Research Statement

Namrata Vaswani, Iowa State University

My research lies at the intersection of machine learning for high dimensional problems, signal and information processing and applications in video, big-data analytics and bio-imaging. More specifically, I have worked on designing and analyzing online algorithms for various high-dimensional structured data recovery problems and on demonstrating their usefulness in dynamic magnetic resonance imaging (MRI) and in video analytics. In the last two decades, the sparse recovery problem, or what is now more commonly referred to as compressive sensing (CS), has been extensively studied, see for example [1, 2, 3, 4, 5, 6] and later works. More recently various other structured data recovery problems, such as low-rank or low-rank plus sparse matrix recovery, have also been studied in detail.

Sparse recovery or CS refers to the problem of recovering a sparse signal from a highly reduced set of its projected measurements. Many medical imaging techniques image cross-sections of human organs non-invasively by acquiring their linear projections one at a time and then reconstructing the image from these projections. For example, in magnetic resonance imaging (MRI), one acquires Fourier projections one at a time, while in Computed Tomography (CT), one acquires the Radon transform coefficients one a time. For all these applications, the ability to accurately reconstruct using fewer measurements directly translates into reduced scan times and hence sparse recovery methods have had a huge impact in these areas. Low-rank matrix recovery (or matrix completion) refers to the problem of recovering a low-rank matrix from a subset of its entries. A motivating application for low-rank matrix completion is the Netflix problem where the goal is to recover the movie ratings' matrix when each user rates only a few movies.

Low-rank plus sparse matrix recovery refers to recovering a sparse matrix and a low-rank matrix from their sum or from a subset of entries of their sum. As explained in [7], this can equivalently be understood as a robust principal components' analysis (PCA) problem. An important application where it occurs is video layering (separating a video sequence into foreground and background layers). The foreground usually consists of one or more moving persons or objects and hence is a sparse image. In case of a static camera video, the background image usually changes only gradually over time and the changes are global (dense) [7], e.g., moving lake waters or moving trees in a forest. Thus it is well modeled as lying in a low-dimensional subspace that is fixed or slowly changing. Video layering is a key first step to simplifying many video analytics and computer vision tasks, e.g., video surveillance (to track moving foreground objects), background video recovery and subspace tracking in the presence of frequent foreground occlusions, or low-bandwidth mobile video chats or video conferencing (can transmit only the foreground layer). Besides video, other applications where robust PCA is needed include detecting anomalous and suspicious behavior from dynamic social network connectivity patterns, detecting data traffic anomalies in dynamic computer networks, dynamic MRI based sparse region of interest tracking from dense slowly changing backgrounds [8], or outlier-robust recommendation system design [7].

In the last several years, in my group, we have developed provably accurate online (recursive) solutions to two structured data recovery problems: (a) online sparse matrix recovery, and (b) online sparse plus low-rank matrix recovery and online matrix completion. Problem (a) can also be understood as a problem of recursive recovery of sparse vector (signal) sequences. Problem (b) can be interpreted either as a problem of online robust PCA or robust subspace tracking, or as a problem of online sparse recovery in large but structured (low-dimensional and dense) noise. Online algorithms are needed for real-time applications; and even for offline applications, they are faster and need less storage compared to batch techniques. Moreover, as we will see, online approaches often provide an easy way to use temporal dependencies in the dataset without greatly increasing the number of unknown parameters to be estimated. We describe below our recent work on problem (b) first and then on problem (a).

