Scalable Laplacian K-modes



Imtiaz Masud Ziko, Eric Granger and Ismail Ben Ayed

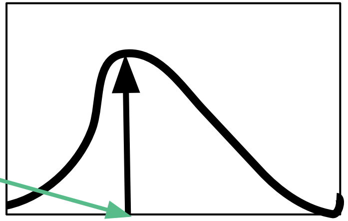
Laplacian K-modes (LK) [*Wang and Carreira-Perpiñán 2014*]

Laplacian K-modes (LK) [*Wang and Carreira-Perpiñán 2014*]

K-modes

$$-\sum_{p=1}^{N}\sum_{l=1}^{L} z_{p,l}\, k(\mathbf{x}_p, \mathbf{m}_l)$$

Mode ($\mathbf{m}_l$)



Laplacian K-modes (LK) [*Wang and Carreira-Perpiñán 2014*]

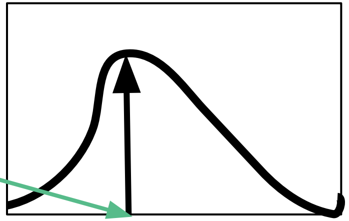
K-modes

$$-\sum_{p=1}^{N}\sum_{l=1}^{L} z_{p,l}\, k(\mathbf{x}_p, \mathbf{m}_l)$$

Laplacian

$$\frac{\lambda}{2}\sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\|\mathbf{z}_p - \mathbf{z}_q\|^2$$

Mode ($\mathbf{m}_l$)



Zhu '02, Weston '08, Shi '00, Belkin '03, '06 etc

Laplacian K-modes (LK) [*Wang and Carreira-Perpiñán 2014*]

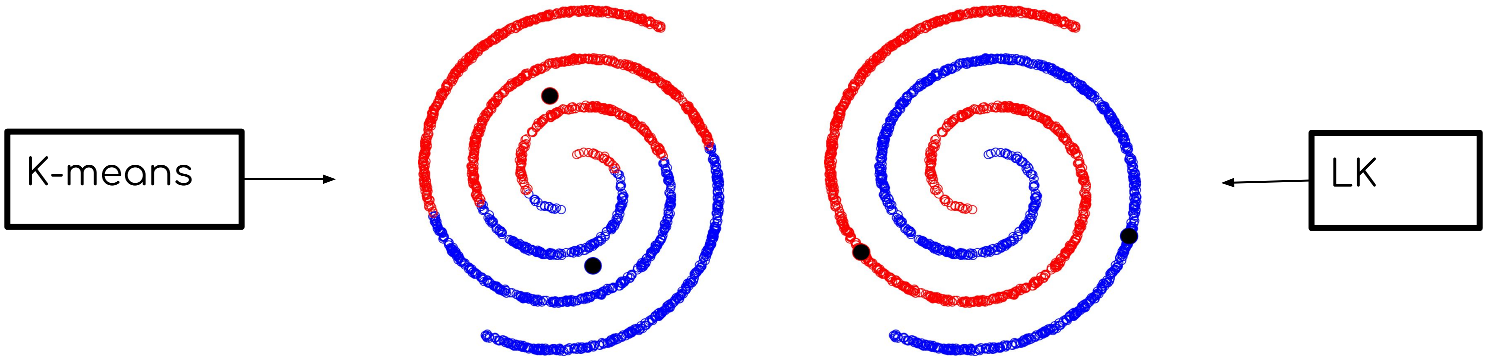
$$\min_{\mathbf{Z}} \left[\underbrace{- \sum_{p=1}^{N} \sum_{l=1}^{L} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l)}_{\text{K-modes}} + \underbrace{\frac{\lambda}{2} \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q) \|\mathbf{z}_p - \mathbf{z}_q\|^2}_{\text{Laplacian}} \right]$$

Laplacian K-modes (LK) [*Wang and Carreira-Perpiñán 2014*]

$$\min_{\mathbf{Z}} \left[\underbrace{- \sum_{p=1}^{N} \sum_{l=1}^{L} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l)}_{\text{K-modes}} + \underbrace{\frac{\lambda}{2} \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q) \|\mathbf{z}_p - \mathbf{z}_q\|^2}_{\text{Laplacian}} \right]$$

