Comparison of Cluster Expansion Fitting Algorithms for Interactions at Surfaces

Laura M. Herder, Jason M. Bray, William F. Schneider

PII: S0039-6028(15)00054-0
DOI: doi: 10.1016/j.susc.2015.02.017
Reference: SUSC 20444

To appear in: *Surface Science*

Received date: 14 December 2014
Revised date: 25 February 2015
Accepted date: 27 February 2015

Please cite this article as: Laura M. Herder, Jason M. Bray, William F. Schneider, Comparison of Cluster Expansion Fitting Algorithms for Interactions at Surfaces, *Surface Science* (2015), doi: 10.1016/j.susc.2015.02.017

Comparison of Cluster Expansion Fitting Algorithms for Interactions at Surfaces

Laura M. Herder[a], Jason M. Bray[b], William F. Schneider[a,b,1]

[a] *Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN, USA*
[b] *Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN, USA*

Abstract

Cluster expansions (CEs) are Ising-type interaction models that are increasingly used to model interaction and ordering phenomena at surfaces, such as the adsorbate-adsorbate interactions that control coverage-dependent adsorption or surface-vacancy interactions that control surface reconstructions. CEs are typically fit to a limited set of data derived from density functional theory (DFT) calculations. The CE fitting process involves iterative selection of DFT data points to include in a fit set and selection of interaction clusters to include in the CE. Here we compare the performance of three CE fitting algorithms—the MIT Ab-initio Phase Stability code (MAPS, the default in ATAT software), a genetic algorithm (GA), and a steepest descent (SD) algorithm—against synthetic data. The synthetic data is encoded in model Hamiltonians of varying complexity motivated by the observed behavior of atomic adsorbates on a face-centered-cubic transition metal close-packed (111) surface. We compare the performance of the leave-one-out cross-validation score against the true fitting error available from knowledge of the hidden CEs. For these systems, SD achieves lowest overall fitting and prediction error independent of the underlying system complexity. SD also most accurately predicts cluster interaction energies without ignoring or introducing extra interactions into the CE. MAPS achieves good results in fewer iterations, while the GA performs least well for these particular problems.

Keywords: cluster expansion, adsorbate interactions, surface ordering, genetic algorithm, MAPS, steepest descent

1. Highlights:

- Cluster expansion (CE) represents energy of an arrangement of adsorbates in terms of two-body, three-body, . . . interactions
- Interaction terms typically selected through fitting to DFT database

*corresponding author
Email address: wschneider@nd.edu (William F. Schneider)