



**POLYTECHNIQUE
MONTRÉAL**

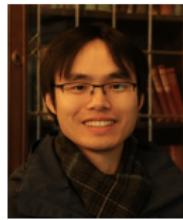
UNIVERSITÉ
D'INGÉNIERIE



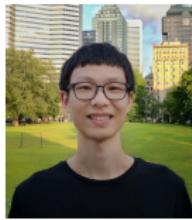
Laplacian Convolutional Representation for Traffic Time Series Imputation

Xinyu Chen

July 11, 2024



Xinyu Chen



Zhanhong Cheng



HanQin Cai



Nicolas Saunier



Lijun Sun

Preprint:

- X. Chen, Z. Cheng, N. Saunier, L. Sun (2022). Laplacian convolutional representation for traffic time series imputation. arXiv preprint arXiv:2212.01529.
https://xinychen.github.io/papers/Laplacian_convolution.pdf

GitHub repository:

- **transdim**: Machine learning for spatiotemporal traffic data imputation and forecasting. (1,000+ stars & 270+ forks on GitHub)
<https://github.com/xinychen/transdim>

Slides:

- <https://xinychen.github.io/slides/LCR.pdf>

Outline

- **Motivation**

- Data-Driven ITS

- Time Series Imputation

- Speed Field Reconstruction

- **Revisit Laplacian Matrix & Circular Convolution**

- Laplacian Matrix

- Laplacian Regularization

- **Circulant Matrix Nuclear Norm Minimization**

- **Laplacian Convolutional Representation**

- Model Description

- Solution Algorithm

- **Experiments**

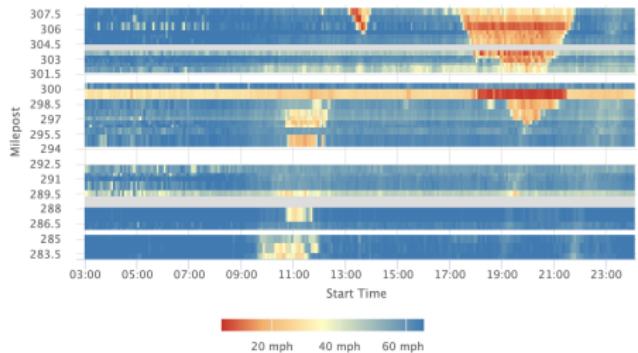
- Univariate Traffic Time Series Imputation

- Speed Field Reconstruction

- **Conclusion**

Motivation

- Portland highway traffic flow data¹



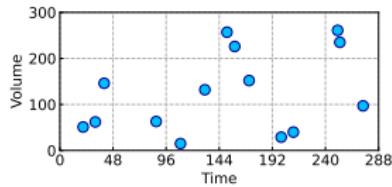
Traffic speed field

Highway network & sensor locations

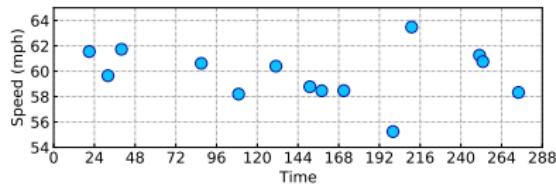
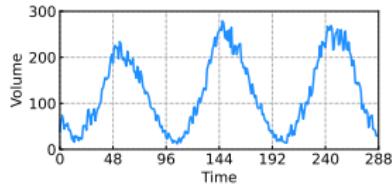
- Speed field $\mathbf{Y} \in \mathbb{R}^{N \times T}$ (N locations & T time steps)
- Speed field shows strong spatial/temporal dependencies

¹<https://portal.its.pdx.edu/home>

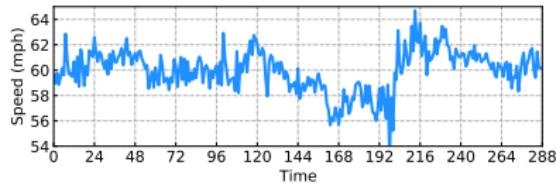
Motivation



↓ Reconstruct
traffic volume?

