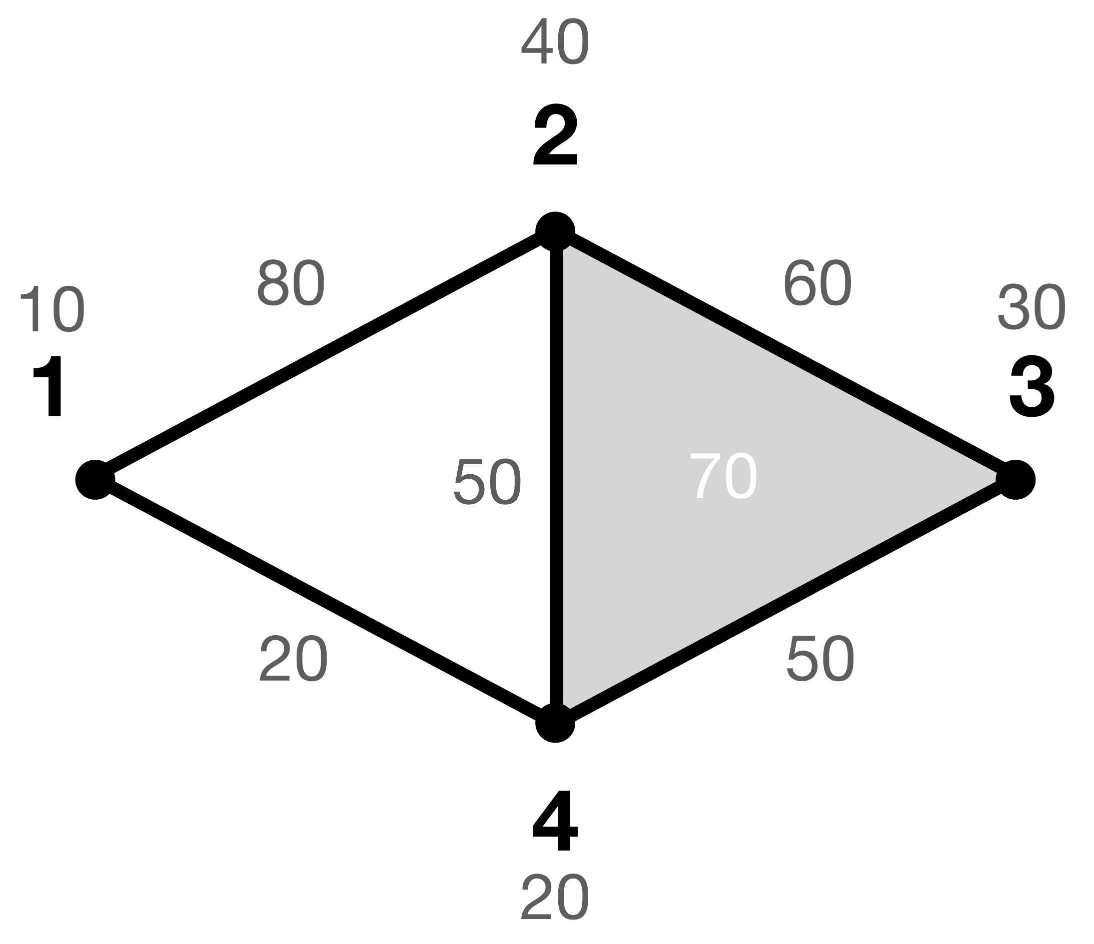


$\mathbf{R} =$

								[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

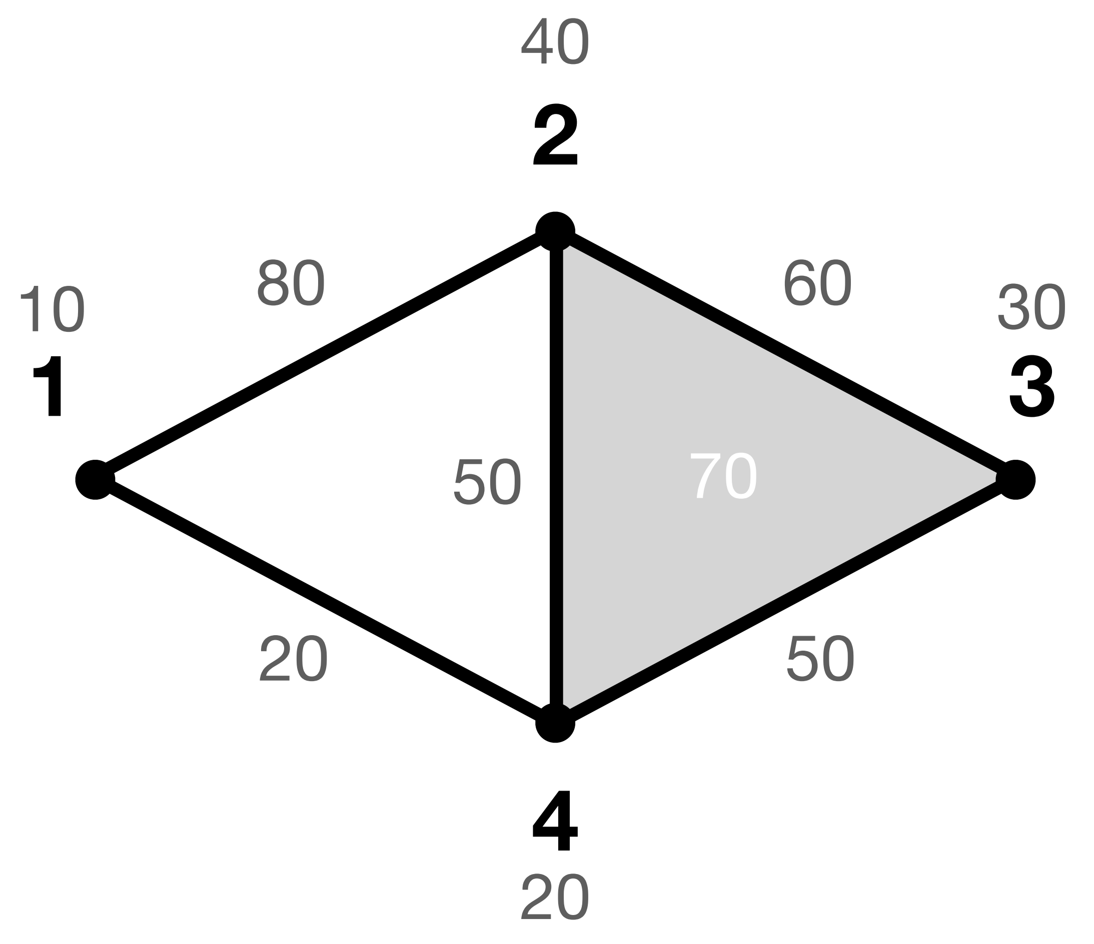
(3, 34) 0

(23, 234) 1

(1, $\underline{\varnothing}$) 0

(12, $\underline{\varnothing}$) 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

