Efficient Spatio-Temporal Sampling via Low-Rank Tensor Sketching

Rose Yu, Sanjay Purushotham, Yan Liu
{qiyu, spurusho, yanliu.cs}@usc.edu
USC UNIVERSITY OF SOUTHERN CALIFORNIA

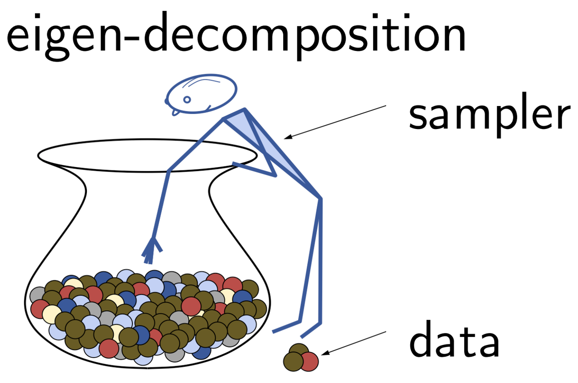
1. Abstract

- We formulate spatio-temporal sampling task as tensor sketching problem.
- We generalize sparse subspace embedding to low-rank tensor domain.
- Our algorithm achieves accurate predictions with significant speed-up in social media and climate applications.

2. Spatio-Temporal Sampling

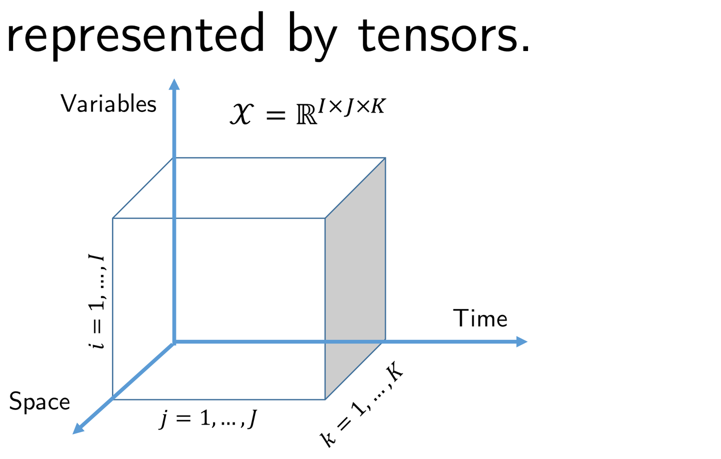
Definition:
Identify important *locations* or *time stamps* and extract samples from them with the merit of computational efficiency

Challenge:
- Voronoi diagram: no theoretical guanreee
- Sequential sampling [Krause 2008]:require submodular assumption
- Determinal point process [Kulesza 2012]: requires expensive eigen-decomposition

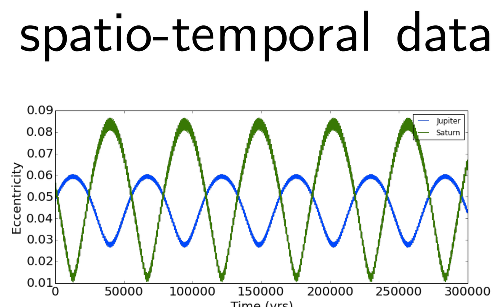


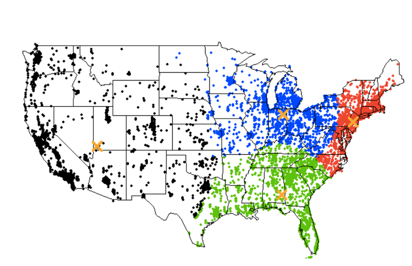
3. Tensor Representation

- Multivariate spatio-temporal data can be naturally represented by tensors.

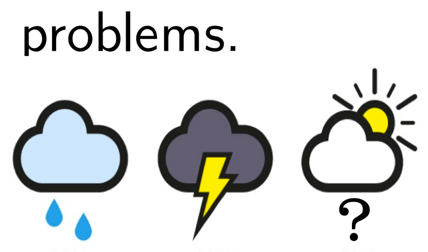


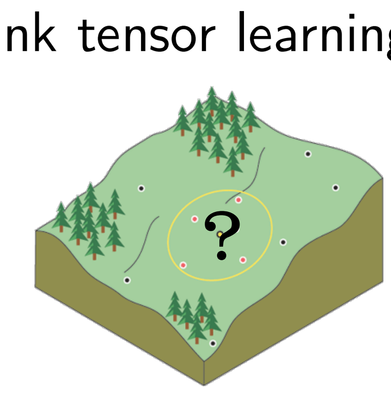
- Low-rank tensor can capture structures in spatio-temporal data [Yu 2014, 2015].