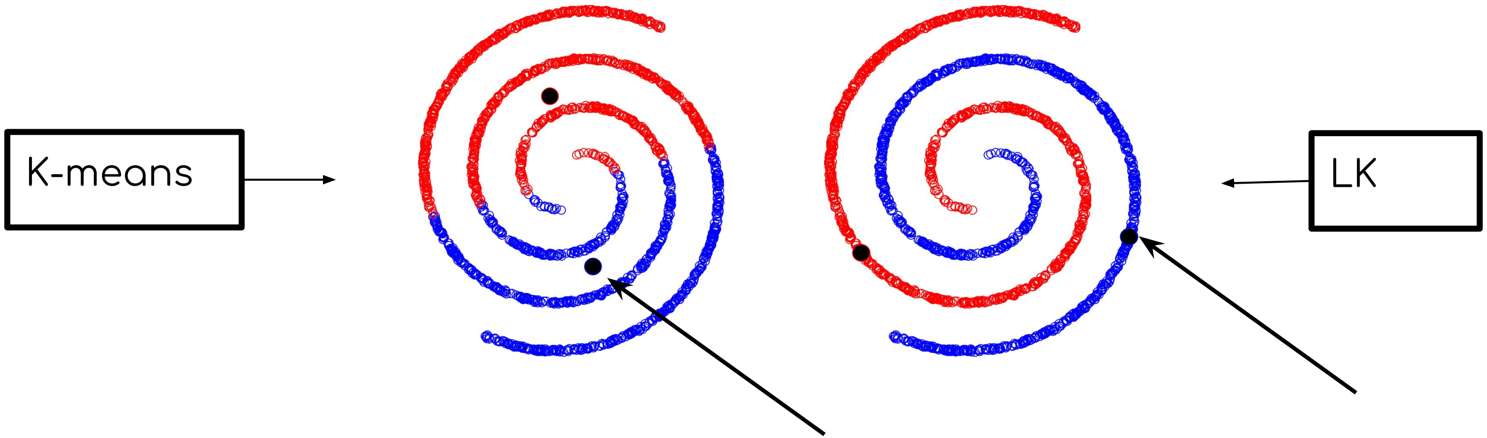
$$\underbrace{\mathbf{1}^t \mathbf{z}_p = 1,}_{\text{Simplex constraint}} \ \underbrace{\mathbf{z}_p \in \{0, 1\}^L}_{\text{Discrete}} \ \forall p$$

Why Laplacian K-modes?



★ Handles **non convex (manifold structured) clusters.**

Why Laplacian K-modes?



- ★ Handles **non convex (manifold structured) clusters.**

- ★ **Mean** or **Mode ?**

Why Laplacian K-modes?

Mode images

Mean images

★ Handles **non convex (manifold structured) clusters.**

★ **Mean** or **Mode** ?

☑ Prototypes from input set

Laplacian k-modes: challenges

- **Challenging Optimization problem:**
  - simplex/integer constraint.
  - Dependance of modes on $\mathbf{z}_p$
  - Laplacian over discrete variable!

Laplacian k-modes: challenges

- Challenging Optimization problem:

- Well-known Spectral relaxation *[Shi & Malik '00]* :

Laplacian k-modes: challenges

- Challenging Optimization problem:

- Well-known Spectral relaxation *[Shi & Malik '00]* :
  - Eigen-decomposition of Laplacian (***N*** *x* ***N***).

Laplacian k-modes: challenges

- **Challenging Optimization problem:**
- **Well-known Spectral relaxation** *[Shi & Malik ‘00]*:
  - Eigen-decomposition of Laplacian ($N \times N$).
- **Convex relaxation (relax integer constraint)** *[Wang and Carreira-Perpiñán ‘14]*:
  - Solve over $N \times L$ variables altogether.
  - Projection to L-dimensional simplex.

Not applicable in large scale clustering

Laplacian k-modes: challenges

- Challenging Optimization problem:

- **Well-known Spectral relaxation** *[Shi & Malik '00]* :
    - Eigen-decomposition of **Laplacian** (*N x N*).