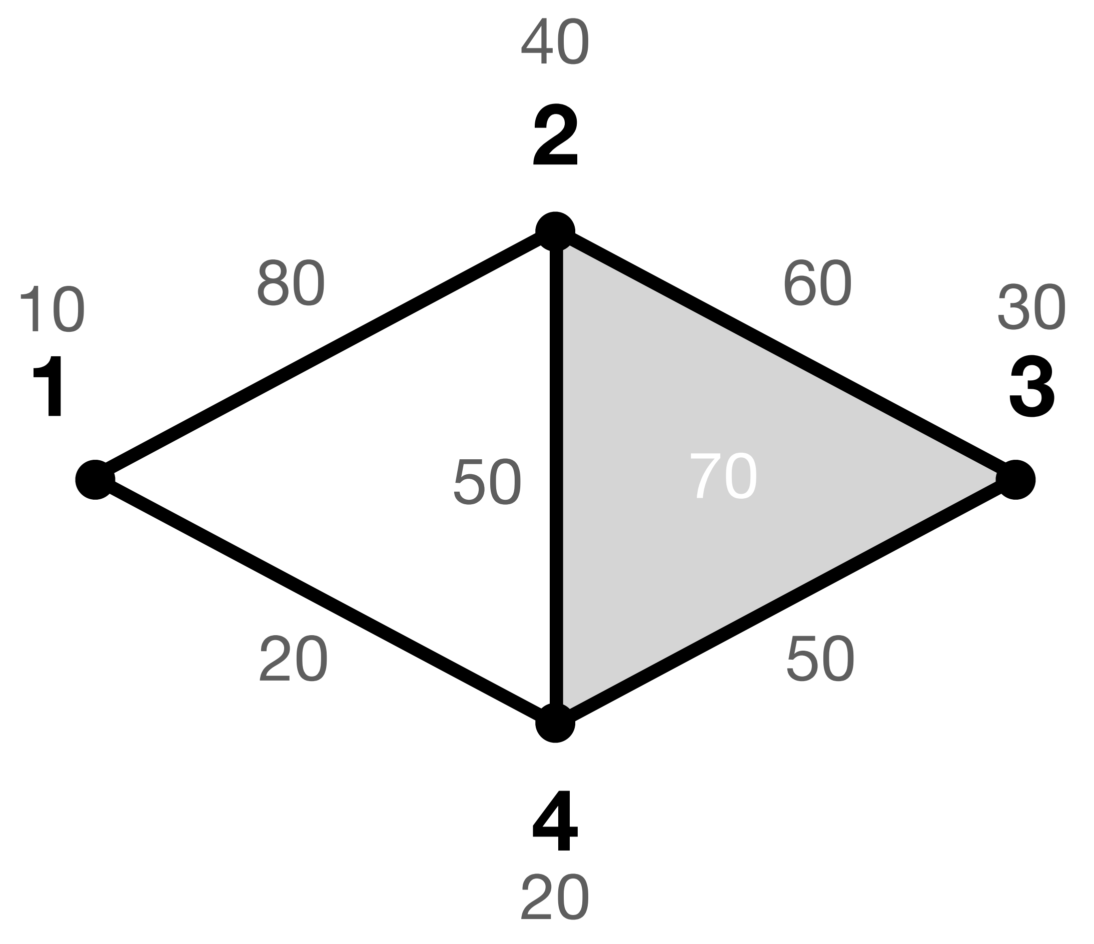
(1, Ø) 0
(12, Ø) 1

Persistence diagram

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

(10, Ø) 0
(80, Ø) 1

Extracting information



$\mathbf{R} =$

									[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										

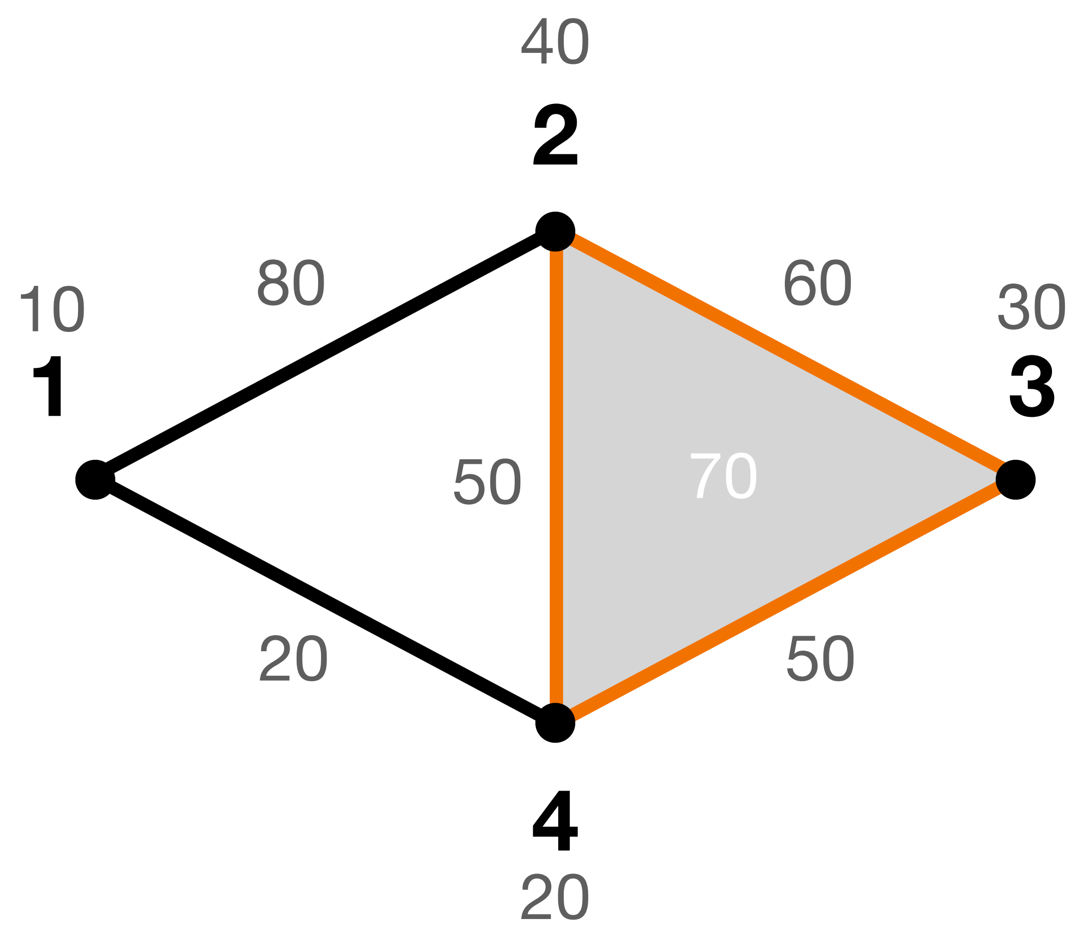
Persistence pairing [representatives]