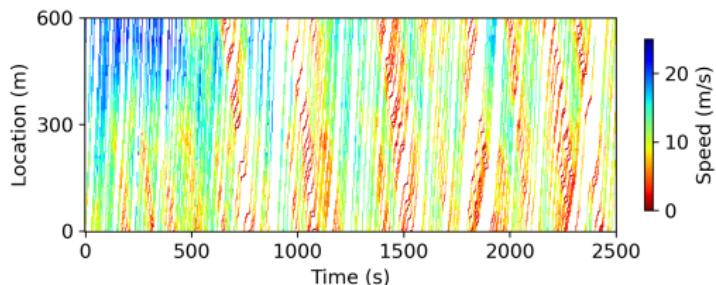


↓ Reconstruct
traffic speed?



- How to utilize the global trends of traffic time series?
- How to produce local consistency of traffic data?

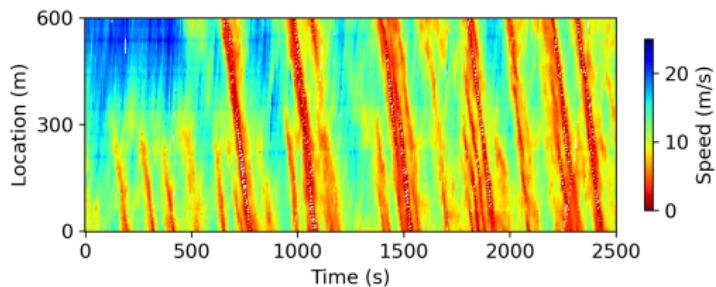
Motivation



200-by-500 matrix
(NGSIM)



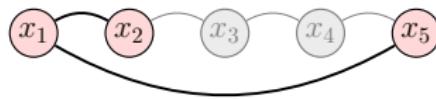
Reconstruct speed field from
20% sparse trajectories?



- How to learn from sparse spatiotemporal data?
- How to characterize spatial/temporal local dependencies?

Revisit Laplacian Matrix & Circular Convolution

- Intuition of (circulant) Laplacian matrix.



Undirected and circulant graph

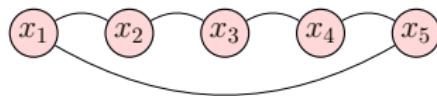
Modeling

$$\mathbf{L} = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}$$

(Circulant) Laplacian matrix

Revisit Laplacian Matrix & Circular Convolution

- Intuition of (circulant) Laplacian matrix.



Undirected and circulant graph

Modeling

$$\mathbf{L} = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}$$

(Circulant) Laplacian matrix

Revisit Laplacian Matrix & Circular Convolution

Reformulate Laplacian regularization with circular convolution.

- Intuition of (circulant) Laplacian matrix.

Undirected and circulant graph

$$\text{Modeling} \longrightarrow \mathbf{L} = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}$$

(Circulant) Laplacian matrix

- Laplacian kernel: $\ell = (2, -1, 0, 0, -1)^\top$.

$$\mathbf{L}\mathbf{x} = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \star \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \ell \star \mathbf{x}$$

where \star denotes the circular convolution.

- Local trend modeling via (Laplacian) temporal regularization:

$$\mathcal{R}(\mathbf{x}) = \frac{1}{2} \|\mathbf{L}\mathbf{x}\|_2^2 = \frac{1}{2} \|\ell \star \mathbf{x}\|_2^2$$

Revisit Laplacian Matrix & Circular Convolution

Reformulate Laplacian regularization with circular convolution.

- Define Laplacian kernel:

$$\ell \triangleq (\underbrace{2\tau}_{\text{degree}}, \underbrace{-1, \dots, -1}_{\tau}, 0, \dots, 0, \underbrace{-1, \dots, -1}_{\tau})^\top \in \mathbb{R}^T$$

for any time series $\mathbf{x} = (x_1, \dots, x_T)^\top \in \mathbb{R}^T$.

