Extracting functional dependence from sparse data and dimensionality reduction for potential energy surfaces

Sergei Manzhos[1], Koichi Yamashita
Department of Chemical Systems Engineering, University of Tokyo, Tokyo, Japan

Tucker Carrington
Department of Chemistry, Queen's University, Kingston, ON, Canada

Abstract

We address the problem of the reconstruction of a continuous potential energy surface (PES) from sparse *ab initio* data. PES govern reaction dynamics, and the availability of a continuous PES function greatly simplifies molecular modelling. However, for dimensions $D>6$, PES functions are not available for most molecular and reactive systems, such as polyatomic molecules on surfaces, the modelling of which is important for catalysis. The reason is the difficulty of constructing a function from very sparse (down to ~2-3 data per dimension) *ab initio* samples which are expensive to compute.
We develop Neural Network (NN) based methods to construct such functions from sparse data. The function is represented as one or a sum of N component functions depending on $d<D$ original or new coordinates. NN are used as building blocks that approximate the component functions. For a given density of data, only component functions with $d<dmax$ can reliably be recovered (or only coupling up to a certain order can be recovered if original coordinates are used). This justifies using functions of lower dimensionality.
We then make the component functions depend on $d<D$ new, adapted coordinates. The coordinate transformation (we like a linear transformation, but it can also be non-linear) is done by a NN and is adjusted during the fit for the best approximation error. When $Nd<D$, a dimensionality reduction regime obtains. The resulting dimensionality reduction method has properties of both traditional linear techniques, in that the coordinate transformation is linear, and of NN autoencoders, in that it is optimised during NN training by seeking the best fit to data. An important difference between this method and principal component methods is that no constraints are imposed on the possible transformations, which increases the efficiency of the embedding. The method is illustrated by fitting a PES for the reaction of nitrous oxide with a copper surface.

---

1. presenting author, email: Sergei.Manzhos@gmail.com