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

Persistence diagram

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

Representatives

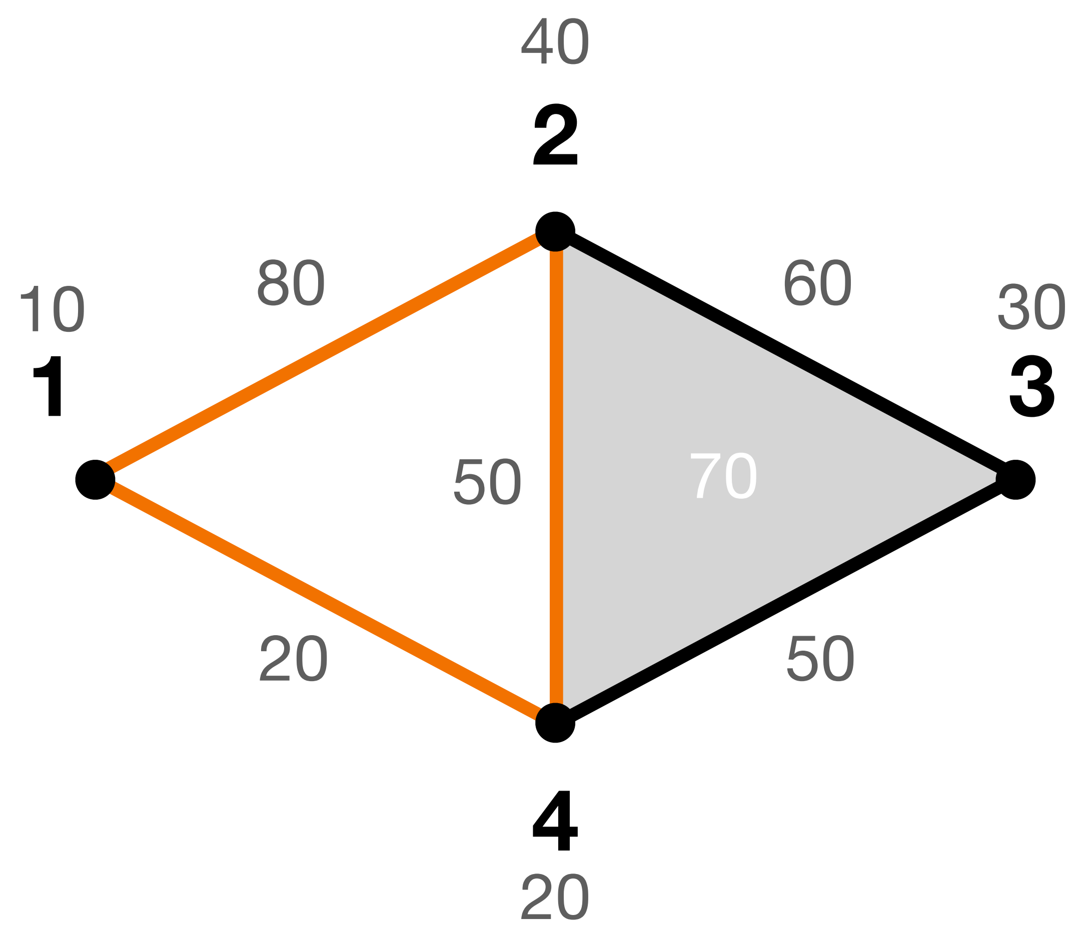


$\mathbf{R} =$

								[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

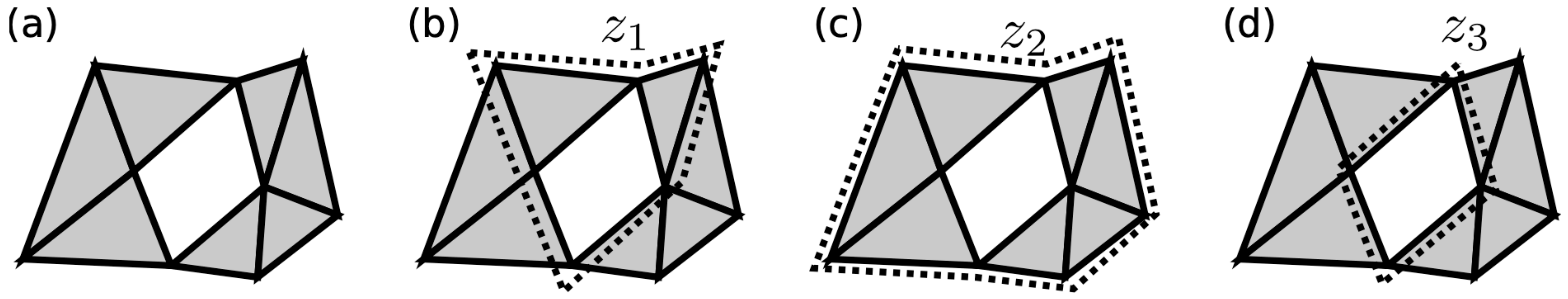


$\mathbf{R} =$

								[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

Homologous cycles

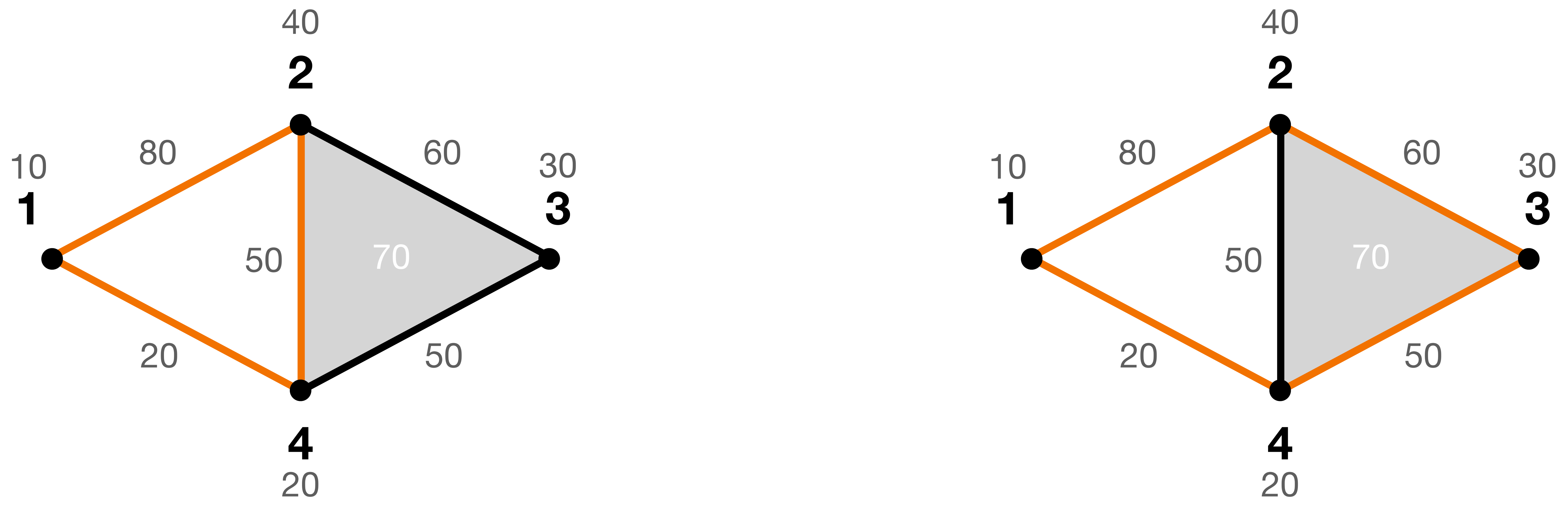


Homologous cycles

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

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

Homologous cycles



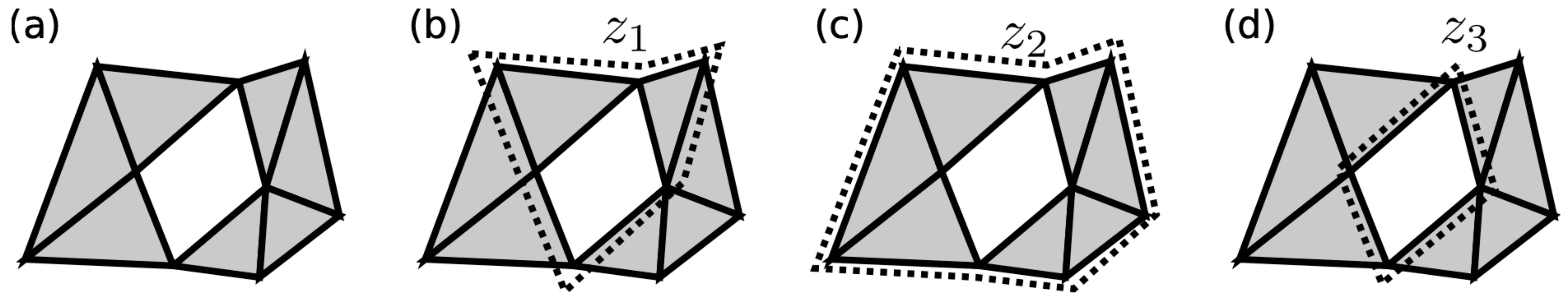
Homologous cycles

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

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

Representatives

Optimal representatives



Optimal cycle

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

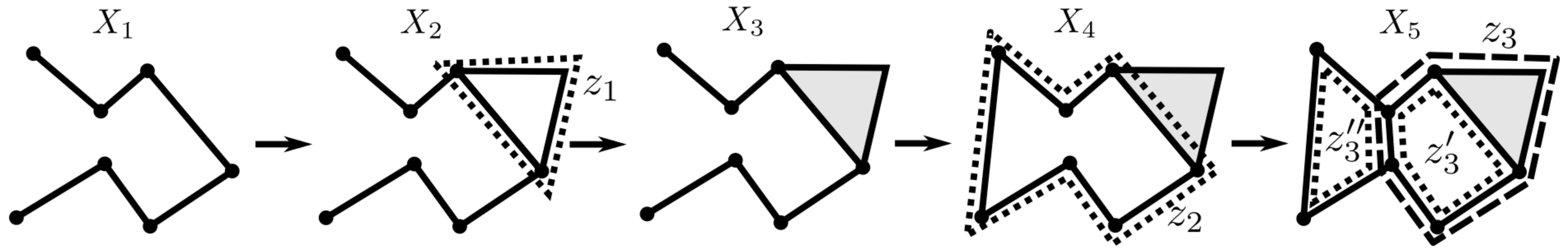
$$\begin{aligned} &\text{minimize } \|z\|_0 \quad \text{subject to} \\ &\qquad z = z_1 + \partial w, \\ &\qquad w \in C_2(X). \end{aligned}$$

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)*

Persistence representatives

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



$$\text{minimize } \|z\|_0 \ \text{ subject to}$$
$$z = z_3 + \partial w + k z_2,$$
$$w \in C_2(X_5),$$
$$k \in \mathbb{k}.$$

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:

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

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

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

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

Betti numbers via Hodge Laplacian

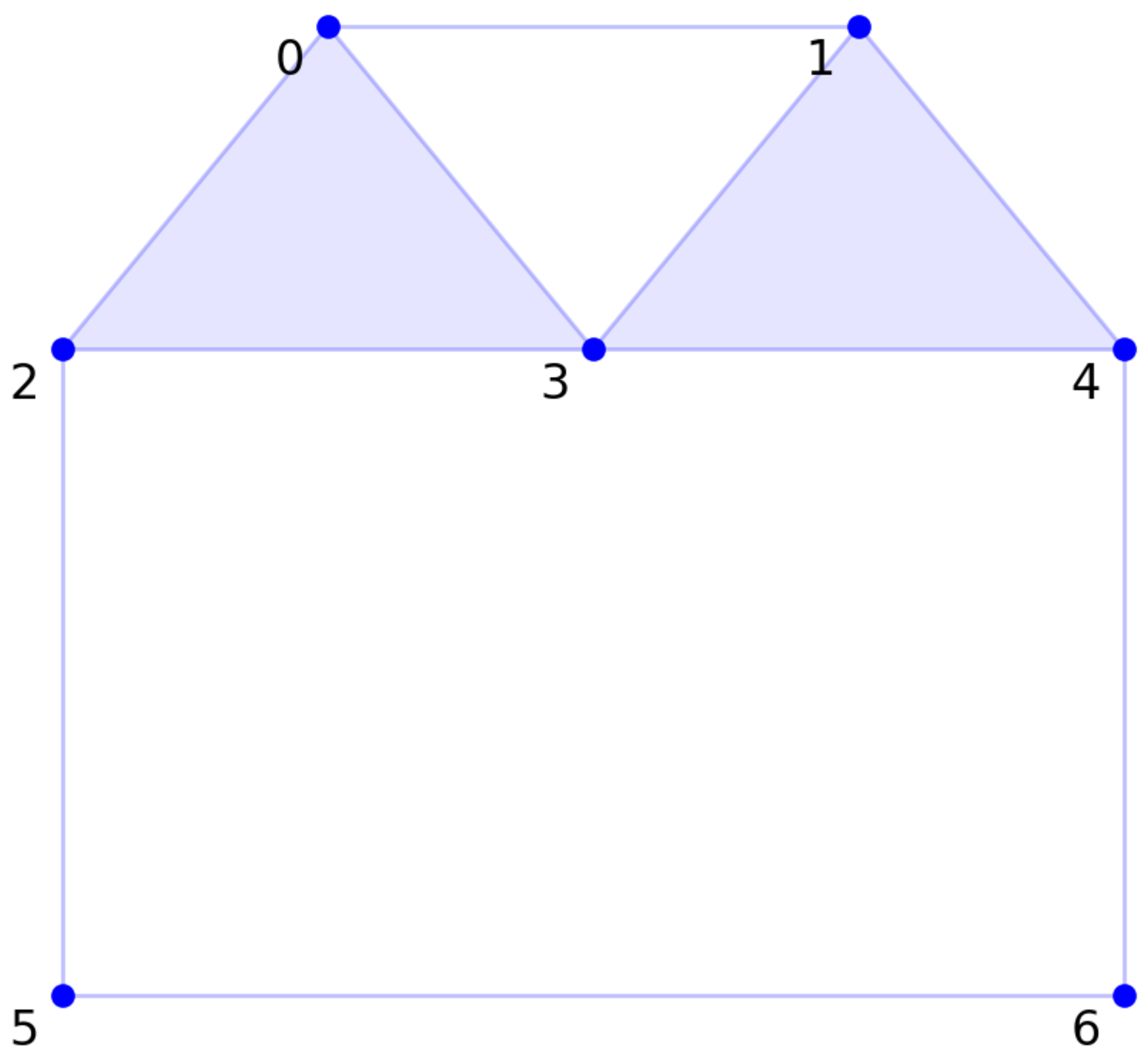
$$\beta_k = \dim \ker(\mathbf{L}_k)$$

Harmonic cycles

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

Hodge decomposition

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



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

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

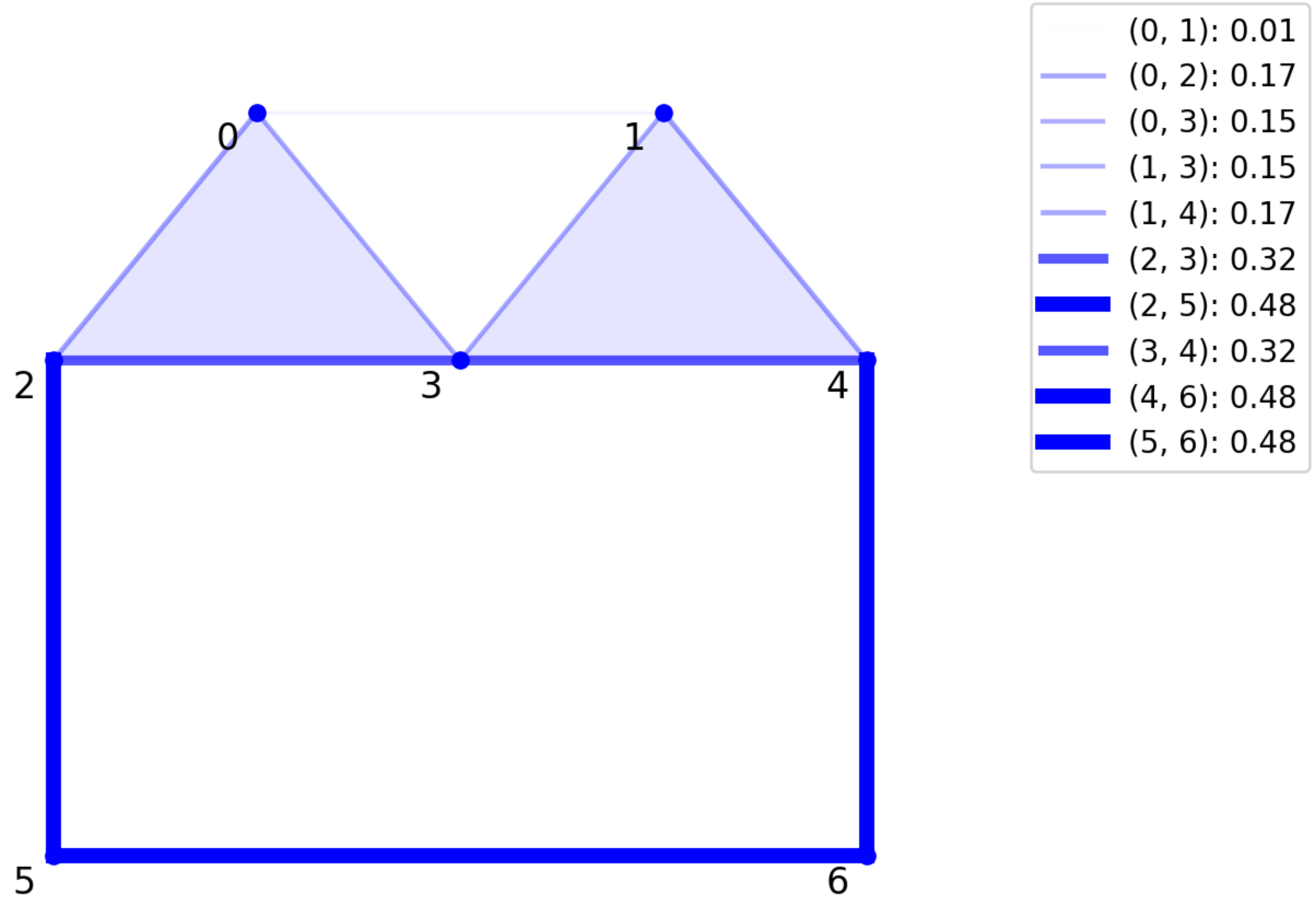
Betti numbers via Hodge Laplacian

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

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

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