- Local trend modeling via (Laplacian) temporal regularization:

$$\mathcal{R}_\tau(\mathbf{x}) = \frac{1}{2} \|\mathbf{L}\mathbf{x}\|_2^2 = \frac{1}{2} \|\ell * \mathbf{x}\|_2^2$$

- Property with discrete Fourier transform (denoted by $\mathcal{F}(\cdot)$)²:

$$\mathcal{R}_\tau(\mathbf{x}) = \frac{1}{2} \|\ell * \mathbf{x}\|_2^2 = \frac{1}{2T} \|\mathcal{F}(\ell) \circ \mathcal{F}(\mathbf{x})\|_2^2$$

²It refers to the Convolution theorem.

Circulant Matrix Nuclear Norm Minimization

Circulant Matrix Nuclear Norm Minimization (CircNNM)

For any partially observed time series $\mathbf{y} \in \mathbb{R}^T$ with observed index set Ω , the optimization problem of CircNNM for reconstructing time series is given by

$$\begin{aligned} & \min_{\mathbf{x}} \|\mathcal{C}(\mathbf{x})\|_* \\ \text{s.t. } & \|\mathcal{P}_\Omega(\mathbf{x} - \mathbf{y})\|_2 \leq \epsilon \end{aligned}$$

where $\mathcal{C} : \mathbb{R}^T \rightarrow \mathbb{R}^{T \times T}$ denotes the circulant operator. $\|\cdot\|_*$ denotes the nuclear norm of matrix, namely, the sum of singular values.

- An important property:

$$\|\mathcal{C}(\mathbf{x})\|_* = \|\mathcal{F}(\mathbf{x})\|_1$$

- CircNNM shows an efficient FFT³ implementation in $\mathcal{O}(T \log T)$ time (Liu'22, Liu & Zhang'23).

³Fast Fourier Transform (FFT).

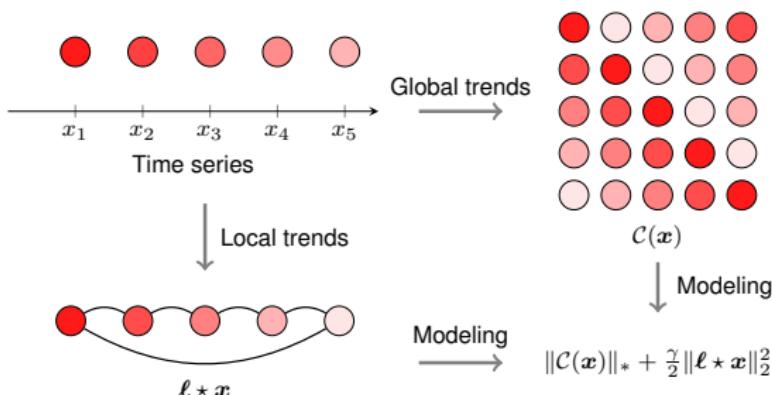
Laplacian Convolutional Representation

Laplacian Convolutional Representation (LCR)

For any partially observed time series $\mathbf{y} \in \mathbb{R}^T$ with observed index set Ω , LCR utilizes circulant matrix and Laplacian kernel to characterize **global and local trends** in time series, respectively, i.e.,

$$\min_{\mathbf{x}} \|\mathcal{C}(\mathbf{x})\|_* + \gamma \cdot \mathcal{R}_{\tau}(\mathbf{x})$$

$$\text{s.t. } \|\mathcal{P}_{\Omega}(\mathbf{x} - \mathbf{y})\|_2 \leq \epsilon$$



Laplacian Convolutional Representation

- LCR model:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathcal{C}(\mathbf{x})\|_* + \gamma \cdot \mathcal{R}_\tau(\mathbf{x}) \\ \text{s.t.} \quad & \|\mathcal{P}_\Omega(\mathbf{x} - \mathbf{y})\|_2 \leq \epsilon \end{aligned}$$

- Augmented Lagrangian function:

$$\mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{w}) = \|\mathcal{C}(\mathbf{x})\|_* + \frac{\gamma}{2} \|\ell * \mathbf{x}\|_2^2 + \frac{\lambda}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 + \langle \mathbf{w}, \mathbf{x} - \mathbf{z} \rangle + \frac{\eta}{2} \|\mathcal{P}_\Omega(\mathbf{z} - \mathbf{y})\|_2^2$$

where $\mathbf{w} \in \mathbb{R}^T$ is the Lagrange multiplier, and $\langle \cdot, \cdot \rangle$ denotes the inner product.

