Downloaded from orbit.dtu.dk on: Apr 28, 2024

DTU Library

mBEEF

An accurate semi-local Bayesian error estimation density functional

Wellendorff, Jess; Lundgård, Keld Troen; Jacobsen, Karsten Wedel; Bligaard, Thomas

Published in:
The Journal of Chemical Physics

Link to article, DOI:
10.1063/1.4870397

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Wellendorff, J., Lundgård, K. T., Jacobsen, K. W., & Bligaard, T. (2014). mBEEF: An accurate semi-local Bayesian error estimation density functional. *The Journal of Chemical Physics*, *140*(14), 144107. https://doi.org/10.1063/1.4870397

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

mBEEF: An accurate semi-local Bayesian error estimation density functional

Jess Wellendorff, Keld T. Lundgaard, Karsten W. Jacobsen, and Thomas Bligaard

Citation: The Journal of Chemical Physics **140**, 144107 (2014); doi: 10.1063/1.4870397
View online: http://dx.doi.org/10.1063/1.4870397
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/140/14?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in

The role of van der Waals forces in water adsorption on metals
J. Chem. Phys. **138**, 024708 (2013); 10.1063/1.4773901

Adsorption of transition-metal atoms on boron nitride nanotube: A density-functional study
J. Chem. Phys. **125**, 044711 (2006); 10.1063/1.2218841

Characterization of methoxy adsorption on some transition metals: A first principles density functional theory study
J. Chem. Phys. **122**, 044707 (2005); 10.1063/1.1839552

General trends in CO dissociation on transition metal surfaces
J. Chem. Phys. **114**, 8244 (2001); 10.1063/1.1372512

Field-dependent chemisorption of carbon monoxide and nitric oxide on platinum-group (111) surfaces: Quantum chemical calculations compared with infrared spectroscopy at electrochemical and vacuum-based interfaces
J. Chem. Phys. **113**, 4392 (2000); 10.1063/1.1288592

AIP Journal of Applied Physics

Journal of Applied Physics is pleased to announce André Anders as its new Editor-in-Chief

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

mBEEF: An accurate semi-local Bayesian error estimation density functional

Jess Wellendorff,[1,2,a)] Keld T. Lundgaard,[1,3] Karsten W. Jacobsen,[3] and Thomas Bligaard[1,2]

[1]*SUNCAT Center for Interface Science and Catalysis, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, USA*
[2]*Department of Chemical Engineering, Stanford University, Stanford, California 94305, USA*
[3]*Center for Atomic-scale Materials Design (CAMD), Department of Physics, Building 307, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

(Received 30 January 2014; accepted 24 March 2014; published online 11 April 2014)

We present a general-purpose meta-generalized gradient approximation (MGGA) exchange-correlation functional generated within the Bayesian error estimation functional framework [J. Wellendorff, K. T. Lundgaard, A. Møgelhøj, V. Petzold, D. D. Landis, J. K. Nørskov, T. Bligaard, and K. W. Jacobsen, Phys. Rev. B **85**, 235149 (2012)]. The functional is designed to give reasonably accurate density functional theory (DFT) predictions of a broad range of properties in materials physics and chemistry, while exhibiting a high degree of transferability. Particularly, it improves upon solid cohesive energies and lattice constants over the BEEF-vdW functional without compromising high performance on adsorption and reaction energies. We thus expect it to be particularly well-suited for studies in surface science and catalysis. An ensemble of functionals for error estimation in DFT is an intrinsic feature of exchange-correlation models designed this way, and we show how the Bayesian ensemble may provide a systematic analysis of the reliability of DFT based simulations. © *2014 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4870397]

I. INTRODUCTION

Electronic structure theory offers key insights into the properties of materials, chemical reactions, and biomolecules. Kohn-Sham density functional theory[1,2] (KS-DFT) has proven a powerful framework for electronic structure studies,[3] particularly due to a favorable tradeoff between the computational speed and accuracy that can be obtained within this theory. Density functional methods have over the past decade reached a level of maturity where they can be applied not just in detailed theoretical studies of a given material, but be used to search for novel materials for technologically relevant applications in materials science[4–10] and chemical engineering.[11–16] Such studies often take a screening approach where massive amounts of DFT calculations are performed using efficient semi-local approximations for the KS exchange-correlation (XC) energy and potential. These include generalized gradient approximations (GGAs) and recent meta-GGA (MGGA) functionals.[17,18]