Concave Relaxation

- **Convex relaxation (relax integer constraint)** *[Wang and Carreira-Perpiñán '14]* :
    - Solve over *N x L* variables altogether.
    - Projection to *L*-dimensional simplex.

We Tackle

Laplacian k-modes: challenges

- Challenging Optimization problem:

- **Well-known Spectral relaxation** *[Shi & Malik '00]*:
  - Eigen-decomposition of **Laplacian** (N x N).

- **Convex relaxation (relax integer constraint)** *[Wang and Carreira-Perpiñán '14]*:
  - Solve over N x L variables altogether. → *Parallel structure*
  - Projection to L-dimensional simplex.

We Tackle

Laplacian k-modes: challenges

- Challenging Optimization problem:

- **Well-known Spectral relaxation** *[Shi & Malik '00]* :
    - Eigen-decomposition of **Laplacian** (*N x N*).

- **Convex relaxation (relax integer constraint)** *[Wang and Carreira-Perpiñán '14]* :

We Tackle

- Solve over *N x L* variables altogether.
- Projection to *L*-dimensional simplex. → *avoid*

SLK Concave-Convex Relaxation

Laplacian

$\mathrm{tr}(\mathbf{Z}^t\mathbf{L}\mathbf{Z})$

SLK Concave-Convex Relaxation

Laplacian

Direct convex relaxaton

Concave relaxation (ours)

$$\boxed{\mathrm{tr}(\mathbf{Z}^t\mathbf{L}\mathbf{Z})} = \boxed{\sum_p \mathbf{z}_p^t\mathbf{z}_p d_p - \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q} = \boxed{\sum_p d_p - \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q}$$

When $\mathbf{z}_p \in \{0, 1\}$

SLK Concave-Convex Relaxation

Laplacian

Direct convex relaxaton

Concave relaxation (ours)

$$\boxed{\mathrm{tr}(\mathbf{Z}^t\mathbf{L}\mathbf{Z})} = \boxed{\sum_p \mathbf{z}_p^t\mathbf{z}_p d_p - \sum_{p,q} k(\mathbf{x}_p,\mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q} = \boxed{\sum_p d_p - \sum_{p,q} k(\mathbf{x}_p,\mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q}$$

When $\mathbf{z}_p \in \{0,1\}$

$$= \boxed{\sum_p \mathbf{z}_p^t\mathbf{z}_p d_p - \sum_{p,q} k(\mathbf{x}_p,\mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q} \neq \boxed{\sum_p d_p - \sum_{p,q} k(\mathbf{x}_p,\mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q}$$

When $\mathbf{z}_p \in]0,1[$

SLK Concave-Convex Relaxation

Laplacian

Direct convex relaxaton

Concave relaxation (ours)

$$\boxed{\mathrm{tr}(\mathbf{Z}^t\mathbf{L}\mathbf{Z})} = \boxed{\sum_p \mathbf{z}_p^t\mathbf{z}_p d_p - \sum_{p,q} k(\mathbf{x}_p,\mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q} = \boxed{\sum_p d_p - \sum_{p,q} k(\mathbf{x}_p,\mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q}$$

When $\mathbf{z}_p \in \{0,1\}$

$$= \boxed{\underline{\sum_p \mathbf{z}_p^t\mathbf{z}_p d_p} - \sum_{p,q} k(\mathbf{x}_p,\mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q} \neq \boxed{\sum_p d_p - \sum_{p,q} k(\mathbf{x}_p,\mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q}$$

When $\mathbf{z}_p \in]0,1[$

SLK Concave-Convex Relaxation

$$\boxed{\sum_p d_p - \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t \mathbf{z}_q}$$

SLK Concave-Convex Relaxation

Laplacian

$$-\sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t \mathbf{z}_q$$

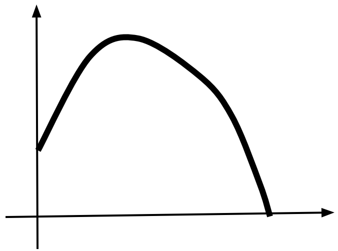
$$\sum_p d_p - \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t \mathbf{z}_q$$