Online Sparse + Low-Rank Matrix Recovery (Online Robust PCA)

Principal Components Analysis (PCA) is a widely used tool for dimension reduction. Given a matrix of data D, PCA seeks to recover a small number of orthogonal directions that contain most of the variability of the data. This is typically accomplished by performing a singular value decomposition (SVD) of D and retaining the left singular vectors corresponding to the largest singular values. A limitation of SVD is that it is highly sensitive to outliers. "Outlier" is a loosely defined term that refers to any corruption that is not small compared to the true data vector and that occurs occasionally. As suggested in [9], an outlier can be nicely modeled as a sparse vector whose nonzero entries can have any magnitude. Robust PCA, which refers to the problem of PCA in the presence of outliers, has been a well-studied problem for a long time and many useful heuristics have been proposed for it, e.g., see [10] and references therein. However none of the practically useful algorithms from older literature come with performance guarantees. In recent works [7, 11], Candès et. al. and Chandrasekaran et. al. posed robust PCA as a problem of separating a sparse matrix S and a low-rank matrix L from their sum, $Y := L + S$. They introduced the Principal Components Pursuit (PCP) program and obtained performance guarantees for it under mostly mild assumptions [7], [11]. Later work by Hsu et. al. [12] improved the result of [11]. Since then, there has been much later work on provably accurate robust PCA solutions but all of it has been for batch methods.

In our work, we consider an online or recursive version of the robust PCA problem where we seek to separate observed data vectors, $y_t := \ell_t + s_t$, into low dimensional components, ℓ_t, and sparse components, s_t, as they arrive, using the previous estimates, rather than re-solving the entire problem at each time t. An application where this type of problem is useful is in video analysis [10]. Imagine a video sequence that has a distinct background and foreground. An example might be a surveillance camera where a person walks across the scene. If the background does not change very much, and the foreground is sparse (both practical assumptions), then separating the background and foreground can be viewed as a robust PCA problem. Other applications where robust PCA can be useful include detecting anomalous behavior in dynamic social networks; or recommendation system design in the presence of outliers and missing data.In most of these applications, an online solution is desirable.

In our recent work [13, 14, 15, 16, 17, 18], we have introduced a novel online solution called Recursive Projected CS (ReProCS); obtained a correctness result for it under mild assumptions; and shown that it significantly outperforms most existing batch as well as recursive robust PCA methods for video layering of difficult videos. To the best of our knowledge, ReProCS is the *first provably accurate algorithm* for online robust PCA. We explain this work below.

[ReProCS for online robust PCA] The Recursive Projected CS (ReProCS) [13, 14, 16] solution uses two extra, but usually practically valid, assumptions beyond what PCP [7, 11] and other batch methods need. First, it assumes that an accurate estimate of the subspace of ℓ_t is available at the initial time (easy to obtain using a short sequence of background-only video frames) and second, that its subspace is either fixed or changes slowly over time (valid assumption for slow changing backgrounds). Like PCP, it also needs a denseness (non-sparseness) assumption on the left singular vectors of the matrix L. Its key idea is as follows. At any time t, it first projects the measurement vector, y_t, perpendicular to the current estimate of the subspace of ℓ_t. Because of the slow subspace change assumption, this approximately nullifies ℓ_t. Recovering s_t from these projected measurements then becomes a standard sparse recovery problem in small noise. The denseness (non-sparseness) assumption on the left singular vectors of L ensures that s_t can be recovered accurately from the projected measurements by ell-1 minimization, or any other sparse

recovery technique. The recovered s_t is used to estimate $\ell_t = y_t - s_t$ by subtraction and this is then used to update the subspace estimate using a modification of standard PCA that we call projection-PCA. Because of how the sparse recovery is done, ReProCS requires a significantly weaker assumption on how frequently the support of the sparse vectors needs to change compared with PCP.

[Video Analytics Application] In foreground and background layering experiments with real video sequences containing large and slow moving or occasionally static objects (infrequent support change), as long as a short initial background-only sequence is available, ReProCS significantly outperforms batch robust PCA approaches, such as PCP and RSL [10], as well as online algorithms such as incremental RSL [19], and a recently introduced algorithm called GRASTA [20]. The experimental comparisons are available in [16, 21] and at `http://www.ece.iastate.edu/~hanguo/PracReProCS.html`.

