List of publications – Naren Sarayu Manoj (nsm@ttic.edu)

Please see below an annotated list of publications, preprints, and unpublished manuscripts that I had a significant contribution to in chronological order. Please email me if you would like a copy of any of the below (including manuscripts that are not publicly available).

1. **Near-Optimal Streaming Ellipsoidal Rounding for General Convex Polytopes** (working manuscript; plan to submit to STOC 2024).

   Yury Makarychev, *Naren Sarayu Manoj*, Max Ovsiankin.[α−β]

   Description: We generalize the results from (6) to the case where the convex body in question is asymmetric. We also use our algorithm to yield the first algorithms for general convex hull coresets, wherein the algorithm must output a subset of the input whose convex hull forms a good approximation to the convex hull of the original input. This is particularly useful in big-data settings where storing the entire convex body is prohibitively expensive.

2. **An Ellipsoidal Sampling Condition for Hypergraph Sparsification and Generalized Matrix Row Sampling** (working manuscript; plan to submit to STOC 2024).

   ***Naren Sarayu Manoj*, Max Ovsiankin.**[α−β]

   Description: Consider the problem of ℓ_p matrix row sampling – we are given a matrix and we want to choose a weighted subset of the rows of this matrix so that the reweighted submatrix approximately preserves the ℓ_p norm of matrix-vector products with respect to the original matrix, for all possible vectors. The best known selection criteria (Lewis weights) had a rather opaque analysis. In this paper, we identify a general geometric condition that allows one to easily and nearly optimally analyze many different sampling schemes, including Lewis weights. Our analysis is also general enough to yield the first results for a particular variant of hypergraph hyperedge sampling, in which we are given a hypergraph and we must output a sub-hypergraph that preserves various cut properties of the original hypergraph.

3. **Dueling Optimization With a Monotone Adversary** (under conference review; oral at OPT 2023).

   Avrim Blum, Meghal Gupta, Gene Li, *Naren Sarayu Manoj*, Aadirupa Saha, Chloe Yang.[α−β]

   Description: We formulate the problem of "dueling convex optimization with a monotone adversary", which is an extension of dueling convex optimization with an adversary motivated by those seen in semi-random models. We give a random walk-based algorithm to solve this more challenging extension of the feedback model studied in dueling convex optimization. The appeal of this algorithm is that it is simple, fast, applicable to many function classes, and requires even less information than what is given in zeroeth-order optimization problems.

4. **Shortest Program Interpolation Learning** (COLT 2023).

   ***Naren Sarayu Manoj*, Nathan Srebro.**[α−β]

   Description: We give the first generalization bound for the shortest-program learning rule when applied to noisy data. We prove the possibly surprising phenomenon that the overfitting of this learning rule can be carefully bounded despite the fact that the output classifier interpolates its training set, which runs counter to conventional wisdom.

5. **An Optimal Algorithm for Certifying Monotone Functions** (SOSA 2023).

   Meghal Gupta, *Naren Sarayu Manoj*.[α−β]

   Description: We resolve an open problem posed in a paper from STOC 2022 regarding finding the optimal query complexity for the problem of certifying Boolean functions. It is helpful to think of this problem as one of explainability – given a function and an input, can we identify the parts of the input that fix the function's value on that input while evaluating only a small number of additional inputs on the function?

6. **Streaming Algorithms for Ellipsoidal Approximation of Convex Polytopes** (COLT 2022).

   Yury Makarychev, *Naren Sarayu Manoj*, Max Ovsiankin.[α−β]

   Description: We gave the first algorithm to round convex polytopes in a stream. The solution is quite elegant – the algorithm is probably the most natural algorithm one could apply for this task, but the analysis is technically involved. The primitive of rounding convex sets has many downstream applications in machine learning and optimization, so building fast algorithms for this task is important. Finally, the algorithm is applicable in low-memory settings, wherein storing the whole convex body in memory is impractical.

7. **Excess Capacity and Backdoor Poisoning** (NeurIPS 2021, Spotlight).

   Avrim Blum, *Naren Sarayu Manoj*.[α−β]

   Description: We theoretically analyze a type of training data corruption that empirically creates dangerous "backdoors" in ML classifiers. The main technical contributions are building the first formal framework within which one can analyze backdoor data poisoning attacks, giving example analyses of learning problems within this framework, and analyzing possible algorithms that yield robust classifiers.

8. **Random Smoothing Might be Unable to Certify ℓ_∞ Robustness for High-Dimensional Images** (JMLR 2020).

   Avrim Blum, Travis Dick, *Naren Sarayu Manoj*, Hongyang Zhang.[α−β]

   Description: We show a hardness result for random smoothing to achieve certified adversarial robustness against attacks in the ℓ_p ball of radius ϵ when $p > 2$, i.e., that any noise distribution $\mathcal{D}$ over $\mathbb{R}^d$ that provides ℓ_p robustness for all base classifiers with $p > 2$ must satisfy $\mathbb{E}\eta_i^2 \gtrsim \Omega(d^{1-2/p})$ for most of the features (pixels) of vector $\eta \sim \mathcal{D}$. Therefore, for high-dimensional images with pixel values bounded in $[0, 255]$, the required noise will eventually dominate the useful information in the images, leading to trivial smoothed classifiers.

9. **Development and Validation of a Deep Learning Algorithm for Gleason Grading of Prostate Cancer From Biopsy Specimens** (JAMA Oncology 2020).

   Kunal Nagpal, Davis Foote, . . . , *Naren Sarayu Manoj*, . . . , Krishna Gadepalli, Greg Corrado, Lily Peng, Martin Stumpe, Craig Mermel.

   Description: We give a deep learning algorithm to grade prostate biopsies on the Gleason scale (which essentially rates the cell deformities present on the slide). I was responsible for building the architecture search that ultimately output the neural network architecture that we used. Our model significantly outpeformed general pathologists on the Gleason grading task.

10. **Quantifying Perceptual Distortion of Adversarial Examples** (arXiv preprint 1902.08265).

    Matt Jordan, *Naren Sarayu Manoj*, Surbhi Goel, Alex Dimakis.

    Description: To demonstrate the value of quantifying the perceptual distortion of adversarial examples, we present and employ a unifying framework fusing different attack styles. We first prove that our framework results in images that are unattainable by attack styles in isolation. We then perform adversarial training using attacks generated by our framework to demonstrate that networks are only robust to classes of adversarial perturbations they have been trained against, and combination attacks are stronger than any of their individual components. Finally, we experimentally demonstrate that our combined attacks retain the same perceptual distortion but induce far higher misclassification rates when compared against individual attacks.