Laplacian Convolutional Representation

- LCR model:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathcal{C}(\mathbf{x})\|_* + \gamma \cdot \mathcal{R}_\tau(\mathbf{x}) \\ \text{s.t.} \quad & \|\mathcal{P}_\Omega(\mathbf{x} - \mathbf{y})\|_2 \leq \epsilon \end{aligned}$$

- Augmented Lagrangian function:

$$\mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{w}) = \|\mathcal{C}(\mathbf{x})\|_* + \frac{\gamma}{2} \|\ell * \mathbf{x}\|_2^2 + \frac{\lambda}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 + \langle \mathbf{w}, \mathbf{x} - \mathbf{z} \rangle + \frac{\eta}{2} \|\mathcal{P}_\Omega(\mathbf{z} - \mathbf{y})\|_2^2$$

where $\mathbf{w} \in \mathbb{R}^T$ is the Lagrange multiplier, and $\langle \cdot, \cdot \rangle$ denotes the inner product.

- The ADMM scheme:

$$\left\{ \begin{array}{l} \mathbf{x} := \arg \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{w}) \\ \mathbf{z} := \arg \min_{\mathbf{z}} \mathcal{L}(\mathbf{x}, \mathbf{z}, \mathbf{w}) \\ = \frac{1}{\lambda + \eta} \mathcal{P}_\Omega(\lambda \mathbf{x} + \mathbf{w} + \eta \mathbf{y}) + \frac{1}{\lambda} \mathcal{P}_\Omega^\perp(\lambda \mathbf{x} + \mathbf{w}) \\ \mathbf{w} := \mathbf{w} + \lambda(\mathbf{x} - \mathbf{z}) \end{array} \right.$$

Laplacian Convolutional Representation

- Optimize \mathbf{x} via FFT (in $\mathcal{O}(T \log T)$ time):

$$\begin{aligned}\mathbf{x} &:= \arg \min_{\mathbf{x}} \|\mathcal{C}(\mathbf{x})\|_* + \frac{\gamma}{2} \|\ell \star \mathbf{x}\|_2^2 + \frac{\lambda}{2} \|\mathbf{x} - \mathbf{z} + \mathbf{w}/\lambda\|_2^2 \\ \implies \hat{\mathbf{x}} &:= \arg \min_{\hat{\mathbf{x}}} \|\hat{\mathbf{x}}\|_1 + \frac{\gamma}{2T} \|\hat{\ell} \circ \hat{\mathbf{x}}\|_2^2 + \frac{\lambda}{2T} \|\hat{\mathbf{x}} - \hat{\mathbf{z}} + \hat{\mathbf{w}}/\lambda\|_2^2\end{aligned}$$

where we introduce $\{\hat{\ell}, \hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{w}}\} \triangleq \mathcal{F}\{\ell, \mathbf{x}, \mathbf{z}, \mathbf{w}\}$ (i.e., FFT).

Laplacian Convolutional Representation

- Optimize \mathbf{x} via FFT (in $\mathcal{O}(T \log T)$ time):

$$\begin{aligned}\mathbf{x} &:= \arg \min_{\mathbf{x}} \|\mathcal{C}(\mathbf{x})\|_* + \frac{\gamma}{2} \|\ell \star \mathbf{x}\|_2^2 + \frac{\lambda}{2} \|\mathbf{x} - \mathbf{z} + \mathbf{w}/\lambda\|_2^2 \\ \implies \hat{\mathbf{x}} &:= \arg \min_{\hat{\mathbf{x}}} \|\hat{\mathbf{x}}\|_1 + \frac{\gamma}{2T} \|\hat{\ell} \circ \hat{\mathbf{x}}\|_2^2 + \frac{\lambda}{2T} \|\hat{\mathbf{x}} - \hat{\mathbf{z}} + \hat{\mathbf{w}}/\lambda\|_2^2\end{aligned}$$

where we introduce $\{\hat{\ell}, \hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{w}}\} \triangleq \mathcal{F}\{\ell, \mathbf{x}, \mathbf{z}, \mathbf{w}\}$ (i.e., FFT).

