

# Shashanka Ubaru

201 E 24th Street, POB 3.134  
University of Texas, Austin, TX, USA 78712  
✉ [Shashanka.Ubaru@ibm.com](mailto:Shashanka.Ubaru@ibm.com)  
📁 [shashankaubaru.github.io](https://github.com/shashankaubaru)

## Research Interests

Machine learning, numerical linear and tensor algebra, quantum computing, coding theory applications, approximation theory and algorithms.

## Positions

- 2018 - present **IBM Research, Yorktown Heights/Austin, USA.**  
*Senior Research Scientist* May, 2022 - present  
*Research Staff Member* Jan, 2020 - May, 2022  
*Goldstine Postdoctoral Fellow* Aug, 2018 - Dec, 2019
- 2022 - present **Oden Institute, University of Texas at Austin, TX, USA.**  
*Visiting Researcher* June, 2022 - present
- 2013 - 2018 **University of Minnesota, Minneapolis, MN, USA.**  
*Research Assistant, Department of Computer Science.* Sep, 2013 - May, 2018
- 2016 & 2017 **Lawrence Berkeley National Laboratory, Berkeley, CA, USA.**  
*Research Intern, Neural Systems and Data Science Lab* June - Aug, 2017  
*Research Intern, Scientific Data Management group* May - Aug, 2016
- 2013 **Seagate Technology, Shakopee, MN, USA.**  
*Signal Processing Intern* May - Aug, 2013
- 2011 - 2012 **Raman Research Institute, Bangalore, India.**  
*Visiting Research Student* June, 2011 - Aug, 2012

## Education

- 2012-18 **University of Minnesota, Minneapolis, MN, USA.**  
*Ph.D. Computer Science* May 2018  
Advisor: Yousef Saad  
Thesis: *Algorithmic advances in learning from large dimensional matrices and scientific data*
- M.S. Computer Science* October 2015  
*M.S. Electrical Engineering* November 2014  
Advisors: Yousef Saad and Arya Mazumdar  
Thesis: *Randomized techniques for matrix decomposition and estimating the approximate rank of a matrix*
- 2008-12 **M.S. Ramaiah Institute of Technology, Bangalore, India.**  
*B.Eng. Electronics and Communication*

## Publications

### Journal articles

- [9] *Dynamic graph and polynomial chaos based models for contact tracing data analysis and optimal testing prescription.*  
**S. Ubaru**, G. Cohen, and L. Horesh.  
Journal of Biomedical Informatics, Novel Informatics Approaches to COVID-19 Research, 122, 103901, 2021.
- [8] *Spectrum-Adapted Polynomial Approximation for Matrix Functions with Applications in Graph Signal Processing.*  
L. Fan, D. Shuman, **S. Ubaru**, and Y. Saad.  
Algorithms, Special Issue: Efficient Graph Algorithms in Machine Learning, 13(11), 295, 2020.

- [7] *Sampling and multilevel coarsening algorithms for fast matrix approximations.*  
S. Ubaru and Y. Saad  
Numerical Linear Algebra with Applications 26.3 : e2234, 2019.
- [6] *Fast estimation of  $\text{tr}(f(A))$  via Stochastic Lanczos Quadrature.*  
S. Ubaru, J. Chen, and Y. Saad  
SIAM Journal on Matrix Analysis and Applications (SIMAX), 38(4), 1075–1099, 2017.
- [5] *Low rank approximation and decomposition of large matrices using error correcting codes.*  
S. Ubaru, A. Mazumdar, and Y. Saad  
IEEE Transactions on Information Theory, 63(9), 5544–5558, 2017.
- [4] *Formation enthalpies for transition metal alloys using machine learning.*  
S. Ubaru, A. Miedlar, Y. Saad, and J R. Chelikowsky  
Physical Review B, (Vol.95, No.21), 2017.
- [3] *Fast estimation of approximate matrix ranks using spectral densities.*  
S. Ubaru, Y. Saad, and A.-K. Seghouane  
Neural Computation, 29(5):1317–1351, 2017.
- [2] *Improving the Incoherence of a Learned Dictionary via Rank Shrinkage.*  
S. Ubaru, A.-K. Seghouane, and Y. Saad  
Neural Computation, 29(1):263–285, 2017.
- [1] *Displaying gray scales by cross pairing select and data voltages in multi-line addressed LCD.*  
S. Ubaru and T.N. Ruckmongathan  
IEEE Journal of Display Technology, 8(11), 669–677, 2012.