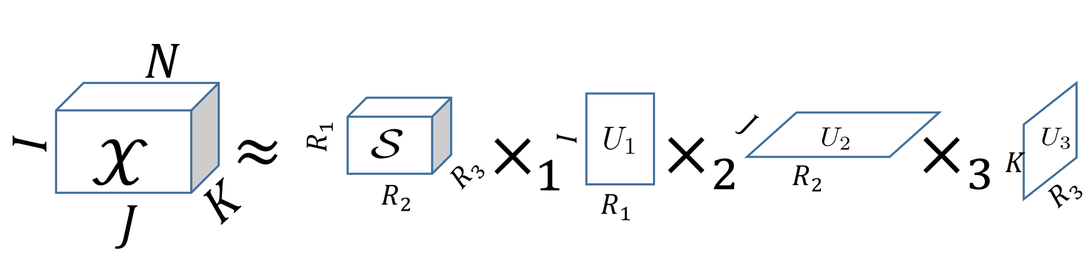
- Many spatio-temporal analysis tasks can be formulated as low-rank tensor learning problems.



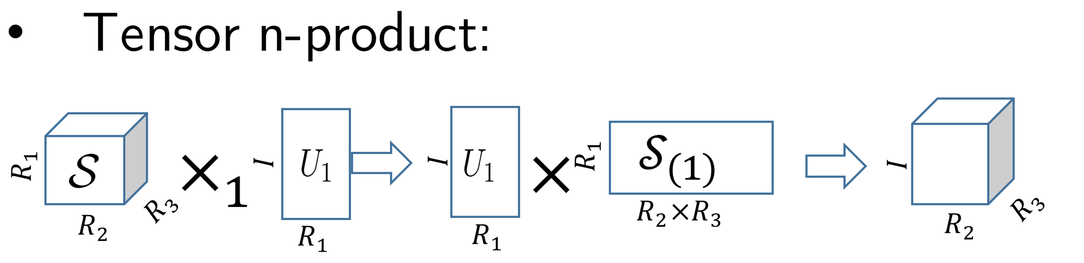


Preliminary :

- Tucker decomposition:



- Tensor n-product:

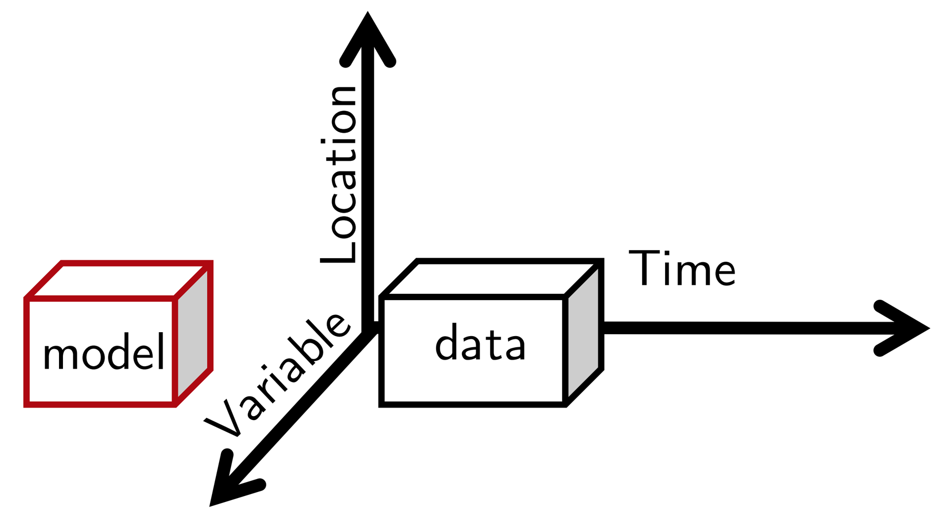


4. Forecasting Formulation

Concatenate historical measurements of L lags $[\mathcal{X}_{t-1,:,m}, \mathcal{X}_{t-2,:,m} \ldots \mathcal{X}_{t-L,:,m}] = \bar{\mathcal{X}}_{t,:,m} \in \mathbb{R}^{T \times (PL) \times M}$

