1. J. Montiel Olea, **C. Rush**, A. Velez, and J. Wiesel, "The out-of-sample prediction error of the square-root-LASSO and related estimators," (resubmitted, after requested revisions, to) *The Annals of Statistics*. Available online: arxiv.org/2211.07608.

   This paper is representative of my work in novel approaches to neighborhood-based techniques for statistical robustness. This work deviates from classical ideas on using neighborhood methods for establishing robust estimators by constructing estimators that robustify against *deviations from the observed empirical distribution of the training data*, instead of deviations from some parametric model. By establishing new connections between distributionally robust optimization, empirical process theory and the so-called max-sliced Wasserstein distances, this work gives estimates of the generalization error of penalized regression problems. To the best of my knowledge, this is the first work to directly connect the optimal transport-based metric used on the space of probability measures to the loss function considered in order to provide robustness guarantees.

2. **C. Rush**, F. Skerman, A. Wein and D. Yang, "Is it easier to count communities than find them?" *14th Innovations in Theoretical Computer Science (ITCS)*. LIPIcs, 2022. Available online: arxiv.org/2212.10872.

   This paper is representative of my work in studying the computational ccomplexity of statistical inference and estimation procedures. Using low-degree polynomials, a class of algorithms that is conjectured to be optimal amongst polynomial time algorithms, this work studies community detection in graphs. The problems of detecting the presence of and recovering community structure in graphs have been well studied and an interesting landscape of statistical and computational phase transitions has emerged. This work shows that even inferring properties of the community structure, like the number and sizes of communities, is computationally hard in settings where actually finding the communities is believed to be computationally hard. This work has implications in clustering, for example, suggesting that determining the k in k-means clustering is as hard of a problem in terms of computational complexity as finding the clusters after k has been given.

3. Z. Bu, J. Klusowski, **C. Rush**, and W. Su, "Characterizing the SLOPE trade-off: a variational perspective and the Donoho-Tanner limit," *The Annals of Statistics*, 51.1 (2023): 33-61. Available online: arxiv.org/2105.13302.

   This paper is representative of my work characterizing exact statistical properties of estimators in the linear proportional asymptotic regime. This work provides the Type I/Type II error trade-off curve for the sorted ℓ_2-penalized estimator (SLOPE). A generalization of the LASSO, SLOPE allows for different levels of regularization for different elements of the estimator based on magnitude, which has been shown to provably control false discoveries in some problem settings. My work proves rigorously that the payoff for these added degrees of freedom is the possibility of simultaneous improvement in terms of both model selection and predictive performance compared to the LASSO. Moreover, for some problems in the linear asymptotic regime, it is known that the proportion of true positives recovered by the LASSO is strictly bounded away from one, even when allowing for false discoveries, something that this work demonstrates is not true for SLOPE.

4. O. Feng, R. Venkataramanan, **C. Rush**, and R. Samworth, "A unifying tutorial on approximate message passing," *Foundations and Trends in Machine Learning*, 15.4 (2022): 335-536. Available online: arxiv.org/2105.02180.

   This book is representative of my work towards advances in the methodology and analysis of approximate message passing (AMP) algorithms. A comprehensive and rigorous introduction to AMP tailored for researchers in statistics and machine learning, this work provides a mathematical perspective that not only highlights the potency and adaptability of the AMP framework, but also unifies and formalizes essential concepts within the extensive body of recent research in this domain.

5. M. Avella Medina, J. Montiel Olea, **C. Rush**, and A. Velez "On the robustness to misspecification of α-posteriors and their variational approximations," *The Journal of Machine Learning Research*, 23.147 (2022): 6579-6629. Available online: arxiv.org/2104.08324.

   This paper is representative of my work on the robustness to model misspecification in Bayesian procedures using posterior tempering. This work rigorously characterizes the robustness to model misspecification of α-posteriors, a type of generalized posterior distribution that has been shown empirically to demonstrate nice robustness properties, and of variation approximations to α-posteriors. We demonstrate that optimized variational approximations to α-posteriors enjoy additional robustness beyond optimized α-posteriors themselves in certain settings, and we believe this provides an additional rational for the use of variational inference due to its robustness, beyond solely for its nice computational properties.

6. J. Barbier, N. Macris, and **C. Rush**, "All-or-nothing statistical and computational phase transitions in sparse spiked matrix estimation," *Advances in Neural Information Processing Systems 33 (NeurIPS)*. 2020. Available online: arxiv.org/2006.07971.