#### Conference proceedings

- [15] *Efficient Scaling of Dynamic Graph Neural Networks.*  
V. Chakaravarthy, S. Pandian, S. Raje, Y. Sabharwal, T. Suzumura, and S. Ubaru.  
Supercomputing (SC21), 2021.
- [14] *Analysis of stochastic Lanczos quadrature for spectrum approximation.*  
T. Chen, T. Trogdon, and S. Ubaru.  
International Conference on Machine Learning (ICML), 2021. *Long presentation*
- [13] *Projection techniques to update the truncated SVD of evolving matrices.*  
V. Kalantzis, G. Kollias, S. Ubaru, A. Nikolakopoulos, L. Horesh, and K.L. Clarkson.  
International Conference on Machine Learning (ICML), 2021.
- [12] *Sparse graph based sketching for fast numerical linear algebra.*  
D. Hu, S. Ubaru, A. Gittens, K. Clarkson, L. Horesh, and V. Kalantzis.  
International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2021.
- [11] *Dynamic Graph Convolutional Networks Using the Tensor M-Product.*  
O. Malik, S. Ubaru, L. Horesh, M. Kilmer, and H. Avron.  
SIAM International Conference on Data Mining (SDM), 2021.
- [10] *Multilabel Classification by Hierarchical Partitioning and Data-dependent Grouping.*  
S. Ubaru, S. Dash, O. Gunluk, and A. Mazumdar.  
Advances in Neural Information Processing Systems (NeurIPS), 2020.
- [9] *Spectrum-Adapted Polynomial Approximation for Matrix Functions.*  
L. Fan, D. Shuman, S. Ubaru, and Y. Saad  
International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2019.
- [8] *Find the dimension that counts: Fast dimension estimation and Krylov PCA.*  
S. Ubaru, A.-K. Seghouane, and Y. Saad  
SIAM International Conference on Data Mining (SDM), 2019.
- [7] *Spectral Sums Beyond Fast Matrix Multiplication: Algorithms and Hardness.*  
C. Musco, P. Netrapalli, A. Sifford, S. Ubaru, and D. P. Woodruff  
Innovations in Theoretical Computer Science (ITCS), 2018.
- [6] *UoI-NMF<sub>cluster</sub>: A Robust Nonnegative Matrix Factorization Algorithm for Improved Parts-Based Decomposition and Reconstruction of Noisy Data.*  
S. Ubaru, K. Wu, and K. E. Bouchard  
IEEE International Conference on Machine Learning and Applications (ICMLA), 2017.  
**\*Best Paper Award.\***
- [5] *Union of Intersections (UoI) for Interpretable Data Driven Discovery and Prediction.*  
K. E. Bouchard, A. F. Bujan, F. Roosta-Khorasani, S. Ubaru, Prabhat, A. M. Snijders, J.-H. Mao, E. F. Chang, M. W. Mahoney, and S. Bhattacharyya  
Advances in Neural Information Processing Systems (NeurIPS), 2017.

- [4] *Multilabel Classification with Group Testing and Codes.*  
**S. Ubaru** and A. Mazumdar  
International Conference on Machine Learning (ICML), 2017.
- [3] *Fast methods for estimating the Numerical rank of large matrices.*  
**S. Ubaru** and Y. Saad  
International Conference on Machine Learning (ICML), 2016.
- [2] *Group testing schemes from low-weight codewords of BCH codes.*  
**S. Ubaru**, A. Mazumdar, and A. Barg  
IEEE International Symposium on Information Theory (ISIT), 2016.
- [1] *Low rank approximation using error correcting coding matrices.*  
**S. Ubaru**, A. Mazumdar, and Y. Saad  
International Conference on Machine Learning (ICML), 2015.

