- **Brian Neal and Tao He. "An adaptive multivariate kernel-based test for association with multiple quantitative traits in high-dimensional data."** MA thesis (in progress). Presented at the Joint Statistical Meetings, 2021. Planned submission to Genetic Epidemiology.

Abstract:

Identifying genetic features associated with a complex disease has benefited from recent advances in set-based and multi-trait testing methods. Jointly testing meaningful sets of genetic features (e.g., variants belonging to a gene set or genes belonging to a biological pathway) for association with a set of phenotypic disease traits can be more powerful than testing features or traits individually, thanks to aggregation of weak marginal signals and the ability to account for pleiotropy, trait correlations and interactions among features. Common challenges in set-based testing include high-dimensionality, where the number of genetic features vastly exceeds the sample size, and a large proportion of noise variables in the feature set. To address this, we develop an adaptive kernel-based multivariate association test (AMKAT) applicable to quantitative disease traits. AMKAT tests the genetic effect nonparametrically using function embeddings in reproducing kernel Hilbert spaces, avoiding the need to assume a particular distribution for the genetic effect or the disease traits; it includes a flexible kernel selection method that accommodates genetic effects with heterogeneous functional forms across disease traits; and it uses a fast, permutation-based filter, introduced in our paper, to screen the feature set and reduce noise and dimension. We compare the empirical performance of AMKAT to similar tests in simulations and using data from an Alzheimer's study. An R package containing a fast implementation of our test, written in C/C++, is presented in our paper.