SLK Concave-Convex Relaxation

Laplacian

$$-\sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t\mathbf{z}_q$$

concave

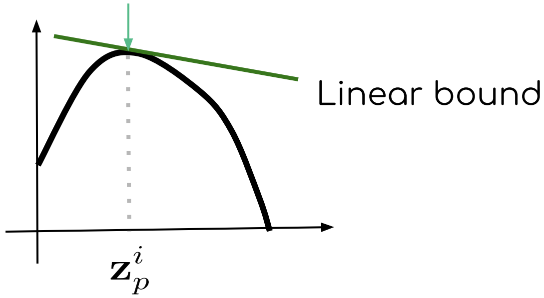


SLK Concave-Convex Relaxation

Laplacian

$$-\sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t \mathbf{z}_q$$

concave

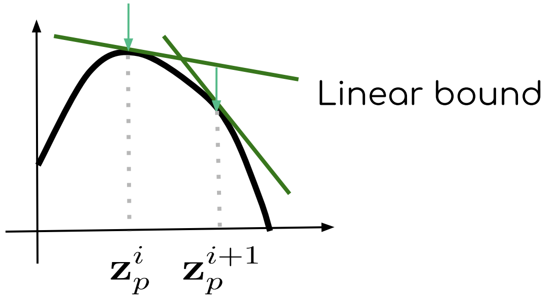


SLK Concave-Convex Relaxation

Laplacian

$$-\sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t \mathbf{z}_q$$

concave



SLK Concave-Convex Relaxation

K-modes

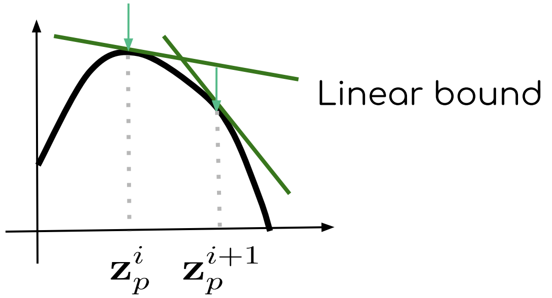
$$-\sum_{p=1}^{N}\sum_{l=1}^{L} z_{p,l}\, k(\mathbf{x}_p, \mathbf{m}_l)$$

+

Laplacian

$$-\lambda \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t \mathbf{z}_q$$

concave



SLK Concave-Convex Relaxation

K-modes

$$-\sum_{p=1}^{N}\sum_{l=1}^{L} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l)$$

$$-\sum_{p=1}^{N} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l) \leq -\sum_{p=1}^{N} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l^i)$$

+

Laplacian

$$-\lambda \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t \mathbf{z}_q$$

concave

SLK Concave-Convex Relaxation

K-modes

Laplacian

$$\sum_{p=1}^{N} \mathbf{z}_p^t \log(\mathbf{z}_p) + \left(- \sum_{p=1}^{N} \sum_{l=1}^{L} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l) \right) + \left(-\lambda \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q) \mathbf{z}_p^t \mathbf{z}_q \right)$$

convex

concave

$$- \sum_{p=1}^{N} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l) \leq - \sum_{p=1}^{N} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l^i)$$

SLK Concave-Convex Relaxation

K-modes

Laplacian

$$\sum_{p=1}^{N} \mathbf{z}_p^t \log(\mathbf{z}_p) \;+\; -\sum_{p=1}^{N}\sum_{l=1}^{L} z_{p,l} k(\mathbf{x}_p, \mathbf{m}_l) \;+\; -\lambda \sum_{p,q} k(\mathbf{x}_p, \mathbf{x}_q)\mathbf{z}_p^t \mathbf{z}_q$$

convex

concave

- ☞ Avoids extra dual variables for constraints: $\mathbf{z}_p \geq 0$
- ☞ Closed- form update duel : $\mathbf{1}^t \mathbf{z}_p = 1$