## Book chapters

- [2] *Quantum Computing Algorithms For Decision Making Under Uncertainty.*  
L. Horesh, K. Clarkson, V. Kalantzis, M. Squillante, **S. Ubaru**, and A. Abboud  
Air Force Research Lab Technical Report, AFRL-RI-RS-TR-2021-121, 2021.
- [1] *Applications of trace estimation techniques.*  
**S. Ubaru** and Y. Saad  
High Performance Computing in Science and Engineering, LNCS book series, vol. 11087, ch. 2, pp 19–33, 2018.

## Workshop papers

- [4] *On Quantum Algorithms for Random Walks in the Nonnegative Quarter Plane.*  
V. Kalantzis, M. S. Squillante, **S. Ubaru**, and L. Horesh.  
Workshop on Mathematical performance Modeling and Analysis (MAMA), 2022.
- [3] *Multilabel prediction in log time and data-dependent grouping.*  
**S. Ubaru**, S. Dash, O. Gunluk, and A. Mazumdar  
NeurIPS - Workshop on Information Theory and Machine Learning, 2019.
- [2] *Tensor graph neural networks for prediction on time varying graphs.*  
O. Malik, **S. Ubaru**, L. Horesh, M. Kilmer, and H. Avron  
NeurIPS - Workshop on Graph Representation Learning, 2019.
- [1] *Union of Intersections (UoI) for interpretable data driven discovery and prediction in neuroscience.*  
K. E. Bouchard, P. Sachdeva, S. Bhattacharyya, M. Balasubramanian, and **S. Ubaru**  
Computational and Systems Neuroscience (Cosyne), 2019.

## Preprints

- [6] *Randomized matrix-free quadrature for spectrum and spectral sum approximation.*  
T. Chen, T. Trogon, and **S. Ubaru**, 2022.
- [5] *PCENet: High Dimensional Surrogate Modeling for Learning Uncertainty.*  
P.F. Shustin, **S. Ubaru**, V. Kalantzis, L. Horesh, and H. Avron, 2022.
- [4] *Efficient Quantum Computation of the Fermionic Boundary Operator.*  
I.Y. Akhalwaya, Y.H. He, L. Horesh, V. Jejjala, W. Kirby, K. Naidoo, and **S. Ubaru**, 2022.
- [3] *Quantum Topological Data Analysis with Linear Depth and Exponential Speedup.*  
**S. Ubaru**, I. Akhalwaya, M. Squillante, K. Clarkson, and L. Horesh, 2021.
- [2] *Unsupervised Hierarchical Graph Representation Learning with Variational Bayes.*  
**S. Ubaru** and J. Chen, 2020.
- [1] *Provably convergent acceleration in factored gradient descent with applications in matrix sensing.*  
T. Ajayi, D. Mildebrath, A. Krylidis, **S. Ubaru**, G. Kollias, and K. E. Bouchard, 2019.

## Patents

- [10] *P202010357 - Detecting network patterns using random walks.*  
K.P. Onak, **S. Ubaru**, A. Abboud, and T. Suzumura. (filed)
- [9] *P202101832 - Hardware acceleration with preconditioners.*  
V. Kalantzis, L. Horesh, and **S. Ubaru**. (filed)