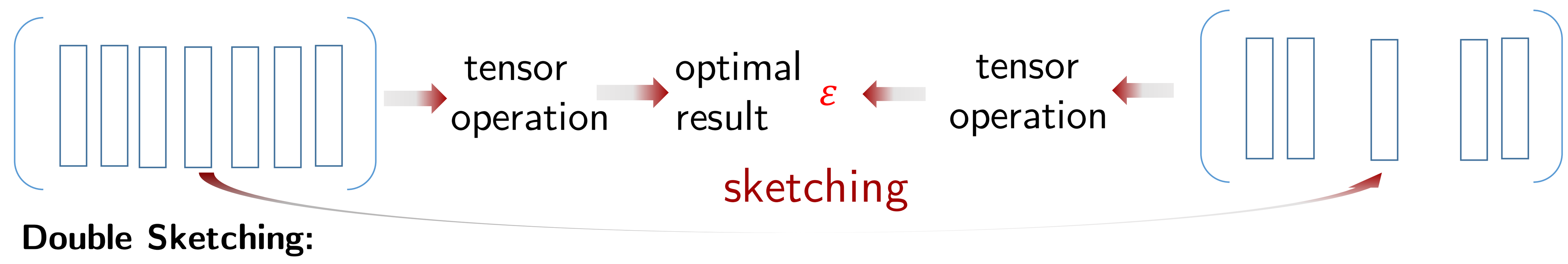
Goal: Learn a model tensor $\mathcal{W} \in \mathbb{R}^{(PL) \times P \times M}$

$$\widehat{\mathcal{W}} = \underset{\mathcal{W}}{\operatorname{argmin}} \left\{ \|\widehat{\mathcal{X}} - \mathcal{X}\|_F^2 + \mu \sum_{m=1}^{M} \operatorname{trace}(\widehat{\mathcal{X}}_{:,:,m} L \widehat{\mathcal{X}}_{:,:,m}^T) \right\}$$
$$\text{subject to} \quad \operatorname{rank}(\mathcal{W}) \le R$$
$$\bar{\mathcal{X}}_{t,:,m} \mathcal{W}_{:,p,m} = \widehat{\mathcal{X}}_{t,p,m}$$



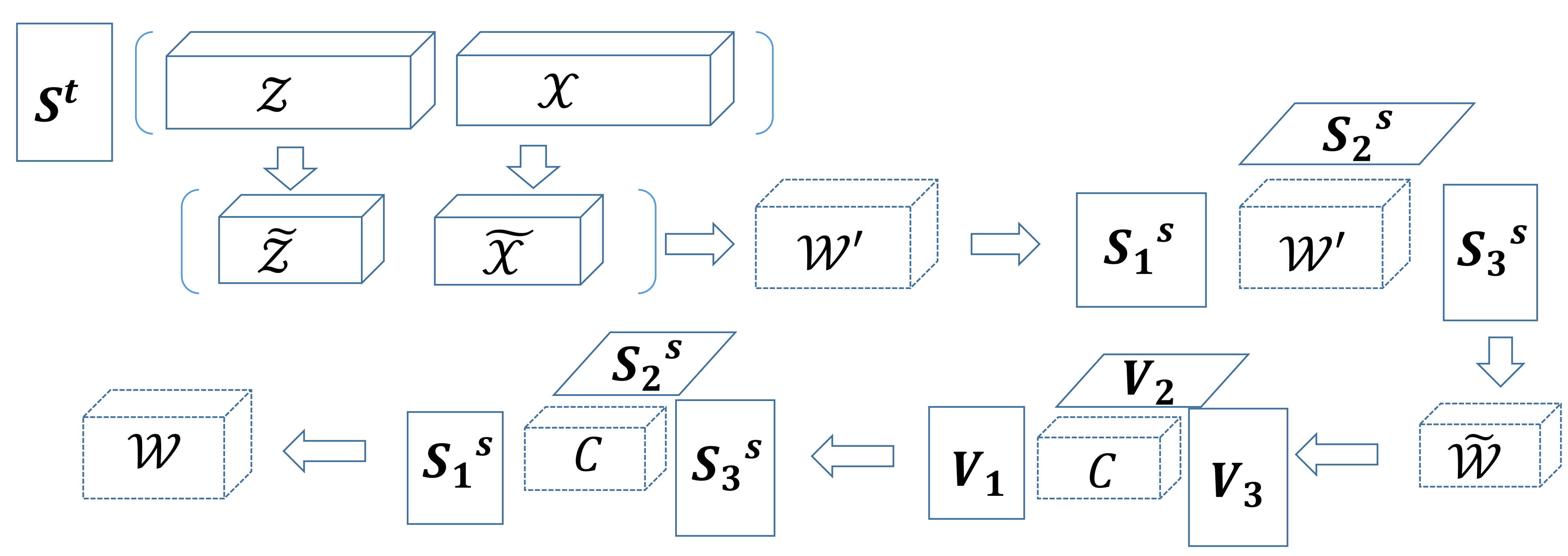
$$\widehat{\mathcal{W}} = \underset{\mathcal{W}}{\operatorname{argmin}} \left\{ \sum_m \|\mathcal{Z}_{:,:,m} \mathcal{W}_{:,:,m} - \mathcal{X}_{:,:,m}\|_F^2 \right\}$$
$$\text{subject to} \quad \operatorname{rank}(\mathcal{W}) \le R$$

5. Subsampled Randomized Low-rank Tensor Learning



Double Sketching:
1. Sketching data tensor along time dimension.2. Sketching model tensor along locations and variables.