The reliability of semi-local density functional approximations (DFAs) is, however, unfortunately not universal. No such single functional appears to offer sovereign accuracy with zero bias in prediction of materials properties across the board of condensed matter and chemistry.[19–22] The GGA and MGGA exchange-correlation model spaces are flexible but incomplete and cannot accommodate an approximation that represents the exact XC functional in all aspects of practical importance. The result is an exchange-correlation model compromise on accuracy between different chemical and materials properties. However, semi-local DFT remains a favorite workhorse method within many research areas, so useful XC model compromises are warranted. Semi-empirical optimization lends itself well as a method for finding reasonably accurate compromises, but will never completely eliminate DFT errors. Even so, recent developments[23,24] clearly indicate that advanced machine learning methods have great potential for generating accurate density functionals. The BEEF class of functionals combines machine learning with a Bayesian point of view to generalize the fitting procedure for XC functionals, thereby allowing for estimation of the errors on calculated quantities. The traditional assumption underlying functional fitting is that a "best-fit" exchange-correlation model fitted to a suitable set of systems might be transferable, meaning that it hopefully calculates the properties well for systems not included in the training data. The generalization of this concept, which underlies BEEF-type functionals, is, that if one defines an "optimal" ensemble of exchange-correlation models, such that the ensemble on average reproduces errors on the training data, then the errors predicted by a well-constructed ensemble could be transferable. The ensemble can then be used to estimate computational uncertainties on calculations for systems not included in the fitted data set.

We have in Ref. 25 established a semi-empirical framework for developing model-compromise optimized density functionals with error ensembles as a practical implementation of the ideas proposed in Refs. 26 and 27. That study led to the first practically useful Bayesian error estimation functional, the BEEF-vdW, containing a somewhat expensive nonlocal correlation term. This BEEF framework uses machine learning tools to find the optimal compromise between model

a)Electronic mail: jewe@slac.stanford.edu

 © 2014 AIP Publishing LLC

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

complexity and model accuracy for a fitted general-purpose DFA in a highly flexible exchange-correlation model space. It furthermore uses ideas from Bayesian statistics[26] to construct an ensemble of XC functionals directly from the cost function that was minimized to find the optimally accurate and transferable exchange-correlation functional. This subsequently allows for fast and systematic error estimates on simulated quantities, as the ensemble is applied non-selfconsistently on the electron density that results from utilizing the optimally fitted functional. A number of surface science and catalysis studies[28–31] have successfully applied the BEEF-vdW functional, and have demonstrated significant improvements over traditionally used GGAs[32] for similar studies.

We here take the BEEF development an important step forward by considering a meta-GGA exchange model space and refine the approach to XC model selection. This semi-local functional allows studies of larger and more complex systems than the BEEF-vdW, since the non-local correlation term has been eliminated. We shall show that the endured loss of accuracy, even for hydrogen-bonded systems, is rather limited. This work thus establishes the currently most versatile error estimation functional, particularly useful for systems that are not dominated by long-range dispersion interactions. We first illustrate the model compromise of semi-local DFT. A refined BEEF model selection procedure for addressing the model compromise is then introduced and applied to MGGA exchange, and the resulting density functional (mBEEF) is subsequently benchmarked. Finally, we illustrate the BEE approach to error estimation for a materials property in DFT by analyzing the adsorption-site preference of CO adsorption on late transition metal surfaces.

II. EXCHANGE-CORRELATION MODEL COMPROMISE

The Perdew-Burke-Ernzerhof (PBE)[33] approximation from 1996 has become a default GGA in many branches of the computational materials science research area. However, the hundreds of GGA functionals reported in literature since then clearly indicate that no GGA can be considered truly universal. The PBEsol[34] modification of PBE, for example, predicts bulk lattice constants remarkably well but severely overestimates molecular bond energies, while the RPBE[32] revision of the PBE functional describes covalent chemistry well at the expense of overestimated lattice constants and underestimated cohesive energies.[25,34,35] This is the topography of the XC model compromise in GGA DFT. Inclusion of non-local exact exchange has become a popular approach to obtaining significantly improved model compromises. Such hybrid functionals may also offer better agreement of KS one-electron transition energies with experimental band gaps.[36,37] However, even with screening of the long-range part of the exact exchange potential, the computational cost is significantly increased from that of GGA, particularly for extended systems such as those relevant to surface chemistry.

Meta-GGA density functionals[38–44] augment the GGA model space of electron density and its first-order gradient by including also the second-order density gradient[38] or the orbital kinetic energy density (KED) of the occupied KS eigenstates.[39] Importantly, an electronic structure with vanishing electron density gradient is in MGGA not necessarily modeled as a uniform electron gas (UEG). The UEG limit on exchange for small density gradients does in general not appear to be compatible with semi-local DFAs fully optimized for prediction of molecular bond energies.[25,27,45] Special-purpose GGAs may be designed by modification of known GGA forms, as in the cases of PBEsol and RPBE. The main purpose of applying MGGA exchange in the BEEF framework is, however, the prospect of better XC model compromises than with GGAs at a very modest increase in computational cost.[42]