- [8] *P202008482 - Projection-based techniques for updating SVD in evolving datasets.*  
V. Kalantzis, G. Kollias, **S. Ubaru**, L. Horesh, and K. Clarkson. (filed)
- [7] *P202006530 - Determining triangles in graph data structures using crosspoint array.*  
V. Kalantzis, **S. Ubaru**, L. Horesh, and H. Avron. (filed)
- [6] *US20220107991A1 - Fast PCA of evolving data using analog crossbar array.*  
**S. Ubaru**, V. Kalantzis, L. Horesh, M. Squillante, and H. Avron. (pending)
- [5] *US20220083623A1- Accelerated quasi-newton methods on analog crossbar hardware.*  
V. Kalantzis, **S. Ubaru**, L. Horesh, H. Avron, and M. Onen. (pending)
- [4] *US20210357540A1 - Matrix sketching using analog crossbar architectures.*  
L. Horesh, M. Onen, H. Avron, T. Gokmen, **S. Ubaru**, and V. Kalantzis. (pending)
- [3] *P201809232 - Short-depth quantum amplitude estimation without eigenstate collapse.*  
I. Akhalwaya, K. Clarkson, L. Horesh, M. Squillante, **S. Ubaru**, and V. Kalantzis. (filed)
- [2] *US20210174242A1 - Automatic multilabel classification using machine learning.*  
**S. Ubaru**, S. Dash, O. Gunluk, A. Mazumdar and L. Horesh. (pending)
- [1] *US20210090182A1 - Tensor-based predictions from analysis of time-varying graphs.*  
O. Malik, **S. Ubaru**, L. Horesh, M. Kilmer, and H. Avron. (pending)

## Presentations

- 2022 *Multilabel Classification by Hierarchical Partitioning and Grouping*  
- RISE workshop, Austin, TX.  
*Quantum Topological Data Analysis with Linear Depth and Exponential Speedup*  
- Quantum Information Group, University of Texas at Austin.  
- Quantum computing course, Columbia University, New York
- 2021 *Into another dimension - a new tensor algebra for learning on dynamic graphs*  
- Seminar at University of Albany, NY.  
*Algorithms for estimating spectral sums for large matrices*  
- Oden Institute Seminar Series, University of Texas at Austin  
*Novel Tensor Framework for Data Representation and Compression*  
- SIAM Conference on Computational Science and Engineering (CSE21)
- 2020 *Multilabel Classification by Hierarchical Partitioning and Data-dependent Grouping*  
- Neural Information Processing Systems (NeurIPS) Virtually held  
*Novel tensor framework for model reduction and neural networks*  
- ICERM: Workshop on Algorithms for Dimension and Complexity Reduction, Brown University, RI
- 2019 *Tensor graph neural networks for prediction on time varying graphs*  
- NeurIPS: Workshop on Graph Representation Learning, Vancouver, Canada  
*Multilabel prediction in log time and data-dependent grouping*  
- NeurIPS: Workshop on Information Theory and Machine Learning, Vancouver, Canada  
*Spectrum approximation by Lanczos Quadrature and Preconditioned SVRG*  
- 20th International Conference On Preconditioning, Minneapolis, MN, USA.  
*Find the dimension that counts: Fast dimension estimation and Krylov PCA*  
- SIAM International Conference on Data Mining (SDM), Calgary, Canada.  
*Error Correcting Codes for Machine Learning.*  
- CSA Seminar, Indian Institute of Science (IISc), Bengaluru, India.
- 2018 *Error Correcting Codes for Machine Learning*  
- Theory Seminar, University of Massachusetts, Amherst, MA.  
- IP Seminar, IBM T.J. Watson Research Center, NY.
- 2017 *UoI-NMF<sub>cluster</sub>: Robust Nonnegative Matrix Factorization Algorithm for Noisy Data*  
- International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico.  
*Multilabel Classification with Group Testing and Codes*  
- Neural Systems and Engineering Labs, Lawrence Berkeley National Laboratory, CA.  
- International Conference on Machine Learning (ICML), Sydney, Australia.  
*Error Correcting Codes for Machine Learning*  
- The University of Melbourne, Melbourne, Australia  
*UoI-NMF<sub>cluster</sub> and UoI-CUR: Union of Intersections methods for matrix approximations*  
- Neural Systems and Engineering Labs, Lawrence Berkeley National Laboratory, CA.
- 2016 *Error correcting codes for low rank approximation and group testing*  
- BLISS Seminar, University of California, Berkeley, CA.  
*Fast methods for estimating the Numerical rank of large matrices*  
- International Conference on Machine Learning (ICML), New York, NY.