L2-Sparse subspace embedding [Clark && Woodruff 2013]:
for each column j, uniformly pick a row i ∈ {1,2,⋯M} and assign {−1, 1} with equal probability to $\boldsymbol{S}_{i,j}$.

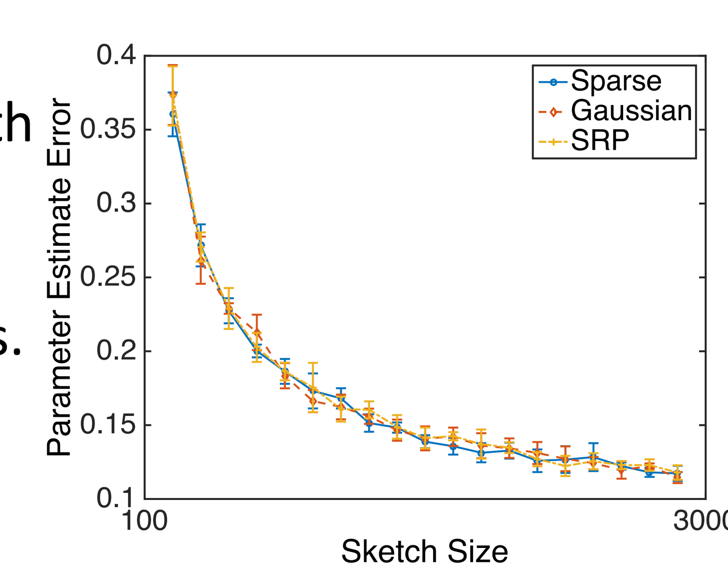


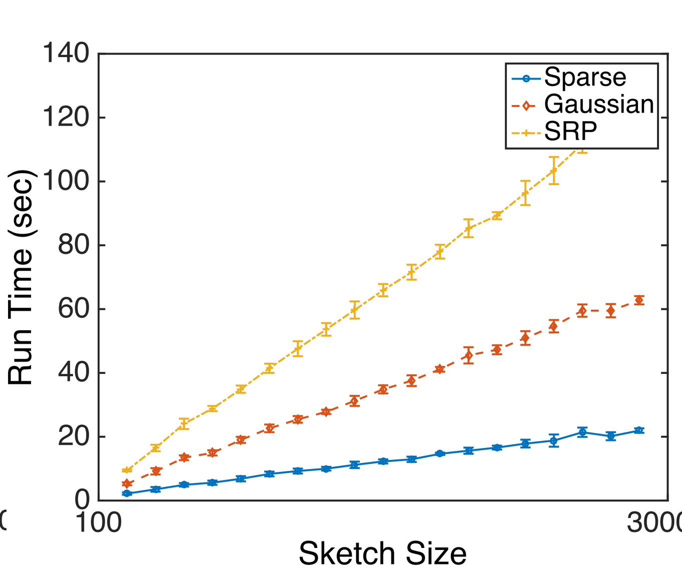
Theoretical Analsysis

Lemma 1 (Adapted from [5]): *For any* $0 < \boldsymbol{\delta} < \mathbf{1}$, *and for* $\boldsymbol{S}^t$ *a* $\boldsymbol{l}_2$ *sparse -subspace embedding matrix with* K = O(P2M2/δε2) *rows, then with probability* $\mathbf{1} - \boldsymbol{\delta}$, *we can achieve* $(\mathbf{1} + \boldsymbol{\varepsilon})$*-approximation for tensor least square in* O(nnz($\mathcal{X}$)) *time.*
Let $\boldsymbol{S_n}^s$, n = 1, 2, 3 *be a sparse* $\boldsymbol{l}_2$*-subspace embedding matrix with* K = O(R/ε) *rows, then with high probability, we can achieve* (1 + ε)*-approximation for low-rank tensor approximation in* O(nnz($\mathcal{W}'$)) + poly(P + Q + M)poly(R/ε) *time,*

6. Experiments

Synthetic: 30000 time stamps with VAR(2) model, parameter tensor W ∈ R $^{30 \times 60 \times 20}$.
Repeat the procedure for 10 times.