SLK Proposed bound

Iterative bound:

$$\mathcal{A}_i(\mathbf{Z}) = \sum_{p=1}^{N} \mathbf{z}_p^t(\log(\mathbf{z}_p) - \mathbf{a}_p^i - \lambda \mathbf{b}_p^i)$$

Where,

$$\mathbf{a}_p^i = [a_{p,1}^i, \ldots, a_{p,L}^i]^t,\ a_{p,l}^i = k(\mathbf{x}_p, \mathbf{m}_l^i)$$

$$\mathbf{b}_p^i = [b_{p,1}^i, \ldots, b_{p,L}^i]^t,\ b_{p,l}^i = \sum_q k(\mathbf{x}_p, \mathbf{x}_q) z_{q,l}^i$$

SLK Proposed bound

Iterative bound:

$$\mathcal{A}_i(\mathbf{Z}) = \sum_{p=1}^{N} \mathbf{z}_p^t(\log(\mathbf{z}_p) - \mathbf{a}_p^i - \lambda \mathbf{b}_p^i)$$

Sum of independent function

Where,

$$\mathbf{a}_p^i = [a_{p,1}^i, \ldots, a_{p,L}^i]^t, \; a_{p,l}^i = k(\mathbf{x}_p, \mathbf{m}_l^i)$$

$$\mathbf{b}_p^i = [b_{p,1}^i, \ldots, b_{p,L}^i]^t, \; b_{p,l}^i = \sum_q k(\mathbf{x}_p, \mathbf{x}_q) z_{q,l}^i$$

SLK Proposed bound

Independent Iterative bound:

$$\min_{\mathbf{z}_p \in \nabla_L} \mathbf{z}_p^t(\log(\mathbf{z}_p) - \mathbf{a}_p^i - \lambda \mathbf{b}_p^i), \; \forall p$$

SLK Proposed bound

Independent Iterative bound:

$$\min_{\mathbf{z}_p \in \nabla_L} \mathbf{z}_p^t (\log(\mathbf{z}_p) - \mathbf{a}_p^i - \lambda \mathbf{b}_p^i), \; \forall p$$

☞ KKT conditions get **closed form solution:**

$$\mathbf{z}_p^{i+1} = \frac{\exp(\mathbf{a}_p^i + \lambda \mathbf{b}_p^i)}{\mathbf{1}^t \exp(\mathbf{a}_p^i + \lambda \mathbf{b}_p^i)}$$

SLK-BO

Modes as byproducts of the formulated z-updates:

- In z - updates: $$\max_{\mathbf{y}\in\mathbf{X}} \left[\underbrace{k(\mathbf{y},\mathbf{m}_l^i)}_{\text{proximity}} + \underbrace{\sum_p z_{p,l} k(\mathbf{x}_p,\mathbf{y})}_{\text{density}}\right]$$

SLK-BO

Modes as byproducts of the formulated z-updates:

- In z - updates: $$\max_{\mathbf{y}\in\mathbf{X}} \Big[\underbrace{k(\mathbf{y}, \mathbf{m}_l^i)}_{\text{proximity}} + \underbrace{\sum_p z_{p,l} k(\mathbf{x}_p, \mathbf{y})}_{\text{density}}\Big]$$

- $$\mathbf{z}_p^{i+1} = \frac{\exp(\mathbf{a}_p^i + \lambda \mathbf{b}_p^i)}{\mathbf{1}^t \exp(\mathbf{a}_p^i + \lambda \mathbf{b}_p^i)}$$ take the form of **soft approximation of hard max** as:

$$\mathbf{m}_l^{i+1} = \mathbf{x}_p, \ \text{with } p = \arg\max_q [z_{q,l}]^i$$

SLK-BO

Modes as byproducts of the formulated z-updates:

☞ $\mathbf{z}_p^{i+1} = \frac{\exp(\mathbf{a}_p^i + \lambda \mathbf{b}_p^i)}{\mathbf{1}^t \exp(\mathbf{a}_p^i + \lambda \mathbf{b}_p^i)}$ take the form of **soft approximation of hard max** as:

$$\mathbf{m}_l^{i+1} = \mathbf{x}_p, \ \text{with } p = \arg\max_q [z_{q,l}]^i$$

Linear in *N*

Unlike Mean-shift :