ℓ_1 -norm Minimization in Complex Space (Liu & Zhang'23)

For any optimization problem in the form of ℓ_1 -norm minimization in complex space:

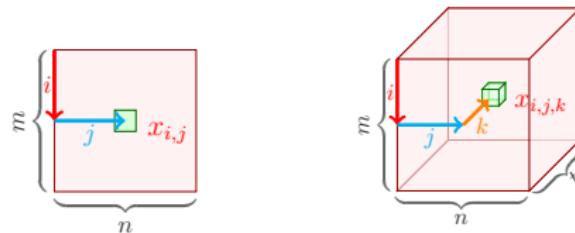
$$\min_{\hat{\mathbf{x}}} \|\hat{\mathbf{x}}\|_1 + \frac{\delta}{2} \|\hat{\mathbf{x}} - \hat{\mathbf{h}}\|_2^2$$

with complex-valued $\hat{\mathbf{x}}, \hat{\mathbf{h}} \in \mathbb{C}^T$ and weight parameter δ , element-wise, the solution is given by

$$\hat{x}_t := \frac{\hat{h}_t}{|\hat{h}_t|} \cdot \max\{0, |\hat{h}_t| - 1/\delta\}, t = 1, \dots, T.$$

Laplacian Convolutional Representation

- What is tensor? $\mathbf{X} \in \mathbb{R}^{m \times n}$ vs. $\mathbf{X} \in \mathbb{R}^{m \times n \times t}$



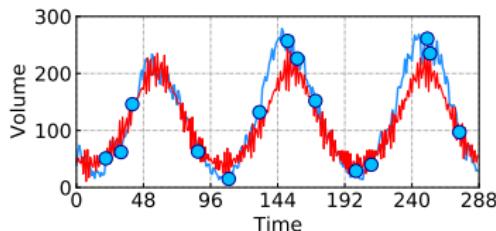
Two-Dimensional LCR (LCR-2D)

For any partially observed time series $\mathbf{Y} \in \mathbb{R}^{N \times T}$ with observed index set Ω , LCR can be formulated as follows,

$$\begin{aligned} \min_{\mathbf{X}} \quad & \|\mathcal{C}(\mathbf{X})\|_* + \frac{\gamma}{2} \|(\ell_s \ell^\top) \star \mathbf{X}\|_F^2 \\ \text{s.t. } & \|\mathcal{P}_\Omega(\mathbf{X} - \mathbf{Y})\|_F \leq \epsilon \end{aligned}$$

where $\mathcal{C} : \mathbb{R}^{N \times T} \rightarrow \mathbb{R}^{N \times N \times T \times T}$ denotes the circulant operator.

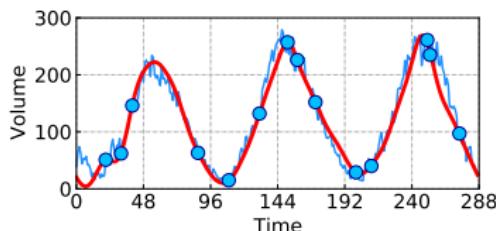
Experiments



CircNNM:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|\mathcal{C}(\boldsymbol{x})\|_* \\ \text{s. t. } \quad & \|\mathcal{P}_\Omega(\boldsymbol{x} - \boldsymbol{y})\|_2 \leq \epsilon \end{aligned}$$

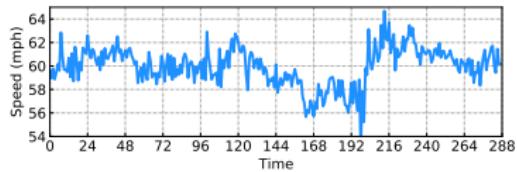
↓ Plus temporal regularization (TR)