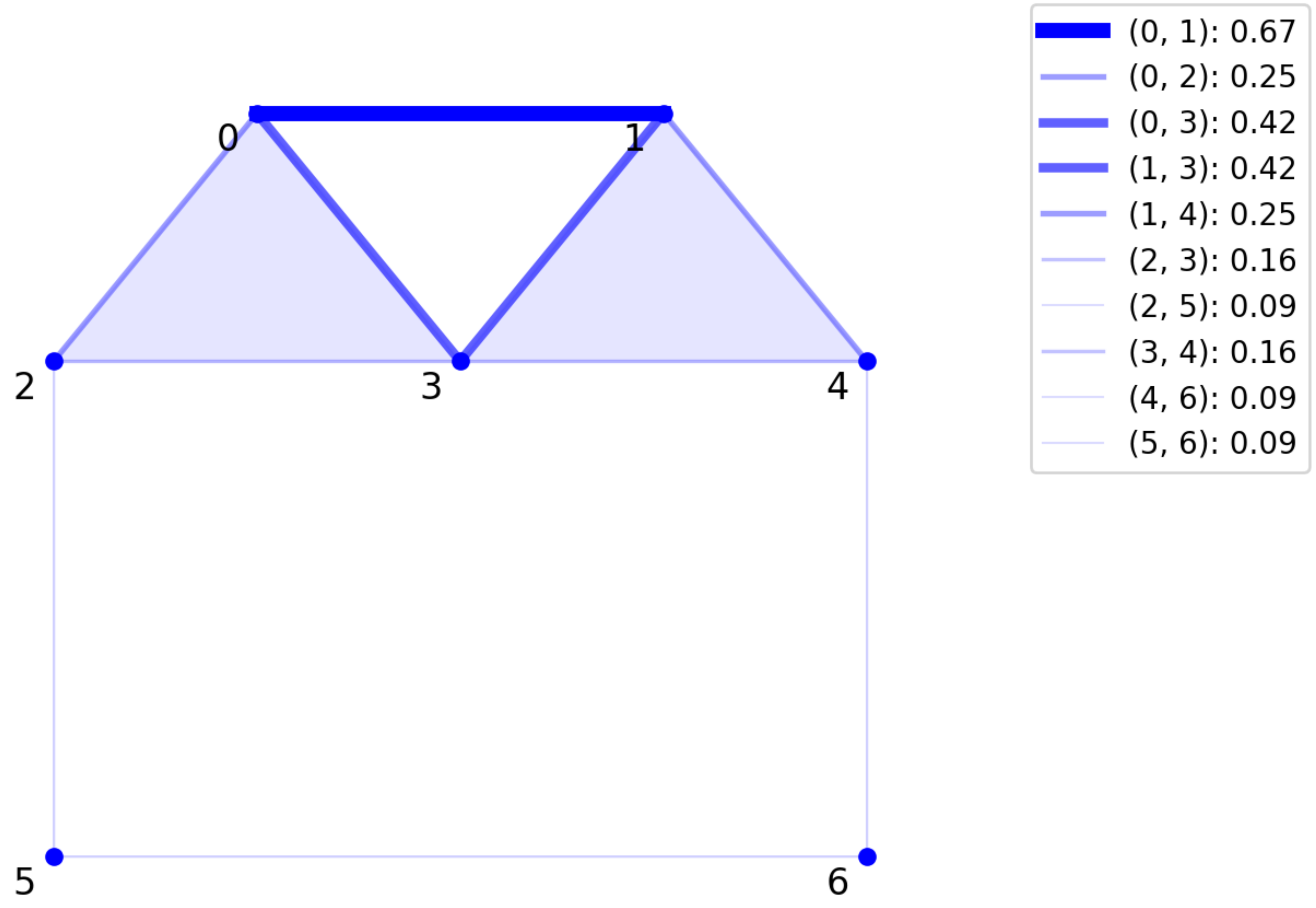
Betti numbers via Hodge Laplacian

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

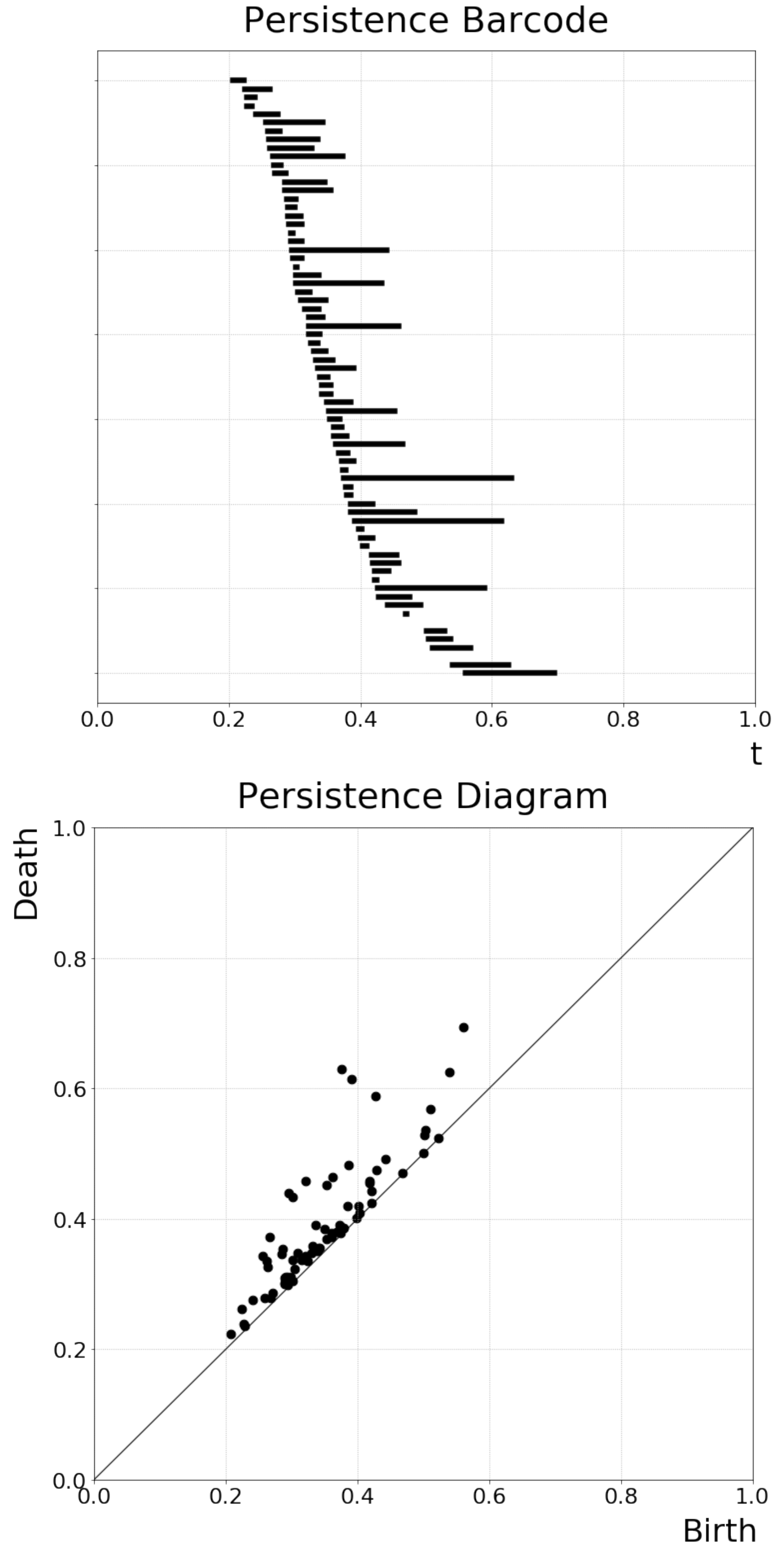
Harmonic cycles

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

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

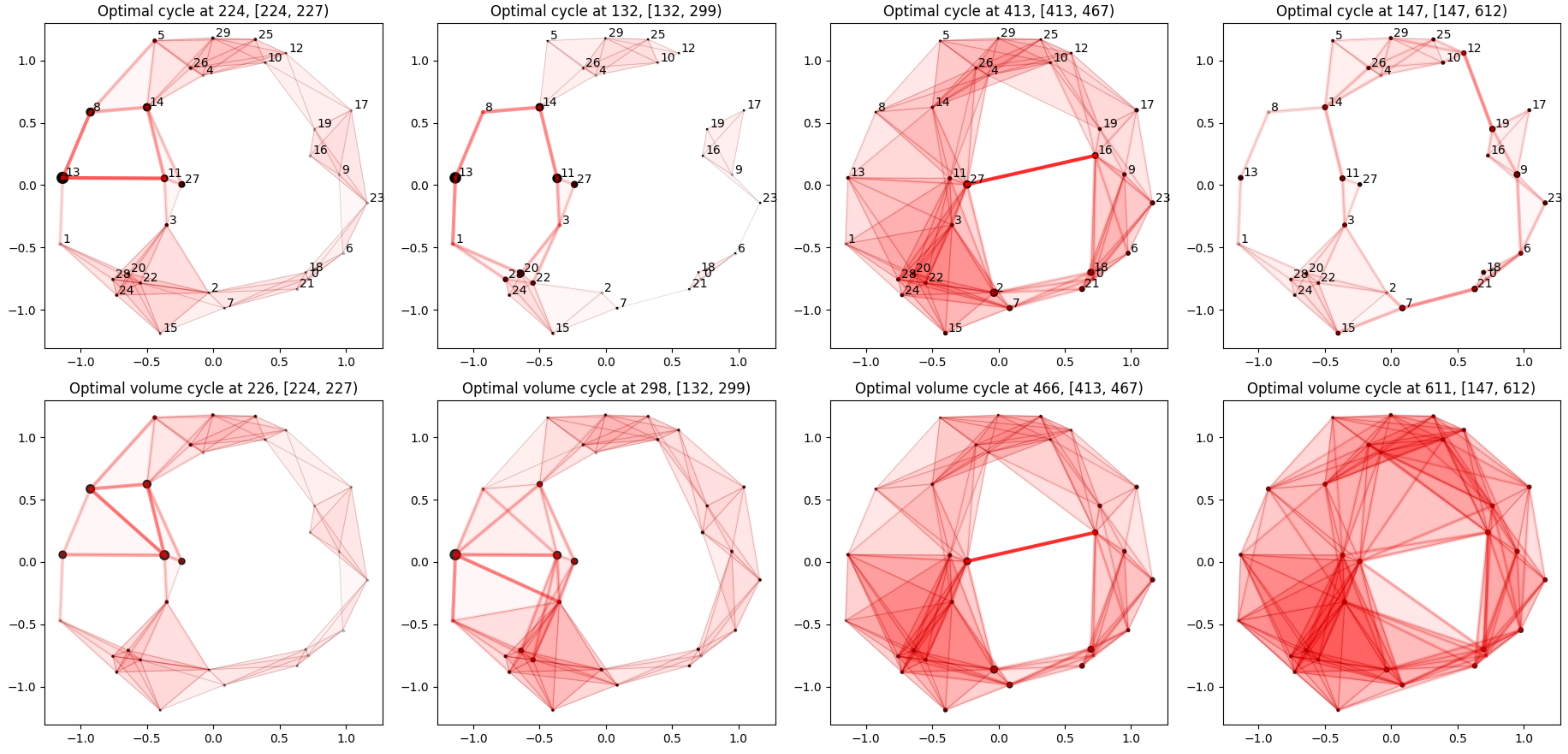


Persistent harmonic representatives

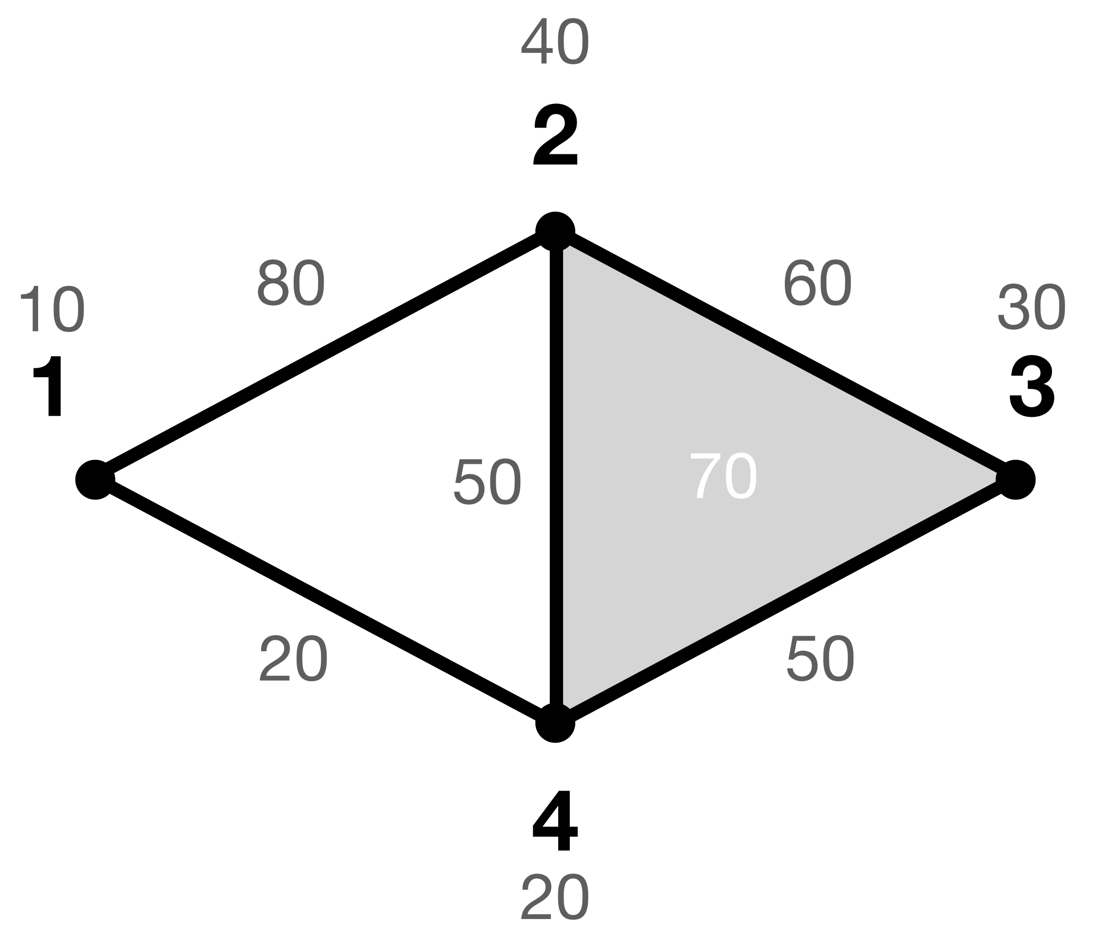


Persistence diagram

$$D_{f_\theta}(X) = \{(f(b_i), f(d_i))\}_{i \in I}$$



Boundary matrix at filtration step *t*



$\mathbf{B} =$

	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										

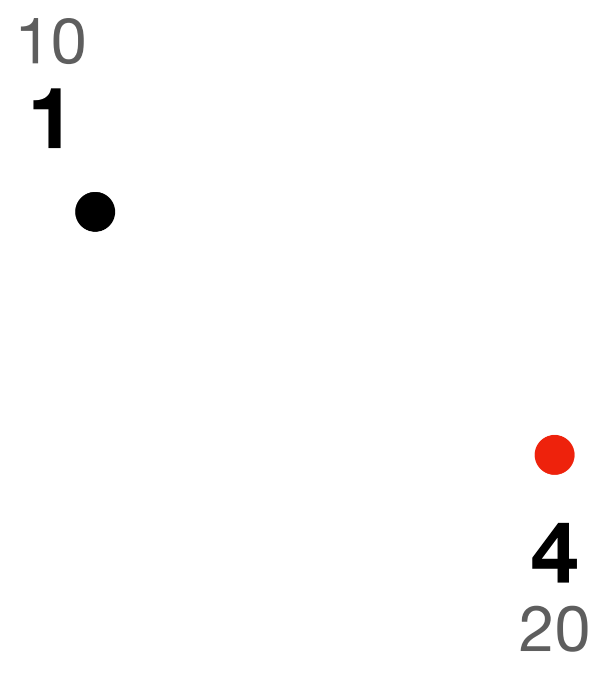
Filtration function provides order on simplices, therefore on columns and rows of the filtration matrix. Ties are broken, first by simplex dimension, second by lexicographic order given by order on vertices.