[Correctness result for ReProCS] In recent work [17, 18], we have shown that, as long as the ReProCS algorithm parameters are set appropriately, a good-enough estimate of the subspace of ℓ_0 is available, slow subspace change (quantified in a certain way) holds, the left singular vectors of the matrix L are dense enough, and there is a certain amount of support change of s_t at least every so often, then the support of s_t can be exactly recovered with high probability. Moreover, the sparse and low-rank matrix columns can be recovered with bounded and small error; and the subspace recovery error decays to a small value within a short delay of a subspace change time. To the best of our knowledge, our result [17, 18] is among the *first* correctness results for an online (recursive) robust PCA algorithm or equivalently for online sparse plus low-rank matrix recovery. We first studied ReProCS in [15] and obtained a performance guarantee for it. However the result was not a correctness result (it depended on intermediate algorithm estimates satisfying a certain property). The same is true for a later result of Feng et al [22].

Online algorithms are needed for real-time applications; and even for offline applications, they are faster and need less storage compared to batch techniques. Moreover, online approaches can provide a natural way to exploit temporal dependencies in the dataset without increasing the problem size (number of unknowns). In our case, we show that ReProCS uses slow subspace change to allow for significantly more correlated support sets of the sparse vectors than do the various results for PCP [7, 11, 12]. Our result for ReProCS allows a constant fraction of nonzeros in any row of S, while the result for PCP from [11, 12] only allows this fraction to be proportional to $1/r$ where r is the rank of L. The result for PCP from [7] uses an even stronger assumption - it requires uniformly randomly selected support sets. Of course this advantage for ReProCS comes at a cost. It needs a tighter bound on the rank-sparsity product compared to what is needed by the PCP result from [7]. Also, the slow subspace change model that we assume for obtaining the ReProCS result can be understood as a piecewise stationary approximation to the nonstationary subspace change at each time (that occurs in practice).

For obtaining our result, we needed to develop new proof techniques and we expect that these can potentially be applied to various other problems involving a PCA step where the data and noise are correlated. New techniques were needed because all existing correctness results were only for batch robust PCA approaches. Moreover, almost all existing work on finite sample PCA assumes that the error between the measured and true data vectors is uncorrelated with the true data, see, e.g., [23] and references therein. However, in case of ReProCS, because of how the estimate $\hat{\ell}_t$ is computed, the error $e_t := \hat{\ell}_t - \ell_t$ is correlated with ℓ_t and so our proof could not just combine a result for sparse recovery with a result for standard PCA. This is also the reason a projection-PCA algorithm was needed in the subspace update step of ReProCS.

[Modified-PCP] A key limitation of ReProCS is that it does not use the fact that the noise seen by its projected sparse recovery step also has structure; it lies close to a very low-dimensional subspace. The only way to use this fact is in a piecewise batch fashion (one cannot impose a low-dimensional subspace assumption on a single vector). The resulting solution can be understood as a modification of the PCP program when partial subspace knowledge is available and hence we call it modified-PCP [24, 25]. Modified-PCP solves the problem of robust PCA with partial subspace knowledge using an idea inspired by our older work on modified-CS for sparse recovery with partial support knowledge [26] (explained below). The advantage

of modified-PCP is that it needs a weaker assumption on the rank-sparsity product compared to both PCP and ReProCS. However, its disadvantage is similar to that of PCP, it cannot handle correlated support change as well as ReProCS can.