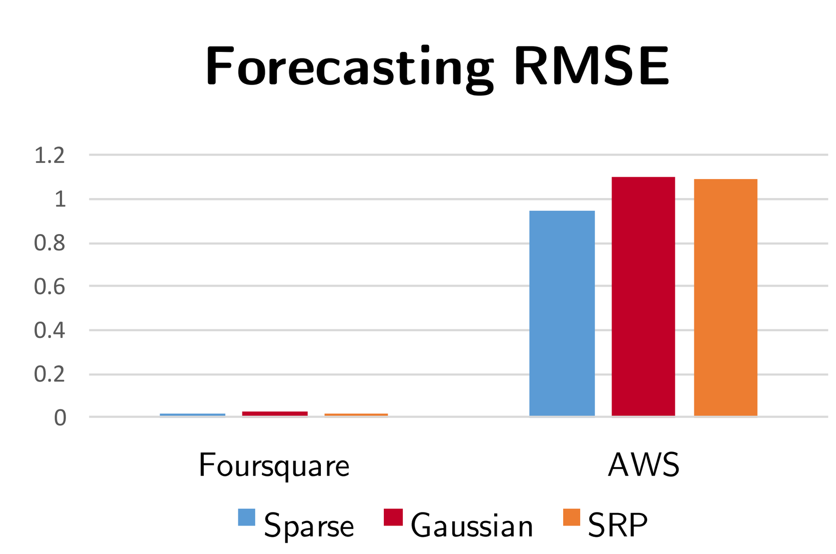




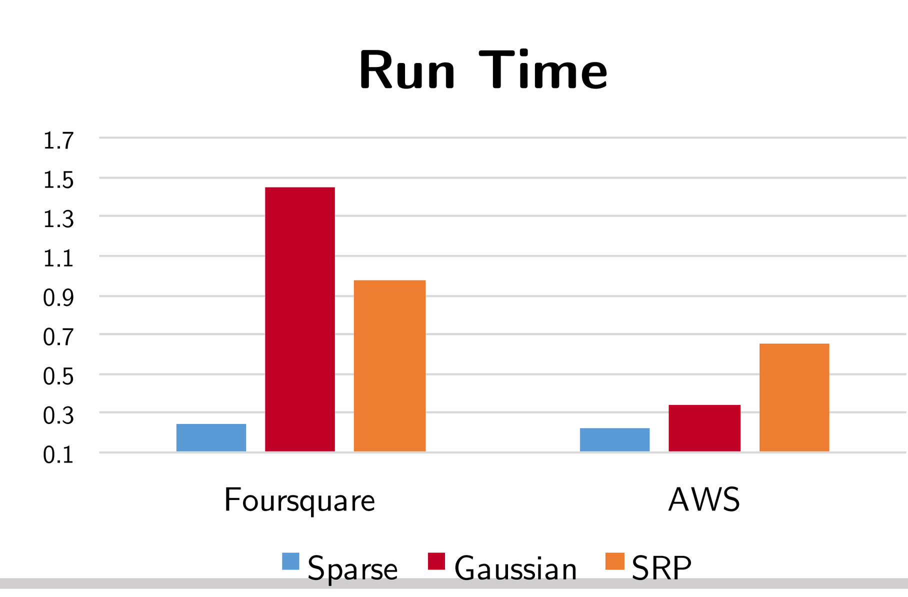
Real-World Datasets :
- **Foursquare**: 121 user check-ins, 15 categories of business venues, 1200 time intervals.
- **AWS**: 153 weather stations measurements, 4 climate variables, 76 time stamps.

Settings: 90 % training data on both datasets for VAR model with different lags and average run time.

Forecasting RMSE



Run Time



7. Reference

[1] **Rose Yu**, Dehua Cheng, Yan Liu. Accelerated Online Low-Rank Tensor Learning for Multivariate Spatio-Temporal Streams *International Conference on Machine Learning* (ICML), 2015
[2] **Rose Yu***, Mohammad Taha Bahadori*, Yan Liu. (*Equal Contributions) Fast Multivariate Spatio-temporal Analysis via Low Rank Tensor Learning *Advances in Neural Information Processing Systems* (NIPS), 2014, Spotlight
[3] Bonilla, Edwin V., Kian M. Chai, and Christopher Williams. "Multi-task Gaussian process prediction." *Advances in neural information processing systems*. 2007.
[4] A. Kulesza and B. Taskar. Determinantal point processes for machine learning. *Machine Learning*, 5(2-3):123–286, 2012.
[5] D. P. Woodruff. Sketching as a tool for numerical linear algebra. *Theoretical Computer Sci- ence*, 10(1-2):1–157, 2014.