- ☑ No gradient ascent iterates
- ☑ Independent of feature dimensions
- ☑ Arbitrary kernels

SLK Result

NMI/Accuracy

Algorithm	MNIST	MNIST (code)	MNIST (GAN)	LabelMe (Alexnet)	LabelMe (GIST)	YTF	Shuttle	Reuters
K-means	0.53/0.55	0.66/0.74	0.68/0.75	0.81/0.90	0.57/0.69	0.77/0.58	0.22/0.41	0.48/0.73
K-modes	0.56/0.60	0.67/0.75	0.69/0.80	0.81/0.91	0.58/0.68	0.79/0.62	0.33/0.47	0.48/0.72
NCUT	0.74/0.61	0.84/0.81	0.77/0.67	0.81/**0.91**	0.58/0.61	0.74/0.54	0.47/0.46	-
KK-means	0.53/0.55	0.67/0.80	0.69/0.68	0.81/0.90	0.57/0.63	0.71/0.50	0.26/0.40	-
LK	-	-	-	0.81/**0.91**	0.59/0.61	0.77/0.59	-	-
Spectralnet*	-	0.81/0.80	-	-	-	-	-	0.46/0.65
SLK-MS	**0.80**/0.79	0.88/0.95	0.86/**0.94**	**0.83/0.91**	**0.61/0.72**	**0.82/0.65**	0.45/0.70	0.43/0.74
SLK-BO	0.77/**0.80**	**0.89/0.95**	**0.86/0.94**	**0.83/0.91**	**0.61/0.72**	0.80/0.64	**0.51/0.71**	0.43/0.74
K-means	119.9s	**16.8**s	51.6s	11.2s	132.1s	210.1s	1.8s	36.1s
K-modes	90.2s	20.2s	20.3s	7.4s	12.4s	61.0s	**0.5**s	51.6s
NCUT	26.4s	28.2s	**9.3**s	7.4s	10.4s	19.0s	27.4s	-
KK-means	2580.8s	1967.9s	2427.9s	4.6s	17.2s	40.2s	1177.6s	-
LK	-	-	-	33.4s	180.9s	409.0s	-	-
Spectralnet*	-	3600.0s	-	-	-	-	-	9000.0s
SLK-MS	101.2s	82.4s	37.3s	4.7s	37.0s	83.3s	3.8s	**12.5**s
SLK-BO	**14.2**s	23.1s	10.3s	**1.8**s	**7.1**s	**12.4**s	1.3s	53.1s

Time (seconds)