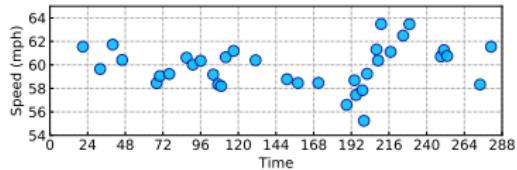
LCR:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|\mathcal{C}(\boldsymbol{x})\|_* + \frac{\gamma}{2} \|\boldsymbol{\ell} * \boldsymbol{x}\|_2^2 \\ \text{s. t. } \quad & \|\mathcal{P}_\Omega(\boldsymbol{x} - \boldsymbol{y})\|_2 \leq \epsilon \end{aligned}$$

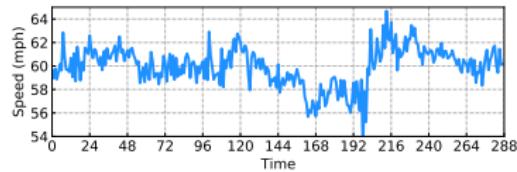
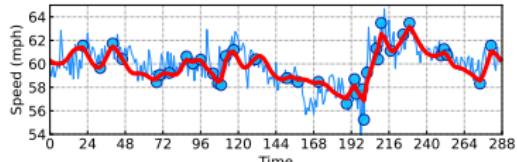
Experiments



