

# Variable selection using Adaptive Non-linear Interaction Structures in High dimensions.

*Peter Radchenko*  
*IOM department, Marshall School*  
*University of Southern California*  
*Los Angeles, CA 90089*  
`peter.radchenko@marshall.usc.edu`

## Abstract

Numerous penalization based methods have been proposed for fitting a traditional linear regression model in which the number of predictors,  $p$ , is large relative to the number of observations,  $n$ . Most of these approaches assume sparsity in the underlying coefficients and perform some form of variable selection. Recently, some of this work has been extended to non-linear additive regression models. However, in many contexts one wishes to allow for the possibility of interactions among the predictors. This poses serious statistical and computational difficulties when  $p$  is large, as the number of candidate interaction terms is of order  $p$  squared. We introduce a new approach, “Variable selection using Adaptive Non-linear Interaction Structures in High dimensions” (VANISH), that is based on a penalized least squares criterion and is designed for high dimensional non-linear problems. Our criterion is convex and enforces the heredity constraint, in other words if an interaction term is added to the model, then the corresponding main effects are automatically included in the model. We provide theoretical conditions under which VANISH will select the correct main effects and interactions. These conditions suggest that VANISH should outperform certain natural competitors when the true interaction structure is sufficiently sparse. Detailed simulation results are also provided, demonstrating that VANISH is computationally efficient and can be applied to non-linear models involving thousands of terms while producing superior predictive performance over other approaches.