   This paper is representative of my work in studying the computational complexity of statistical inference using approximate message passing (AMP). In particular, this paper uses AMP to derive the computational limits of a low-rank matrix estimation problem where the sparsity shrinks as the problem size grows. Low-rank matrix estimation is an important problem with numerous applications in image processing, principal component analysis, machine learning, DNA microarray data, and tensor decompositions. This is the first work in the literature to study AMP algorithms for problems with changing sparsity and the first work to study information-theoretic and computation limits of low-rank matrix estimation in the shrinking sparsity regime, a practical but challenging model of sparsity for high-dimensional problems.

7. Z. Bu, J. Klusowski, **C. Rush**, and W. Su, "Algorithmic analysis and statistical estimation of SLOPE via approximate message passing," *IEEE Transactions on Information Theory*, 67.1 (2021): 506-537. Available online: arxiv.org/1907.07502. (also appeared at NeurIPs 2019).

   This paper is representative of my work that uses the theory of approximate message passing (AMP) algorithms for the study of exact statistical properties of estimators in the linear proportional asymptotic regime. This work uses the exact asymptotics of AMP to study statistical properties of the sorted ℓ_2-penalized estimator (SLOPE) by providing a characterization of the limiting joint distribution of the SLOPE estimator $\widehat{\beta}$ with the truth β_0, which allows researchers to study properties like the mean-square error, the Type I or Type II error, the sparsity, or the bias, among others. To my knowledge, this is the first line of work using AMP to study a statistical estimator with a *non-separable* form of regularization, a theoretical challenge due to the statistical dependencies created, thereby paving the way for others to use this analysis as a roadmap to study a quite general class of penalized estimators.

8. **C. Rush** and R. Venkataramanan, "Finite sample analysis of approximate message passing," *IEEE Transactions on Information Theory*, 64.11 (2018): 7264-7286. Available online: arxiv.org/1606.01800. (JKW Student Paper Award Finalist, top 1% of papers at ISIT)

   This paper is representative of my work in the finite sample analysis of approximate message passing (AMP) algorithms. AMP is a class of efficient, iterative algorithms that have demonstrated success in many high-dimensional statistical estimation tasks, including generalized linear models and low-rank matrix estimation. Before this work, all other AMP analysis was asymptotic in nature. The non-asymptotic framework is powerful as it allows for the study of the algorithm's performance when the number of iterations grows with the problem dimension, which is not true for the asymptotic analysis, and this is often desired when analyzing the performance of an optimization algorithm. Moreover, AMP can be used to characterize statistical properties of a large class of estimators, like the LASSO, and in these settings we are mostly interested in finite sample, not asymptotic, guarantees.

9. **C. Rush**, A. Greig, and R. Venkataramanan, "Capacity-achieving sparse superposition codes with approximate message passing decoding," *IEEE Transactions on Information Theory*, 63.3 (2017): 1476-1500. Available online: arxiv.org/1501.05892.

   This paper is representative of my work in the design of engineering systems using statistical methods. This work studies a class of codes, referred to as sparse regression codes (SPARCs), that are used for transmitting information over the additive white Gaussian noise (AWGN) channel, which is a practical model of real world communication. Using the framework of sparse high-dimensional regression, SPARCs store information in the location of the non-zeros and channel decoding boils down to solving a structured high-dimensional linear regression problem. This work gives the first fully rigorous proof that approximate message passing decoding for SPARCs is capacity-achieving, mean it attains the information-theoretic limits of the problem asymptotically. While state-of-the-art coding for the AWGN channel shows good empirical performance, no such schemes have been proven to achieve capacity, making SPARCs the only efficient and provably capacity-achieving scheme for the AWGN channel.

10. B. Wexler, M. Iselli, S. Leon, W. Zaggle, **C. Rush**, A. Goodman, E. Ahmet, and E. Bo, "Cognitive priming and cognitive training: immediate and far transfer to academic skills in children," *Scientific Reports*, 6.1 (2016): 32859. Available online: nature.com/articles/srep32859. Press: `NPR`, `Yale News`, `WTNH`.

    This paper is representative of my work in applied statistics. Aiming to evaluate the ability of computer-presented brain training games to increase common intelligence measures like focus, self-control, and memory and to improve learning in a study of 583 second-grade children, this work provides evidence of immediate effects on learning and longer-term brain training to improve academic achievement. Because cognitive skills like attention, self-control, and memory are better predictors of math and reading achievement than I.Q. and because these skills are often compromised in children who have experienced trauma or are living in poverty, we hope that such low-cost brain-training solutions could help to improve educational outcomes for poor students.