1135-62-1141 **Adam Gustafson**, **Matthew Hirn**, **Kitty Mohammed**, **Hariharan Narayanan** and **Jason Xu*** (jqxu@ucla.edu). *A Probably-Approximately-Correct Algorithm for Learning a $C^{1,1}(\mathbb{R}^d)$ Function from Noisy Samples.*

One means of fitting functions to high-dimensional data is by providing smoothness constraints. Recently, an efficient algorithm is proposed by Herbert-Voss, Hirn and McCollum that constructs $\hat{f}$ and yields $\mathrm{Lip}(\nabla\hat{f})$ solving the following function approximation problem: given a finite set $E \subset \mathbb{R}^d$ and a function $f : E \to \mathbb{R}$, interpolate the given information with a function $\hat{f} \in C^{1,1}(\mathbb{R}^d)$, the class of first-order differentiable functions with Lipschitz gradients, such that $\hat{f}(a) = f(a)$ for all $a \in E$, and the value of $\mathrm{Lip}(\nabla\hat{f})$ is minimal.

We address statistical aspects when reconstructing $\hat{f}$ from noisy data: we observe samples $y(a) = f(a) + e_a$ for $a \in E$, where e_a is an IID noise term. We obtain uniform bounds relating the empirical risk and true risk over the class $\mathcal{F} = \{f \in C^{1,1}(\mathbb{R}^d) \mid \mathrm{Lip}(\nabla f) \leq \widetilde{M}\}$, where the quantity $\widetilde{M}$ grows with the number of samples at a rate governed by the metric entropy of the class $C^{1,1}(\mathbb{R}^d)$. We provide a solution method via Vaidya's algorithm, supporting our results via numerical experiments on simulated data. (Received September 20, 2017)