SLK Result

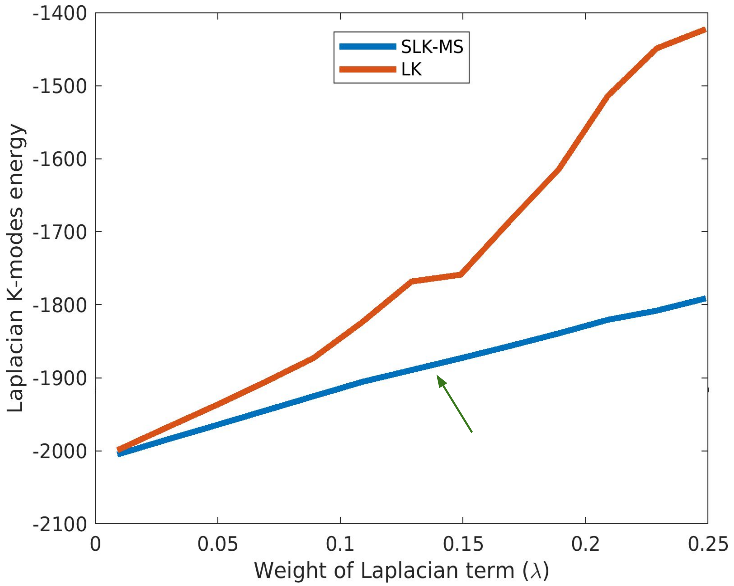
NMI/Accuracy

Algorithm	MNIST	MNIST (code)	MNIST (GAN)	LabelMe (Alexnet)	LabelMe (GIST)	YTF	Shuttle	Reuters
K-means	0.53/0.55	0.66/0.74	0.68/0.75	0.81/0.90	0.57/0.69	0.77/0.58	0.22/0.41	0.48/0.73
K-modes	0.56/0.60	0.67/0.75	0.69/0.80	0.81/0.91	0.58/0.68	0.79/0.62	0.33/0.47	0.48/0.72
NCUT	0.74/0.61	0.84/0.81	0.77/0.67	0.81/**0.91**	0.58/0.61	0.74/0.54	0.47/0.46	-
KK-means	0.53/0.55	0.67/0.80	0.69/0.68	0.81/0.90	0.57/0.63	0.71/0.50	0.26/0.40	-
LK	-	-	-	0.81/**0.91**	0.59/0.61	0.77/0.59	-	-
Spectralnet*	-	0.81/0.80	-	-	-	-	-	0.46/0.65
SLK-MS	**0.80**/0.79	0.88/0.95	0.86/**0.94**	**0.83/0.91**	**0.61/0.72**	**0.82/0.65**	0.45/0.70	0.43/0.74
SLK-BO	0.77/**0.80**	**0.89/0.95**	**0.86/0.94**	**0.83/0.91**	**0.61/0.72**	0.80/0.64	**0.51/0.71**	0.43/0.74
K-means	119.9s	**16.8**s	51.6s	11.2s	132.1s	210.1s	1.8s	36.1s
K-modes	90.2s	20.2s	20.3s	7.4s	12.4s	61.0s	**0.5**s	51.6s
NCUT	26.4s	28.2s	**9.3**s	7.4s	10.4s	19.0s	27.4s	-
KK-means	2580.8s	1967.9s	2427.9s	4.6s	17.2s	40.2s	1177.6s	-
LK	-	-	-	33.4s	180.9s	409.0s	-	-
Spectralnet*	-	3600.0s	-	-	-	-	-	9000.0s
SLK-MS	101.2s	82.4s	37.3s	4.7s	37.0s	83.3s	3.8s	**12.5**s
SLK-BO	**14.2**s	23.1s	10.3s	**1.8**s	**7.1**s	**12.4**s	1.3s	53.1s

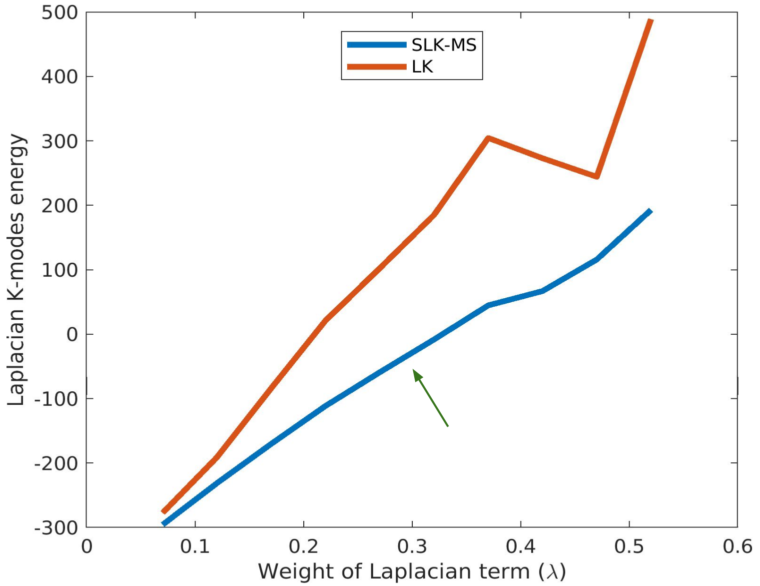
Time (seconds)

SLK Result

Comparison of **optimization quality** w.r.t LK *[Wang and Carreira-Perpiñán 2014]*



LabelMe (Alexnet)



MNIST (small)

Thank you

Code on: https://github.com/imtiazziko/SLK

More at poster session:

Room 210 & 230 AB #96