Boundary matrix at filtration step *t*

	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										

dt=1

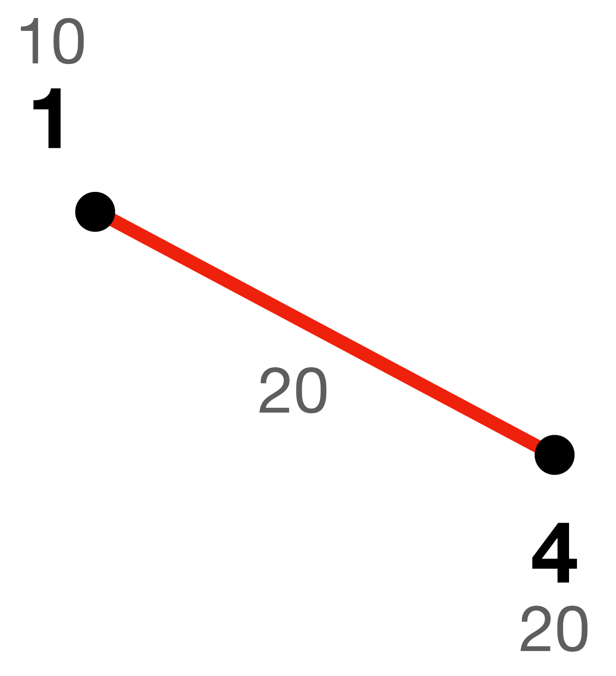
Boundary matrix at filtration step *t*



	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										

dt=4

Boundary matrix at filtration step *t*

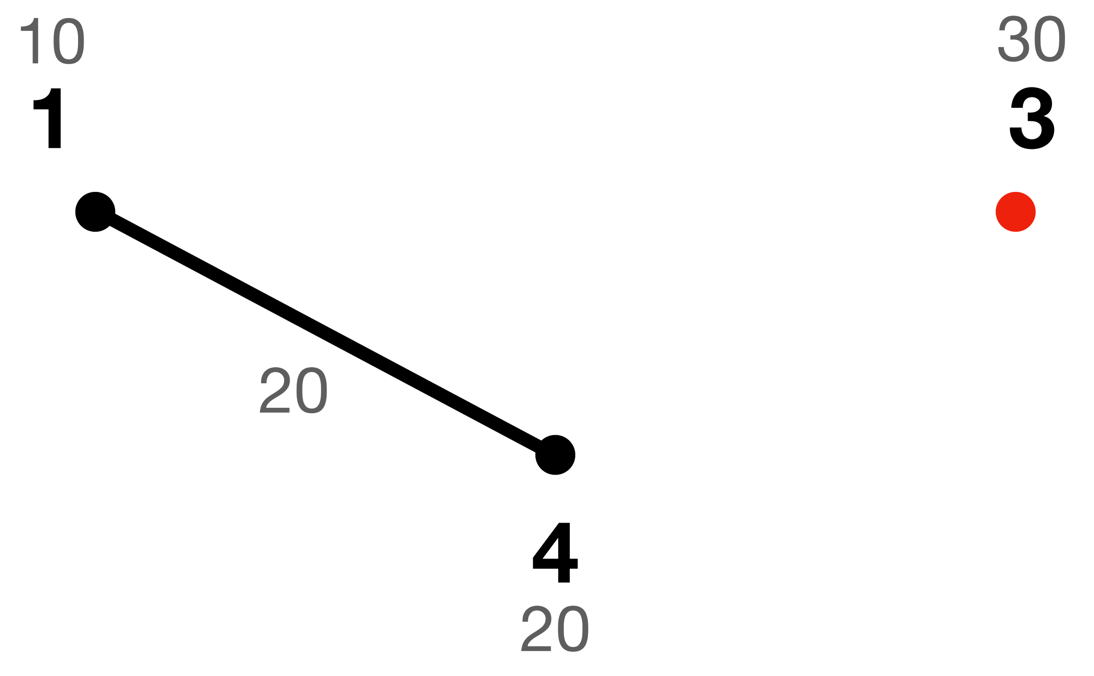


	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

dt=14

$\mathbf{B}_1(t)$

Boundary matrix at filtration step t

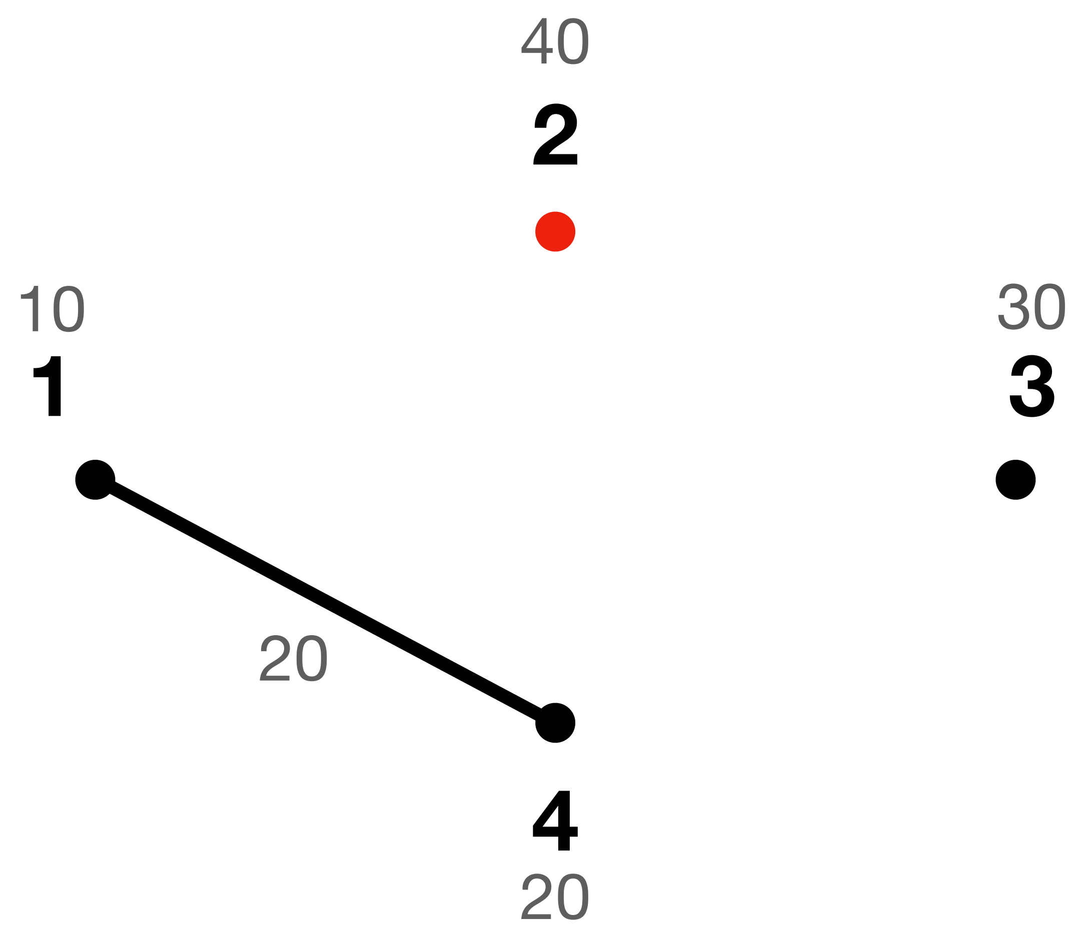


	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

dt=3

$\mathbf{B}_1(t)$

Boundary matrix at filtration step *t*

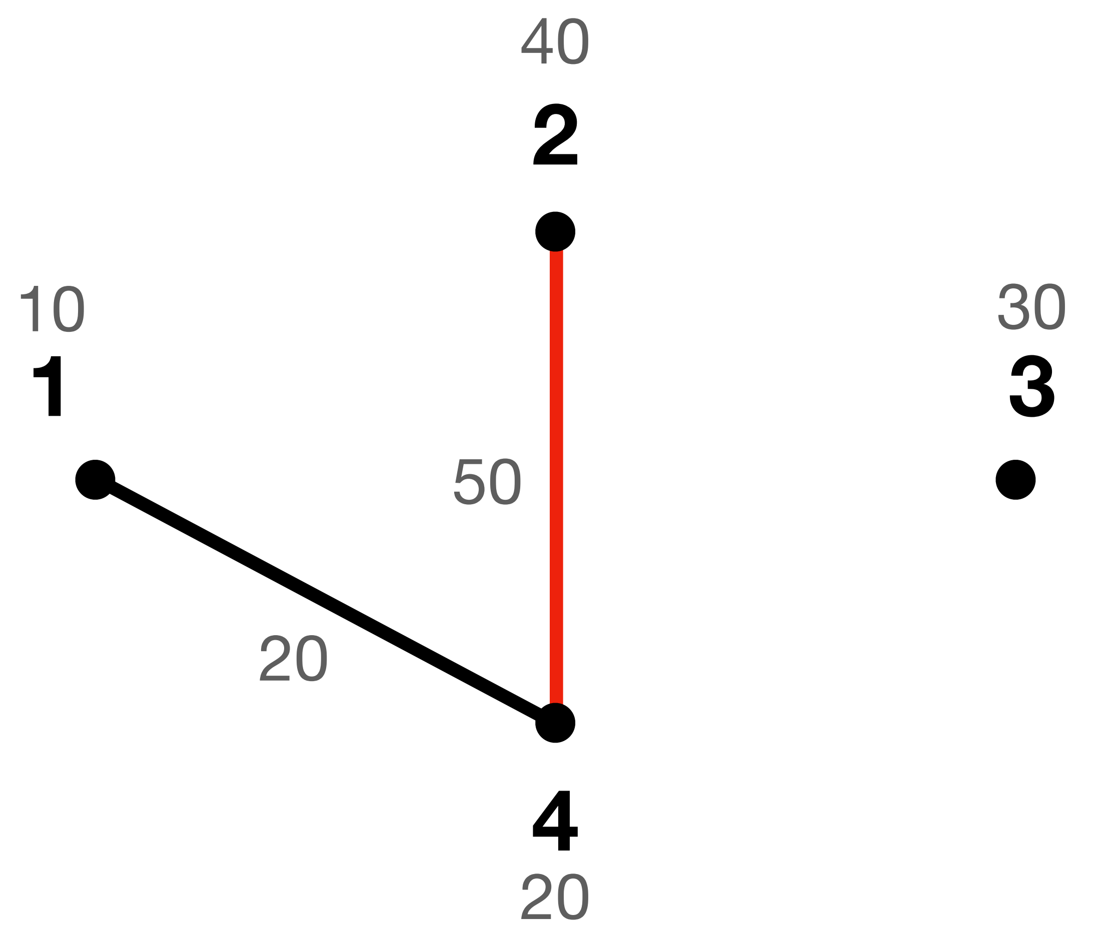


	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

dt=2

$\mathbf{B}_1(t)$

Boundary matrix at filtration step *t*

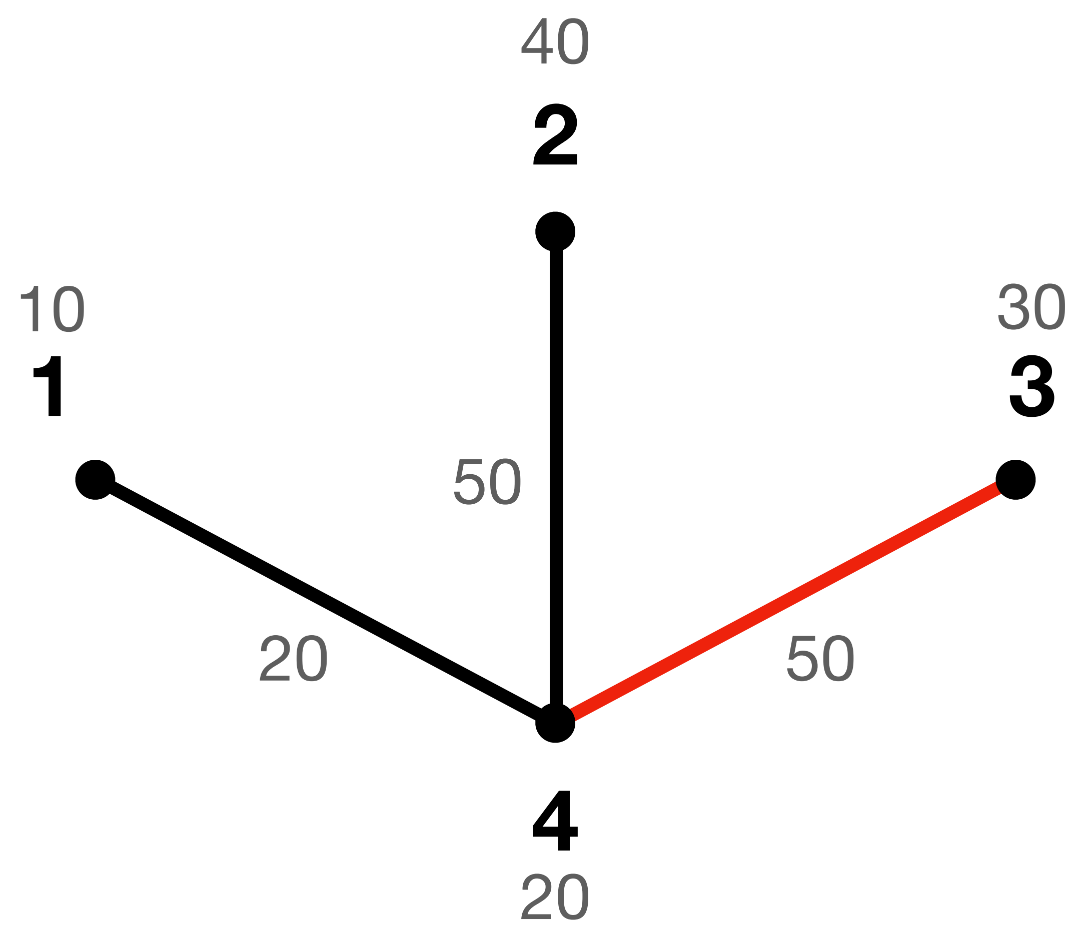


	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

dt=24

$\mathbf{B}_1(t)$

Boundary matrix at filtration step *t*

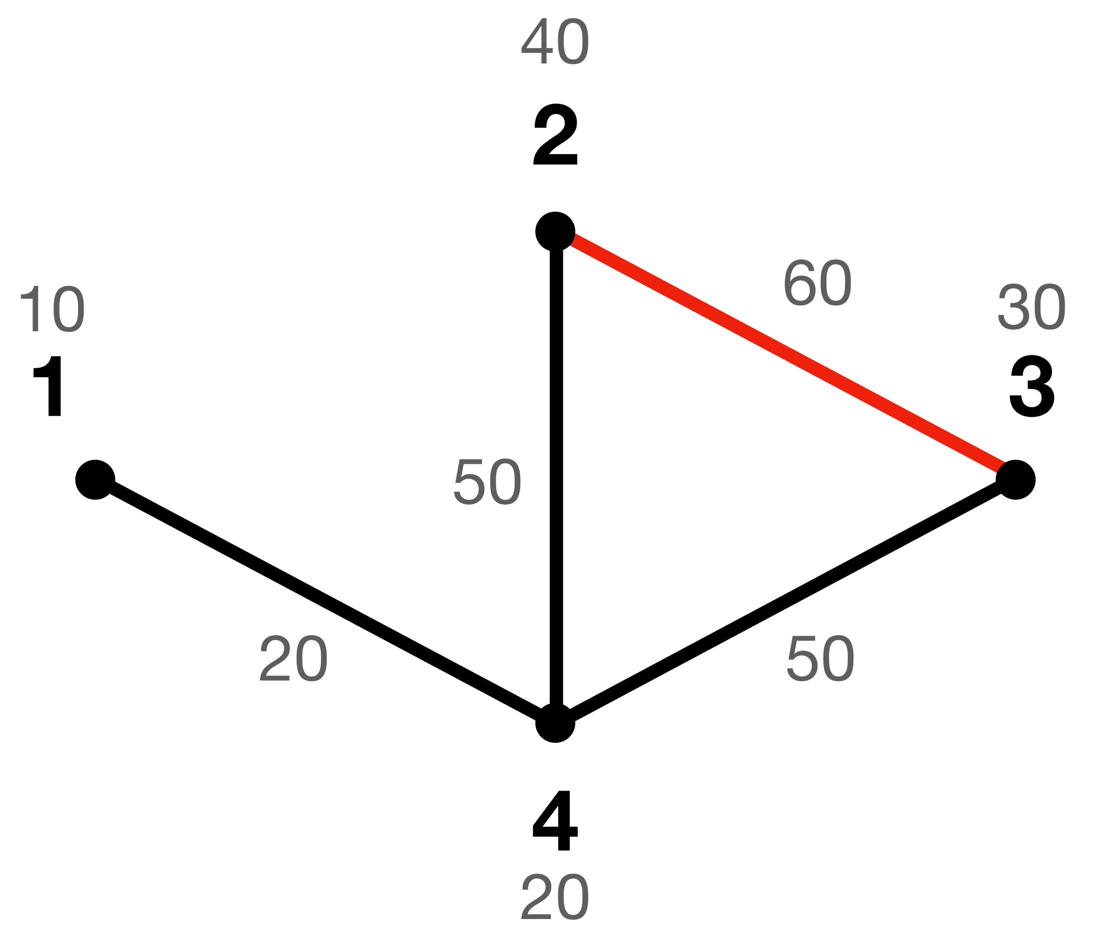


	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

dt=34

$\mathbf{B}_1(t)$

Boundary matrix at filtration step *t*