- 2015 *Low rank approximation using error correcting coding matrices*  
- International Conference on Machine Learning (ICML), Lille, France

## Awards

- 2022 **Outstanding Innovation Award**, IBM Research  
2021 *IBM Research Accomplishment Award*  
2019 *UMN Best Dissertation Award*, Computer Science Department nominee  
2018 **Herman Goldstine Fellowship**, IBM Research  
2017 **Best Paper Award**, International Conference on Machine Learning and Applications (ICMLA)  
2011 *Visiting Research Student Program*, Raman Research Institute

## Teaching

### IBM Research

Summer 2021 Mathematics of Big Data: Sketching and (Multi-) Linear Algebra at MSRI Berkeley, *Instructor*

### University of Minnesota

- Spring 2018 CSci 2033, Elementary Computational Linear Algebra, *Recitation Instructor*  
Fall 2017 CSci 5304, Computational Aspects of Matrix Theory, *Teaching Assistant*  
Spring 2017 CSci 8314, Sparse Matrix Computations, *Temporary Instructor*

## Mentoring

### IBM Research - Interns

- 2022 Elizabeth Yang (PhD Student - U of California, Berkeley)  
2021 - present Paz Fink Sustin (PhD Student - Tel Aviv University)  
2021 Lucky Yerimah (PhD Student - RPI)  
2021 Emily Thompson (REU Student via. DIMACS)  
2020- present Dong Hu (PhD Student - RPI)  
2019 Osman Malik (PhD Student - U of Colorado, Boulder)

## Service

Reviewer: *Conferences:*

Neural Information Processing Systems (NeurIPS) 2016, 2019- 22 ; International Conference on Machine Learning (ICML) 2018 - 22 ; International Conference on Learning Representations (ICLR) 2021-22; AAAI conference on Artificial Intelligence 2020 - 22 ; ACM-SIAM Symposium on Discrete Algorithms (SODA) 2021; Conference on Uncertainty in Artificial Intelligence (UAI) 2019 - 22; IEEE International Conference On Machine Learning And Applications 2018 - 19; IEEE International Symposium on Information Theory (ISIT) 2017.

*Journals:*

Acta Materialia; PLOS One; Journal of Machine Learning Research; SIAM Journal on Matrix Analysis and Applications (3 articles); IEEE Transactions on Signal Processing; IEEE Transactions on Image Processing; Signal Processing Letters, IEEE Transactions on Information Theory; Electronic Transactions on Numerical Analysis (2 articles); IEEE Transactions on Signal and Information Processing over Networks; BIT Numerical Mathematics; Linear Algebra and Its Applications; Journal of Mathematics and Applications, Mathematical Modelling and Analysis; Journal of Imaging.

Organizer: ICMLA Challenge 2018 - Parts based decomposition of noisy data.  
MSRIT ROBONXG-2012, a week long robotics festival.

## Programming skills

C, C++, Matlab, Python, LATEX, Basic web programming.

---

## References

**Yousef Saad**

CSE Distinguished Professor,  
Computer Science and Engineering,  
University of Minnesota- Twin Cities, MN, USA.  
saad@umn.edu.

**Lior Horesh**

Senior Manager  
Mathematics of AI  
IBM T.J. Watson Research Center, NY, USA  
lhoresh@us.ibm.com

**Arya Mazumdar**

Associate Professor,  
The Halicioglu Data Science Institute,  
University of California, San Diego, CA, USA.  
arya@ucsd.edu.

**Sanjeeb Dash**

Manager  
Foundation of Optimization  
IBM T.J. Watson Research Center, NY, USA  
sanjeebd@us.ibm.com