Sparse Subspace Tracking in High Dimensions

Le Trung Thanh[1], Karim Abed-Meraim[1], Adel Hafiane[1], Nguyen Linh Trung[2]

[1] Laboratoire PRISME, Université d'Orléans, INSA CVL, Orléans 45072, France

[2] AVITECH Institute, Vietnam National University, Hanoi, Vietnam

ICASSP 2022 Singapore

Sparse Subspace Tracking

Data Model: $\mathbf{x}_t = \mathbf{A}\mathbf{w}_t + \mathbf{n}_t, \quad 1 \le t \le T$

- T: number of snapshots
- $\mathbf{x}_t \in \mathbb{R}^{n\times 1}$: n-dimensional observation vector
- $\mathbf{A} \in \mathbb{R}^{n\times r}$: basis sparse matrix
- $\mathbf{w}_t \in \mathbb{R}^{r\times 1}$: coefficient vector
- $\mathbf{n}_t \in \mathbb{R}^{n\times 1}$: additive random noise

Objective: On the arrival of new data $\mathbf{x}_t$, we want to update $\mathbf{A}$ or $\text{span}(\mathbf{A})$.

Related Works

Method	High dimension?	Main features
SPCA (ICML 2015)	✓	+ Row-sparsity + Row truncation + QR
OIST (ITW 2016)	✓	+ Rank-1 subspace + Oja method + soft-thresholding
L1-PAST (IEEE TSP 2016)	✗	+ PAST + l1-norm penalty
OVBSL (Elsevier SP 2017)	✗	+ Bayesian inference
SS/DS-OPAST (EUSIPCO 2017)	✗	+ OPAST + l1-norm penalty
SS/GSS-FAPI (Elsevier SP 2020)	✗	+ FAPI + Given rotations

Proposed Method

OPIT Algorithm:

- **Power Iteration**: At ℓ-th iteration
  - Step 1: $\mathbf{S}_\ell \leftarrow \mathbf{C}_t\mathbf{U}_{\ell-1}$
  - Step 2: $\mathbf{U}_\ell \leftarrow \text{QR}(\mathbf{S}_\ell)$
- **OPIT**: $\ell \leftarrow t$
  - Step 1: $\mathbf{S}_t \leftarrow \lambda\mathbf{S}_{t-1}\mathbf{E}_{t-1} + \mathbf{x}_t\mathbf{z}_t^\top$, where
    - $\mathbf{z}_t = \mathbf{U}_{t-1}^\top\mathbf{x}_t$ and $\mathbf{E}_t = \mathbf{U}_{t-2}^\top\mathbf{U}_{t-1}$
    - $0 < \lambda \le 1$: Forgetting factor
  - Step 2:
    - $\hat{\mathbf{S}}_t \leftarrow \tau(\mathbf{S}_t, k)$
    - $\mathbf{U}_t \leftarrow \text{QR}(\hat{\mathbf{S}}_t)$ or $\hat{\mathbf{S}}_t/\|\hat{\mathbf{S}}_t\|_2$

Convergence Analysis:

$$\text{If } \begin{cases} t \ge \dfrac{c_\delta}{W\epsilon^2}\left(\sqrt{r} + (2\sigma_n/\sigma_x + \sigma_n^2/\sigma_x^2)\sqrt{n}\right)^2, \\ \tan\theta(\mathbf{A},\mathbf{U}_0) \le \dfrac{\sigma_x^2+\sigma_n^2}{(1+\sqrt{r}(1+\sqrt{2}))\sigma_x^2 - (2+\sqrt{2})\sigma_n^2} \end{cases}$$

with a small predefined error ϵ and $c_\delta = C\sqrt{\log 2/\delta}$, $0 < \delta \ll 1$, then $\sin(\mathbf{A},\mathbf{U}_t) \le \epsilon$ under mild conditions.

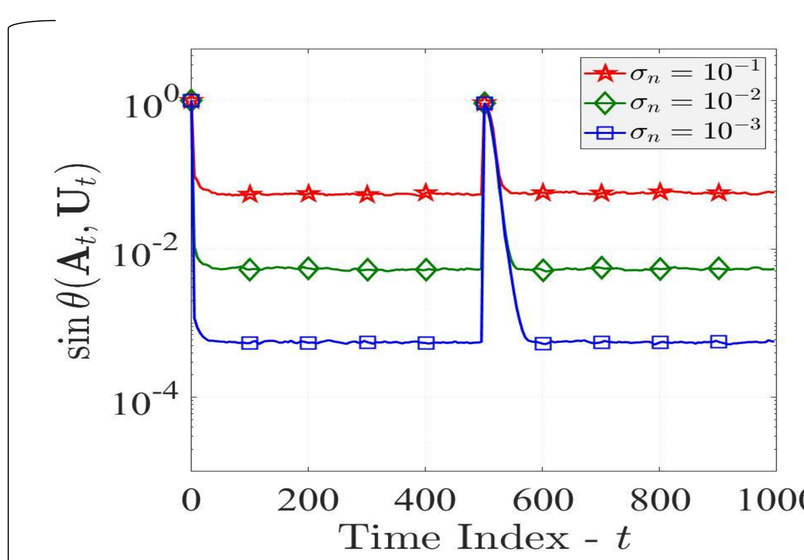
Experiments

Setup:

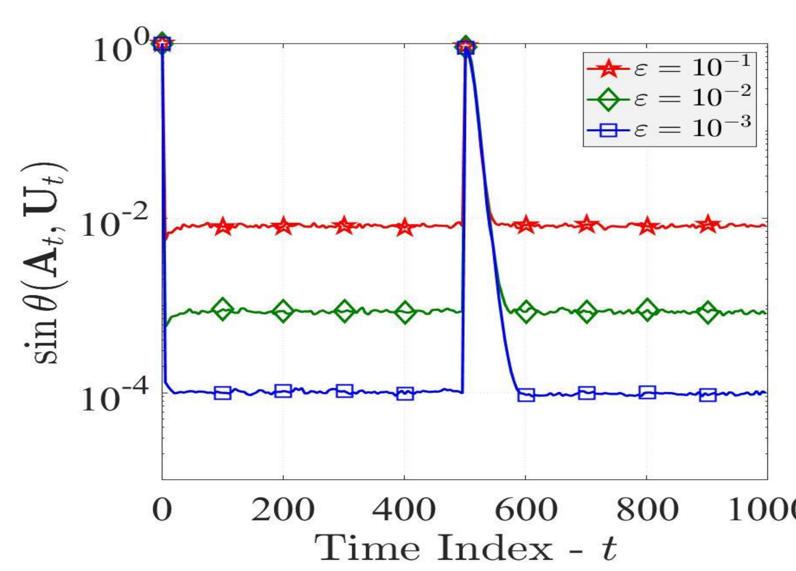
$$\begin{cases} \mathbf{x}_t = \mathbf{A}_t\mathbf{w}_t + \sigma_n\mathbf{n}_t \\ \mathbf{A}_t = \mathbf{\Omega} \circledast (\mathbf{A}_{t-1} + \varepsilon\mathbf{N}_t) \end{cases}$$

- $\mathbf{w}_t \in \mathcal{N}(0, \mathbf{I}_r)$
- $\mathbf{n}_t \in \mathcal{N}(0, \mathbf{I}_n)$ with $\sigma_n > 0$: noise level
- $\mathbf{\Omega}$: Bernoulli with probability $\rho > 0$
- $\mathbf{N}_t$: Normalized Gaussian matrix
- $\varepsilon > 0$: time-varying factor

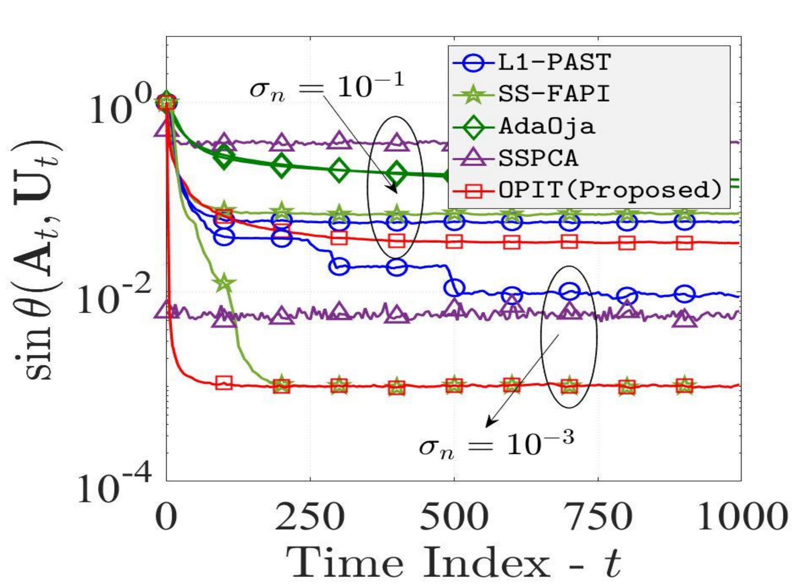
Results:



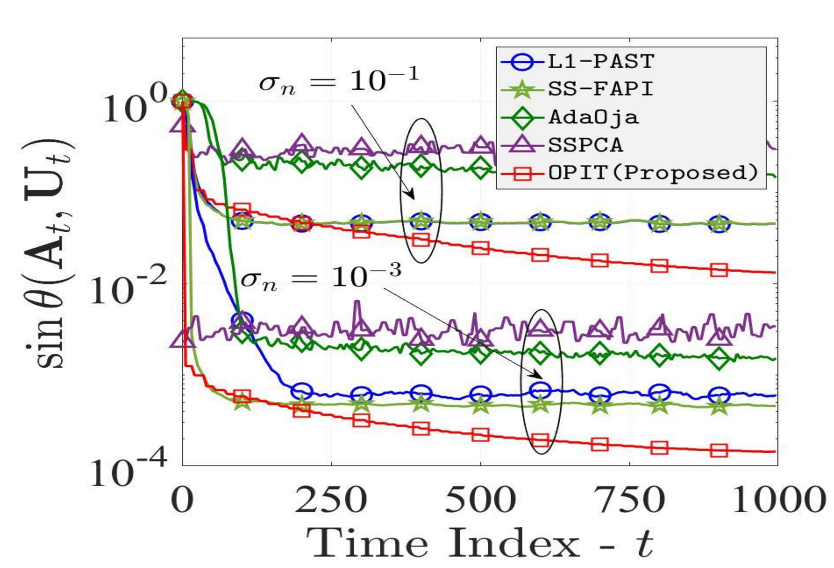
a) Effect of σ_n : $n = 100, r = 5, \rho = 90\%, \varepsilon = 10^{-3}$



b) Effect of ε : $n = 100, r = 5, \rho = 90\%, \sigma_n = 10^{-3}$



c) Classical regime: $n = 50, r = 2, \rho = 90\%, \varepsilon = 10^{-3}$

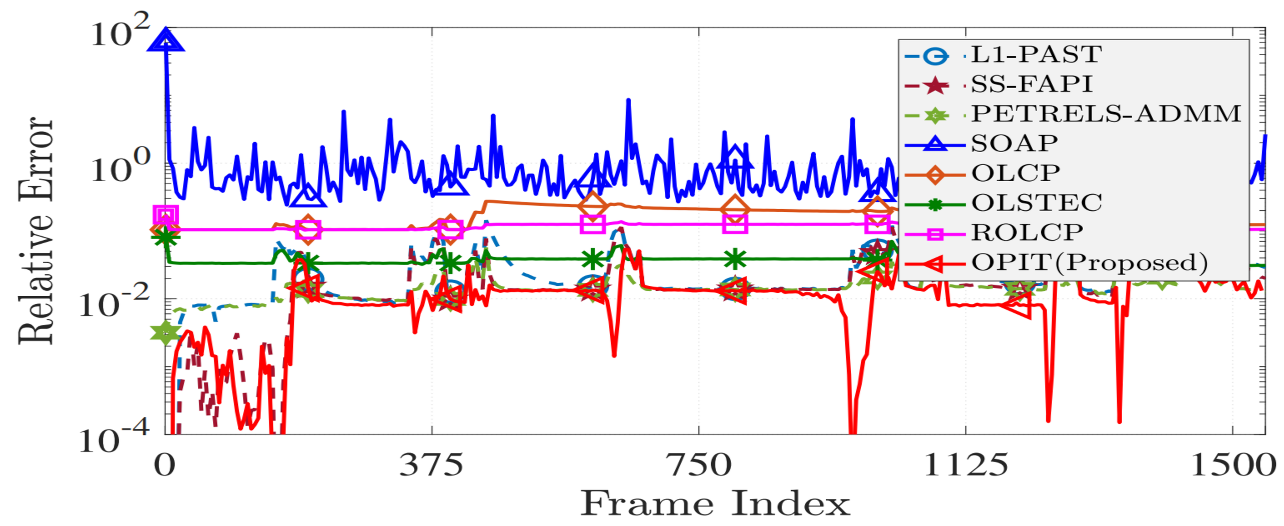


d) High dimension: $n = 10000, r = 10, \rho = 90\%, \varepsilon = 10^{-3}$

Video Tracking:

a) Lobby

b) Hall



Dataset	"Lobby"		"Hall"	
Tensor size	128 × 160 × 1546		174 × 144 × 3584	
Matrix size	20480 × 1546		25056 × 3584	
Evaluation metrics	time(s)	error	time(s)	error
SOAP	14.29	0.842	21.72	0.989
OLCP	10.50	0.161	19.98	0.154
OLSTEC	44.25	0.037	92.82	0.041
ROLCP	4.32	0.114	10.74	0.120
PETRELS-ADMM	118.4	0.015	305.5	0.018
ℓ_1-PAST	14.11	0.031	33.73	0.101
SS-FAPI	12.99	0.023	32.72	0.100
OPIT ($W = 1$)	16.32	0.013	50.78	0.056
OPIT ($W = \lfloor\log(IJ)\rceil$)	1.89	0.021	5.62	0.086

Conclusion

- Proposed a novel sparse subspace tracking algorithm called OPIT.
- Provided a theoretical result on the convergence of OPIT
- Demonstrated the effectiveness of OPIT with several experiments.