Online Sparse Matrix Recovery (Recursive Recovery of Sparse Signal Sequences)

The goal of this work was to design and analyze recursive algorithms for causally reconstructing a time sequence of (approximately) sparse signals from highly undersampled measurements. The signals were assumed to be sparse in some transform domain referred to as the sparsity basis and their sparsity patterns (support set of the sparsity basis coefficients) could change with time. One key application where this problem occurs is dynamic MRI for real-time medical applications such as interventional radiology and MRI-guided surgery, or in functional MRI to track brain activation changes. MRI is a technique for cross-sectional imaging that sequentially acquires the 2D Fourier projections of the cross-section to be reconstructed. Since MR data is acquired one Fourier projection at a time, the ability to accurately reconstruct using fewer measurements directly translates into reduced scan times. Shorter scan times along with online (causal) and fast (recursive) reconstruction can enable real-time imaging of fast changing physiological phenomena, thus making many interventional MRI applications practically feasible. Cross-sectional images of the brain, heart, larynx or other human organ images are piecewise smooth, and thus approximately sparse in the wavelet domain. In a time sequence, their sparsity pattern changes with time, but quite slowly [27, 28, 29]. This simple idea, which was first used in [27], is the key reason our proposed algorithms can achieve accurate reconstruction from much fewer measurements. In recent years, the static sparse recovery or compressive sensing (CS) problem has been thoroughly studied [1, 2, 3, 4, 5, 6]. But most existing algorithms for the dynamic problem, e.g. [30, 31], just use CS solutions to jointly reconstruct the entire time sequence in one go. This is a batch solution and as a result (a) it is very slow in recovering a long sequence and (b) its memory requirement increases linearly with the sequence length. The alternative - solving the CS problem at each time separately (simple-CS) - is online, fast and memory-efficient, but needs many more measurements (higher scan time).

Moreover, batch solutions such as [30, 31] assume Fourier sparsity along the time axis which may not always be valid. An alternative class of batch solutions solve the Multiple Measurements' Vector (MMV) problem, but this assumes that the sparsity pattern of the signal sequence is constant with time. This assumption is often not valid either.

To the best of our knowledge, our recent work [27, 28, 32, 33] proposed the *first* solutions for recursively reconstructing sparse signal sequences using much fewer measurements than those needed for accurate recovery using simple-CS methods. The computational and storage complexity of our proposed algorithms is only as much as that of simple-CS, but their reconstruction performance is much better. For example, in results reported in [33], for reconstructing simulated dynamic MRI sequences, our proposed algorithm (modified-CS) needed only about 16-19% measurements for accurate recovery while a simple-CS solution (basis pursuit solved using the L1-Magic software) needed about twice as many measurements for the same recovery error. Existing MR scanners use the inverse Fourier transform which needs 100% measurements.

Our work described below relies on the following easily verifiable assumptions.

1. The sparsity patterns (support set of the sparse basis vectors) of natural signal/image sequences usually change "slowly" over time [29, 33].

2. In most cases, the values of the nonzero coefficients also change gradually over time [34].

The first assumption above is a new assumption introduced in our work [27, 28]. It has been empirically verified both for medical image sequences and for video [32, 33, 35]. The second assumption is commonly used both for adaptive filtering as well as for various tracking problems.

When using only assumption 1 above, *the above problem can be reformulated as one of sparse reconstruction with partially "known" support*. The support estimate from the previous time serves as the "known" part. We can further improve the proposed algorithm by also using assumption 2.

[LS-CS and KF-CS] The key idea of our first approach (LS-CS-residual or LS-CS) is to solve the ℓ_1 minimization problem with the observation replaced by the least squares (LS) observation residual computed using the "known" part of the support [27, 32]. The LS residual measures a signal that has much fewer large components compared to the original signal (it is what can be called a "sparse-compressible" signal). As a result, when fewer measurements are available, the LS-CS reconstruction error is lower than that of simple-CS methods. By also using fact 2, we can replace the LS residual by the *Kalman filtering residual (KF-CS)* [27]. This improves the reconstruction particularly when the number of measurements is too few even for LS-CS.