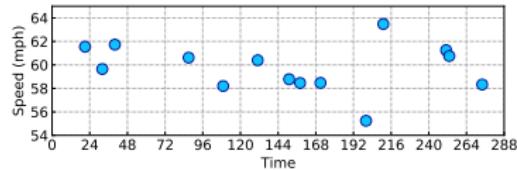
↓ Mask 90% observations



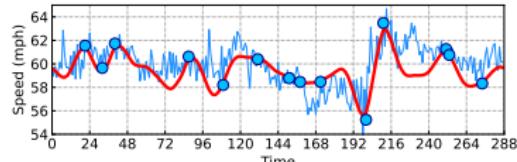
↓ Reconstruct time series



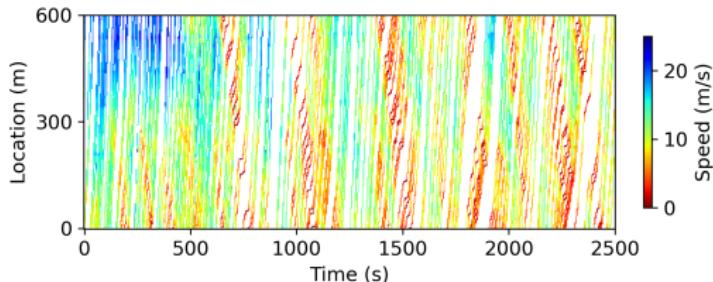
↓ Mask 95% observations



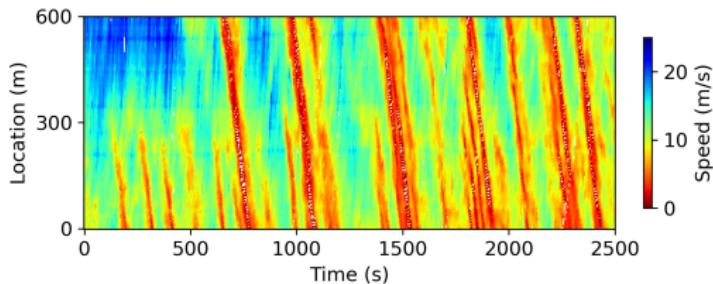
↓ Reconstruct time series



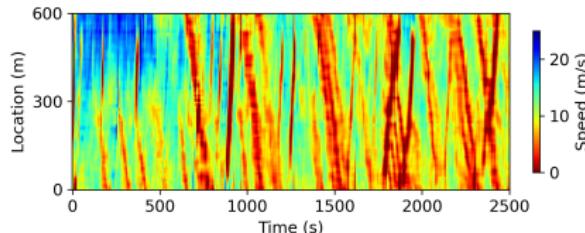
Experiments



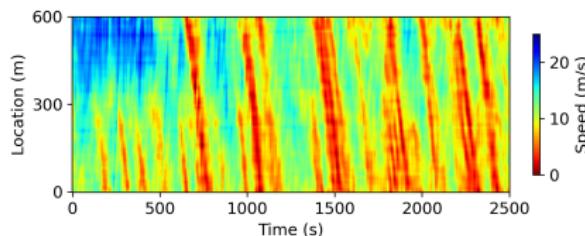
200-by-500 matrix
(NGSIM) \downarrow Reconstruct speed field from
20% sparse trajectories?



Experiments



↓ Plus CircNNM



TR:

$$\min_{\mathbf{X}} \mathcal{R}_\tau(\mathbf{X})$$

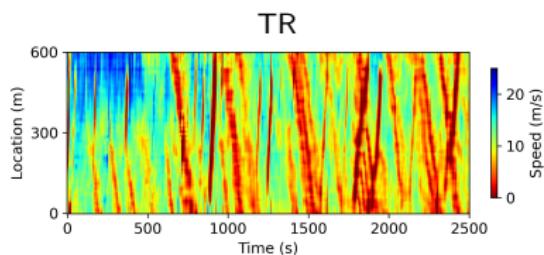
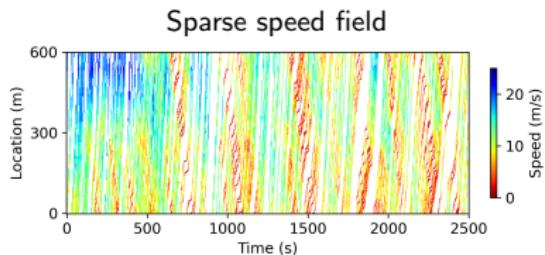
$$\text{s. t. } \|\mathcal{P}_\Omega(\mathbf{X} - \mathbf{Y})\|_F \leq \epsilon$$

LCR:

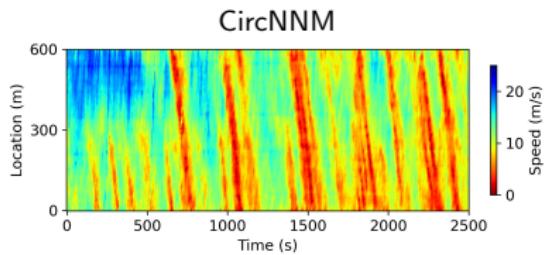
$$\min_{\mathbf{X}} \|\mathcal{C}(\mathbf{X})\|_* + \gamma \cdot \mathcal{R}_\tau(\mathbf{X})$$

$$\text{s. t. } \|\mathcal{P}_\Omega(\mathbf{X} - \mathbf{Y})\|_F \leq \epsilon$$