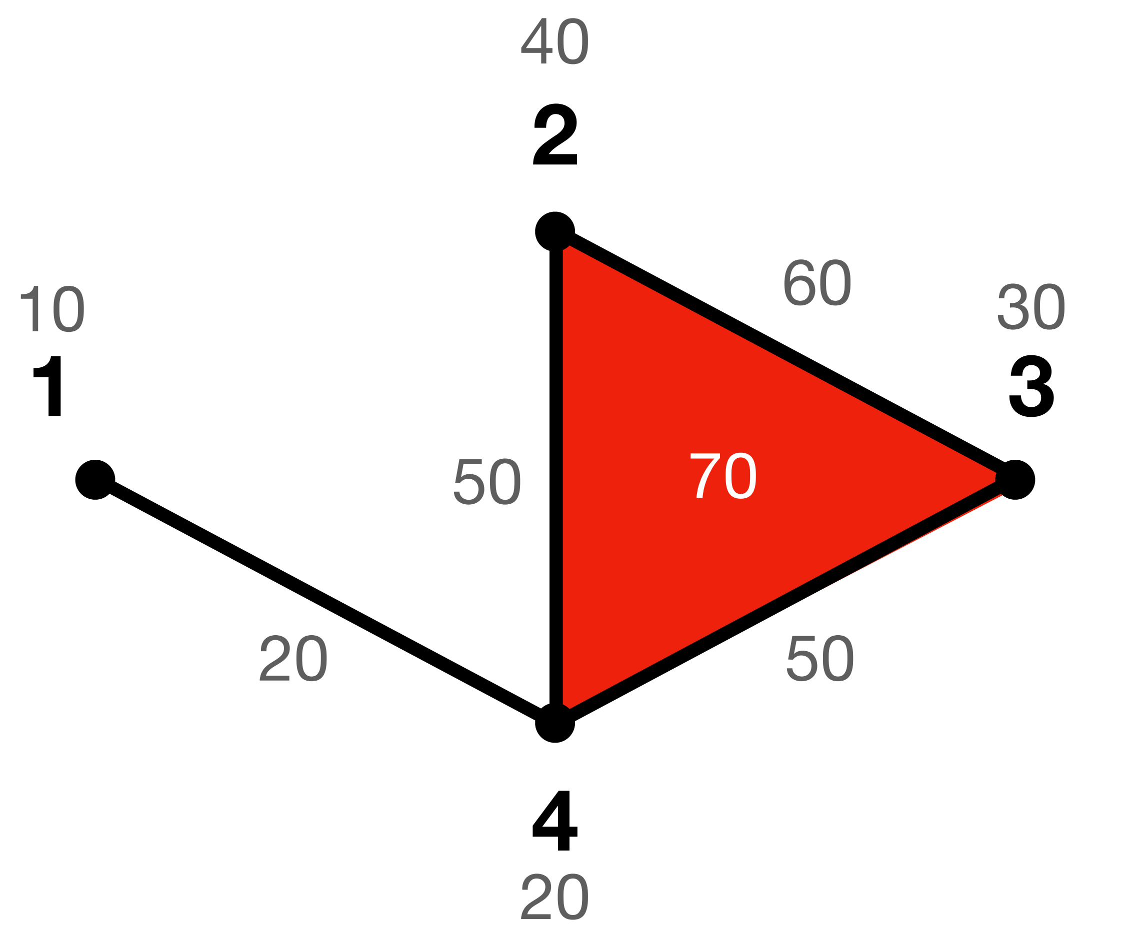


	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

dt=23

$\mathbf{B}_1(t)$

Boundary matrix at filtration step *t*

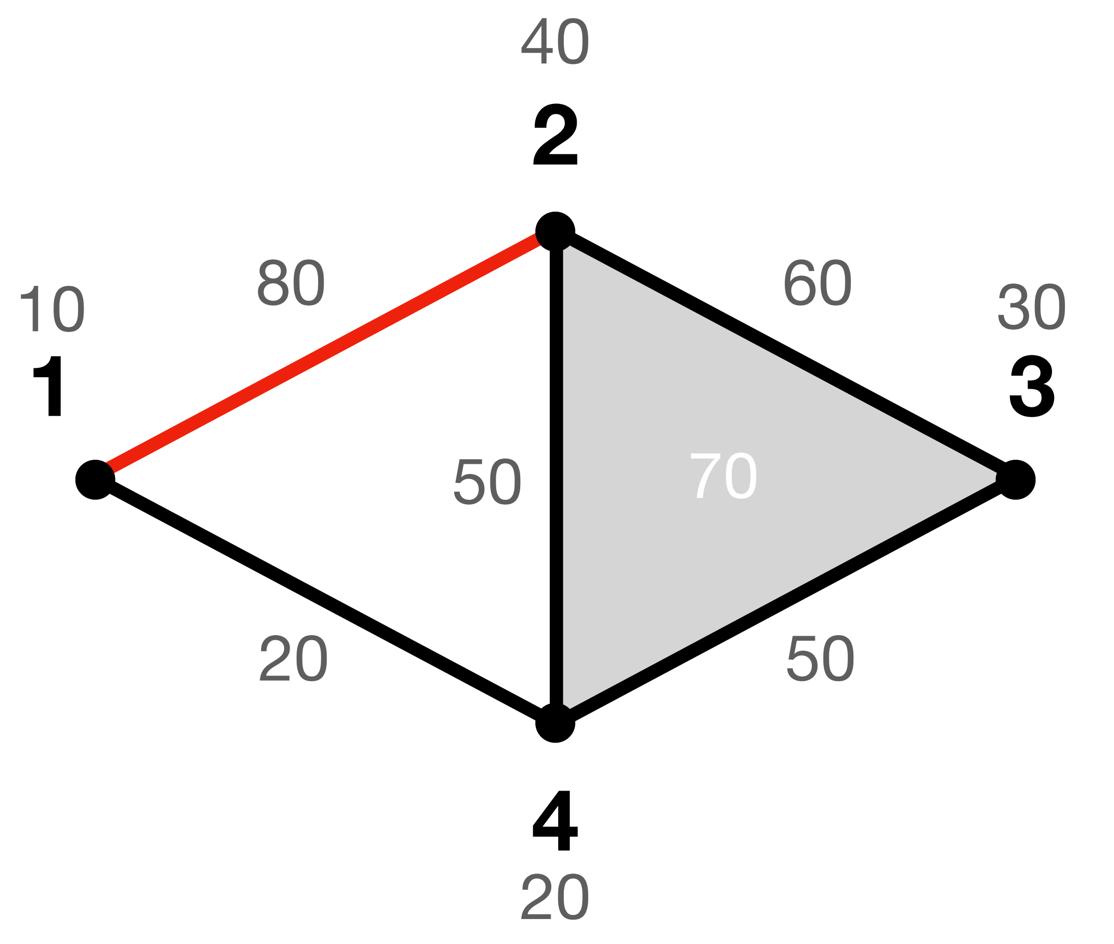


	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

dt=234

$\mathbf{B}_1(t)$ $\mathbf{B}_2(t)$

Boundary matrix at filtration step *t*



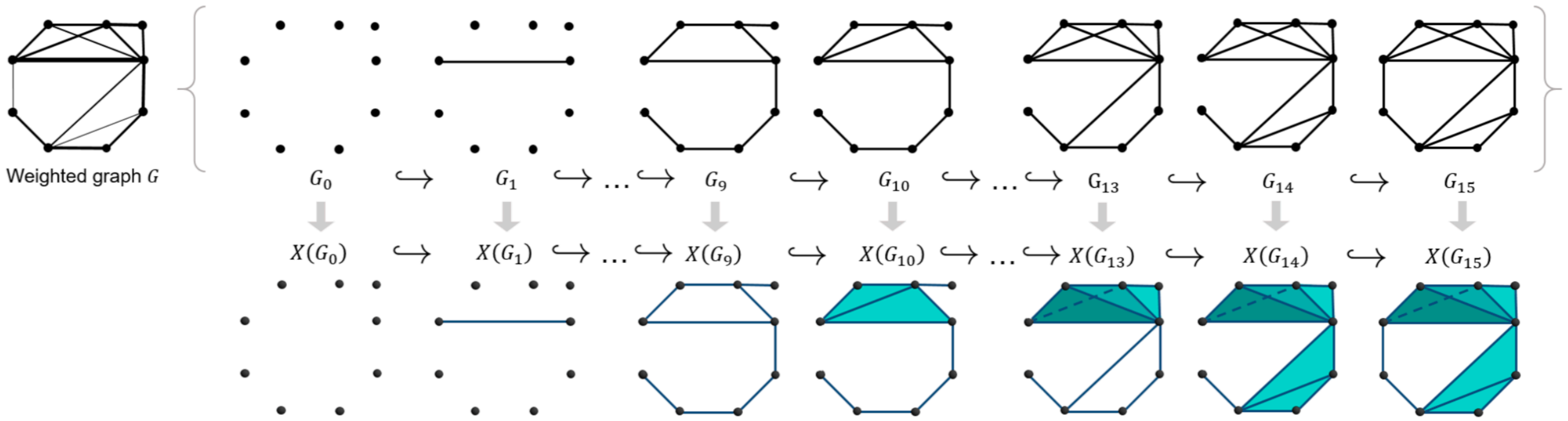
	10	20	20	30	40	50	50	60	70	80
	1	**4**	**14**	**3**	**2**	**24**	**34**	**23**	**234**	**12**
1			1							1
4			1			1	1			
14										
3							1	1		
2						1		1		1
24									1	
34									1	
23									1	
12										

dt=12

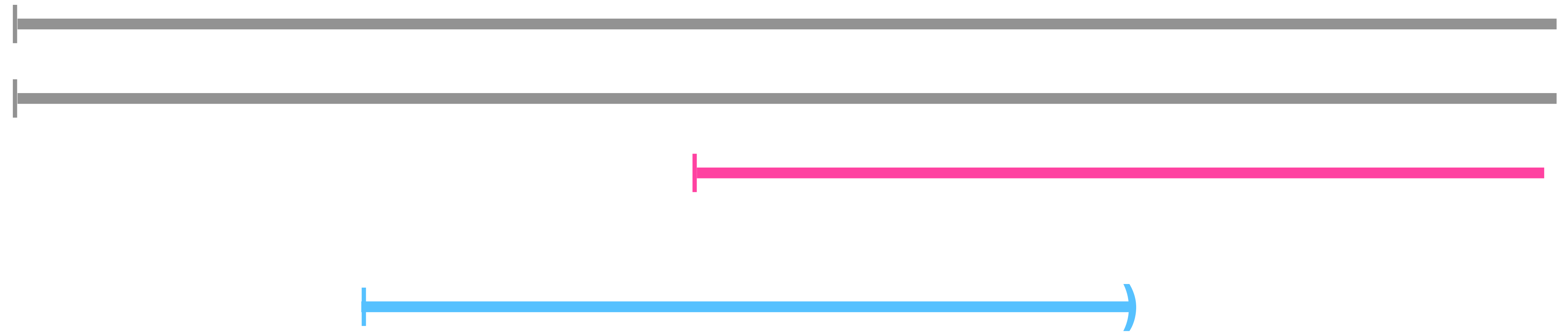
$\mathbf{B}_1(t)$ $\mathbf{B}_2(t)$

Persistent harmonic representatives

Filtration

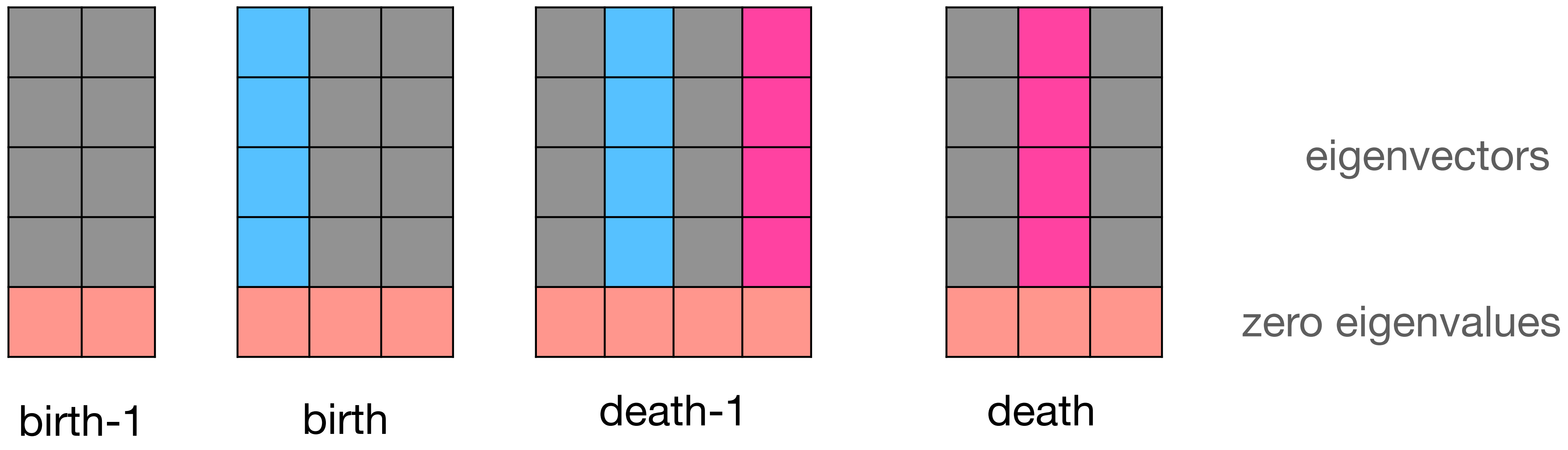


Persistence barcode



Higher-order Laplacian at filtration step *t*

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



Persistent harmonic representatives