[Modified-CS and Regularized modified-CS] Even though LS-CS and KF-CS improved reconstruction accuracy over simple-CS methods, they could not be used for "exact" reconstruction from fewer noise-free measurements. This led to our second and more powerful approach - modified-CS [33]. Denote the "known" part of the support by T. Modified-CS tries to find the signal that is sparsest outside of T while satisfying the data constraint. If T has small error (few extras and misses), modified-CS can achieve *provably exact reconstruction from very few measurements* [33]. By also using slow signal value change, one can design regularized modified-CS which also constrains the change of the nonzero coefficient values along T [34, 36]. In numerical experiments as well as in experiments with simulated dynamic MR imaging, modified-CS significantly outperformed existing work at the time [33].

Under the practically valid assumption of slowly changing support, we have also been able to prove (a) exact reconstruction and (b) error stability over time, using fewer measurements than what simple-CS methods need. We explain these results next.

[Exact reconstruction] We have obtained exact reconstruction conditions for Modified-CS and have argued that it achieves exact reconstruction under much weaker sufficient conditions (using much fewer noise-free measurements) than those needed to provide the same guarantee for simple CS as long as slow support change holds [33].

[Stability over time] For both LS-CS and modified-CS, for the noisy measurements' case, we have shown the following. Under fairly mild assumptions (bounded noise, a large enough nonzero coefficient magnitude increase rate, and weaker requirements on the number of measurements than what is needed for obtaining similar error bounds on simple-CS error, *both the support recovery errors and the reconstruction errors, are "stable", i.e. they remain bounded by time-invariant and small values at all times [32, 37, 38, 39].* Stability is critical for any recursive algorithm since it ensures that the reconstruction error does not blow up over time. In practice, say in real-time MRI, provably stable and small error would mean that the reconstructed images seen by the interventional radiologist or surgeon would always a close approximation of reality.

References

[1] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, vol. 3, 1996, pp. 1688–1691.

[2] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using focuss: A re-weighted norm minimization algorithm," *IEEE Trans. Sig. Proc.*, pp. 600 – 616, March 1997.

[3] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal of Scientific Computing*, vol. 20, pp. 33–61, 1998.

[4] E. Candes and T. Tao, “Decoding by linear programming,” *IEEE Trans. Info. Th.*, vol. 51(12), pp. 4203 – 4215, Dec. 2005.

[5] E. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Info. Th.*, vol. 52(2), pp. 489–509, February 2006.

[6] D. Donoho, “Compressed sensing,” *IEEE Trans. Info. Th.*, vol. 52(4), pp. 1289–1306, April 2006.

[7] E. J. Candès, X. Li, Y. Ma, and J. Wright, “Robust principal component analysis?” *Journal of ACM*, vol. 58, no. 3, 2011.

[8] R. Otazo, E. Candes, and D. K. Sodickson, “Low-rank and sparse matrix decomposition for accelerated dynamic mri with separation of background and dynamic components,” *Submitted to Magnetic Resonance in Medicine*, 2013.

[9] J. Wright and Y. Ma, “Dense error correction via l1-minimization,” *IEEE Trans. on Info. Th.*, vol. 56, no. 7, pp. 3540–3560, 2010.

[10] F. D. L. Torre and M. J. Black, “A framework for robust subspace learning,” *International Journal of Computer Vision*, vol. 54, pp. 117–142, 2003.

[11] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, “Rank-sparsity incoherence for matrix decomposition,” *SIAM Journal on Optimization*, vol. 21, 2011.

[12] D. Hsu, S. Kakade, and T. Zhang, “Robust matrix decomposition with sparse corruptions,” *IEEE Trans. Info. Th.*, Nov. 2011.

[13] C. Qiu and N. Vaswani, “Real-time robust principal components’ pursuit,” in *Allerton Conf. on Communications, Control and Computing*, 2010.

[14] ——, “Recursive sparse recovery in large but correlated noise,” in *Allerton Conf. on Communication, Control, and Computing*, 2011.

[15] C. Qiu, N. Vaswani, B. Lois, and L. Hogben, “Recursive robust pca or recursive sparse recovery in large but structured noise,” *IEEE Trans. Info. Th.*, August 2014, early versions in ICASSP 2013 and ISIT 2013.