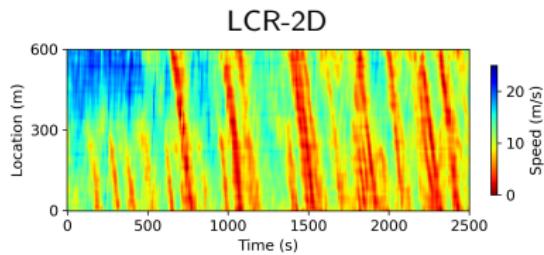
Experiments



MAPE = 46.94% & RMSE = 4.34 m/s



MAPE = 43.51% & RMSE = 1.65 m/s



MAPE = 41.29% & RMSE = 1.55 m/s

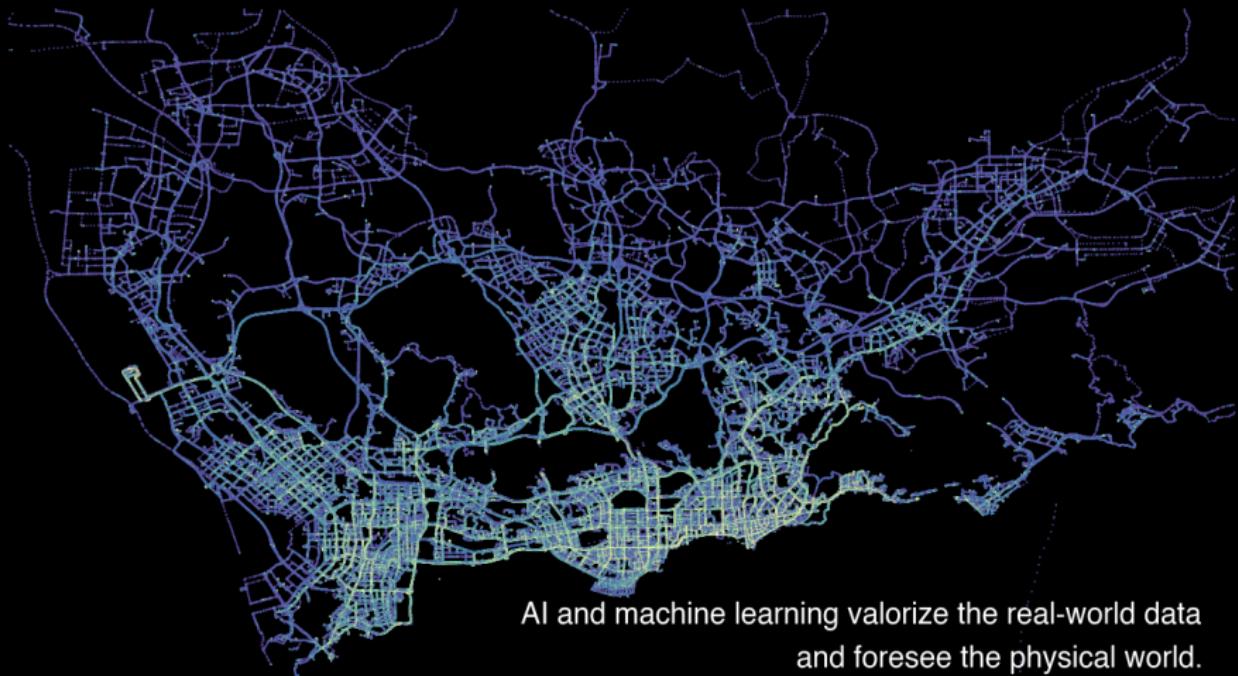
Conclusion

- **(Starting point)** How to impute traffic time series?
 - ✓ Low-rank models ✓ Temporal regularization
- **(Solution)** Time series trend modeling in the low-rank framework?
 - Global time series trend modeling (low-rank model):
$$\min_{\mathbf{x}} \|\mathcal{C}(\mathbf{x})\|_*$$
$$\text{s. t. } \|\mathcal{P}_\Omega(\mathbf{x} - \mathbf{y})\|_2 \leq \epsilon$$
 - Local time series trend modeling (temporal regularization):
$$\mathcal{R}_\tau(\mathbf{x}) = \frac{1}{2} \|\boldsymbol{\ell} \star \mathbf{x}\|_2^2$$
- **(Highlight)** A unified framework with the **FFT** implementation.

References

A short list:

- [Liu'22] G. Liu (2022). Time series forecasting via learning convolutionally low-rank models. *IEEE Transactions on Information Theory*, 68(5): 3362–3380.
- [Liu & Zhang'23] G. Liu and W. Zhang (2023). Recovery of future data via convolution nuclear norm minimization. *IEEE Transactions on Information Theory*, 69(1): 650–665.



AI and machine learning valorize the real-world data
and foresee the physical world.

Source: <https://spatiotemporal-data.github.io>



POLYTECHNIQUE
MONTRÉAL

UNIVERSITÉ
D'INGÉNIERIE



Thanks for your attention!

Any Questions?

About me:

- 🏠 Homepage: <https://xinychen.github.io>
- /github/ GitHub: <https://github.com/xinychen>
- ✉️ How to reach me: chenxy346@gmail.com