Solve two eigenproblems on $\mathbf{L}_{k,t}$ with $t = b_i$ and $t = b_i - 1$ to find two sets of eigenvalues $\boldsymbol{\lambda}_{k,t}$ and corresponding eigenvectors $\mathbf{V}_{k,t}$ of dimension k at time t to obtain zero eigenvalues (up to a tolerance hyperparameter δ) and their corresponding eigenvectors

- eigenvalues $\boldsymbol{\lambda}^0_{k,b_i-1}$ and corresponding eigenvectors $\mathbf{V}^0_{k,b_i-1}$ at time $b_i - 1$ immediately preceding birth time b_i,
- eigenvalues $\boldsymbol{\lambda}^0_{k,b_i}$ and corresponding eigenvectors $\mathbf{V}^0_{k,b_i}$ at birth time b_i.

Find the representative vector $\mathbf{w}^+_{k,b_i}$ at birth such that $\mathbf{w}^+_{k,b_i} \in \mathbf{V}^0_{k,b_i}$, but $\mathbf{w}^+_{k,b_i} \notin \mathbf{V}^0_{k,b_i-1}$. To do so,

$$\mathbf{w}^+_{k,b_i} = \arg \max_{\mathbf{v}^0_{k,b_i} \in \mathbf{V}^0_{k,b_i}} \min \sigma \left(\mathbf{V}^0_{k,b_i-1} \mid \mathbf{v}^0_{k,b_i}[: -1] \right), \tag{11}$$

where $\min \sigma(\mathbf{X} \mid \mathbf{x})$ is the smallest singular value of the augmented matrix $\mathbf{X}$ by a vector $\mathbf{x}[: -1]$, i.e. with the last element truncated.

Higher-order Laplacian at filtration step *t*

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