[16] H. Guo, C. Qiu, and N. Vaswani, “An online algorithm for separating sparse and low-dimensional signal sequences from their sum,” *IEEE Trans. Sig. Proc.*, August 2014.

[17] B. Lois and N. Vaswani, “A correctness result for online robust pca,” in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2015.

[18] ——, “Online matrix completion and online robust pca,” in *IEEE Intl. Symp. Info. Th. (ISIT)*, 2015, longer version at arXiv:1503.03525 [cs.IT].

[19] Y. Li, L. Xu, J. Morphett, and R. Jacobs, “An integrated algorithm of incremental and robust pca,” in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2003, pp. 245–248.

[20] J. He, L. Balzano, and A. Szlam, “Incremental gradient on the grassmannian for online foreground and background separation in subsampled video,” in *IEEE Conf. on Comp. Vis. Pat. Rec. (CVPR)*, 2012.

[21] B. Lois, N. Vaswani, and C. Qiu, "Performance guarantees for undersampled recursive sparse recovery in large but structured noise," in *GlobalSIP*, 2013.

[22] J. Feng, H. Xu, and S. Yan, "Online robust pca via stochastic optimization," in *Adv. Neural Info. Proc. Sys. (NIPS)*, 2013.

[23] B. Nadler, "Finite sample approximation results for principal component analysis: A matrix perturbation approach," *The Annals of Statistics*, vol. 36, no. 6, 2008.

[24] J. Zhan and N. Vaswani, "Robust pca with partial subspace knowledge," in *IEEE Intl. Symp. Info. Th. (ISIT)*, 2014.

[25] ——, "Robust pca with partial subspace knowledge," *IEEE Trans. Sig. Proc.*, 2015, to appear.

[26] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Processing*, September 2010.

[27] N. Vaswani, "Kalman filtered compressed sensing," in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2008.

[28] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," in *IEEE Intl. Symp. Info. Th. (ISIT)*, 2009.

[29] C. Qiu, W. Lu, and N. Vaswani, "Real-time Dynamic MR Image Reconstruction using Kalman Filtered Compressed Sensing," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2009.

[30] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. Baraniuk, "Compressive imaging for video representation and coding," in *Proc. Picture Coding Symposium (PCS), Beijing, China*, April 2006.

[31] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed sensing in dynamic mri," *Magnetic Resonance in Medicine*, vol. 59(2), pp. 365–373, January 2008.

[32] N. Vaswani, "LS-CS-residual (LS-CS): Compressive Sensing on Least Squares residual," *IEEE Trans. Sig. Proc.*, vol. 58(8), pp. 4108–4120, August 2010.

[33] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Sig. Proc.*, vol. 58(9), pp. 4595–4607, September 2010.

[34] W. Lu and N. Vaswani, "Regularized modified bpdn for noisy sparse reconstruction with partial erroneous support and signal value knowledge," *IEEE Trans. Sig. Proc.*, January 2012.

[35] C. Qiu and N. Vaswani, "Compressive sensing on the least squares and kalman filtering residual for real-time dynamic mri and video reconstruction," *IEEE Trans. Image Proc.*, 2009, submitted (shorter version in ICASSP'09).

[36] W. Lu and N. Vaswani, "Exact reconstruction conditions for regularized modified basis pursuit," *IEEE Trans. Sig. Proc.*, May 2012.

[37] J. Zhan and N. Vaswani, "Time invariant error bounds for modified-cs based sparse signal sequence recovery," *IEEE Trans. Info. Th.*, March 2015.

[38] N. Vaswani, "Stability (over time) of Modified-CS for Recursive Causal Sparse Reconstruction," in *Allerton Conf. Communication, Control, and Computing*, 2010.

[39] J. Zhan and N. Vaswani, "Time invariant error bounds for modified-cs based sparse signal sequence recovery," in *IEEE Intl. Symp. Info. Th. (ISIT)*, 2013.