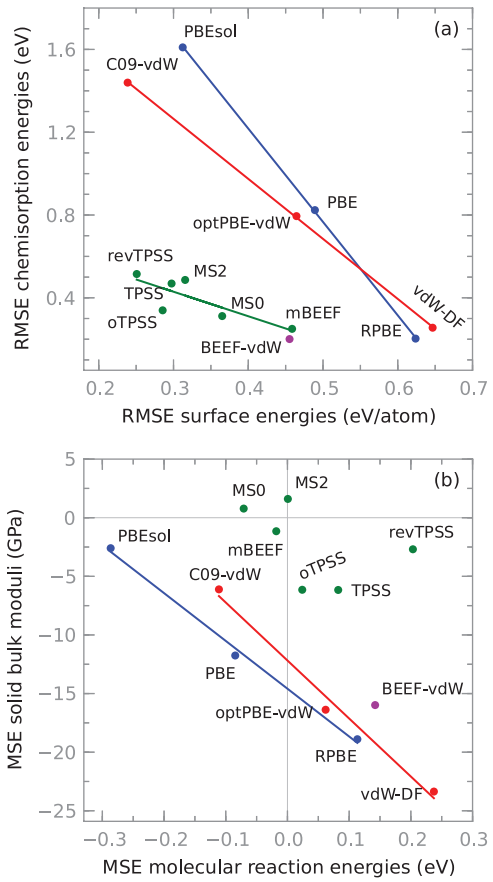

FIG. 1. Bivariate analyses of DFT prediction errors on chemical and materials properties. (a) Root-mean-squared errors on the CE27a chemisorption energies against those on the SE30 surface energies. (b) Mean-signed errors on the BM32 solid bulk moduli against those on the RE42 molecular reaction energies. Straight lines are fits through the GGA (blue), meta-GGA (green), and vdW-DF type (red) data. The meta-GGA points are closest to origo in both panels, indicating improved possibilities for exchange-correlation model compromises in the meta-GGA model space as compared to varying the exchange approximation in the GGA and vdW-DF ones.

We illustrate this point in Fig. 1, where a broad selection of mostly GGAs, MGGAs, and vdW-DF[46] type functionals are applied in calculations of four different quantities; chemisorption energies of small molecules on close-packed transition metal facets, surface energies of various facets, solid bulk moduli, and gas-phase reaction energies. These properties are represented by the CE27a, SE30, BM32, and RE42 data sets, respectively, all discussed in more detail later. The tested GGAs are PBEsol, PBE, and RPBE, while the literature MGGAs are TPSS,[41] revTPSS,[42] oTPSS,[47] MS0,[44] and

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

MS2.[48] Note that three of the van der Waals functionals, vdW-DF, optPBE-vdW,[49] and C09-vdW,[50] are equivalent except for the choice of PBE-like exchange. Figure 1(a) shows root-mean-squared errors (RMSEs) on the CE27a chemisorption energies against those on SE30 surface energies. The points within each class of XC model space fall approximately on straight lines, illustrating the trade-off one is forced to make between accurate adsorbate–surface bond strengths and surface stabilities. However, the MGGA model space offers the most attractive compromises; the green line in Fig. 1(a) is significantly closer to the origo. This is supported by Fig. 1(b), in which mean-signed errors (MSEs) on predicted BM32 bulk moduli are plotted against those on RE42 molecular reaction energies. The relations between mean errors are again approximately linear and the MGGA points fall closest to origo, though not all on the same straight line.

The bivariate prediction error analysis in Fig. 1 confirms the conjectures from earlier studies[51,52] that the exchange-correlation model compromise of typical XC models lead directly to a trade-off between the systematic errors on various chemical and materials properties. Simple but efficient approaches to optimizing density functionals with respect to this trade-off are therefore core components of BEEF-class functional developments.

III. EXCHANGE MODEL SPACE

The spin-unpolarized meta-GGA exchange energy we write as the usual[41] integral over the uniform electron gas exchange energy density $\epsilon_x^{\mathrm{UEG}}$ scaled with a semi-local MGGA exchange enhancement factor F_x,

$$E_x = \int n\epsilon_x^{\mathrm{UEG}}(n)\, F_x(n, \nabla n, \tau)\, d\boldsymbol{r}, \tag{1}$$

where $n = n(\boldsymbol{r})$ is the local electron density, ∇n the density gradient, and the semi-local kinetic energy density $\tau = \frac{1}{2}\sum_{i,\sigma} |\nabla \Psi_{i,\sigma}|^2$ is summed over spins σ and state labels i of the KS eigenstates $\Psi_{i,\sigma}$. Atomic units are used throughout. The exchange enhancement factor we furthermore express in terms of dimensionless electronic structure parameters; the reduced density gradient $s = |\nabla n|/(2k_{\mathrm{F}}n)$, where $k_{\mathrm{F}} = (3\pi^2 n)^{\frac{1}{3}}$, and the reduced kinetic energy density $\alpha = (\tau - \tau^{\mathrm{W}})/\tau^{\mathrm{UEG}}$, where $\tau^{\mathrm{W}} = |\nabla n|^2/8n$ and $\tau^{\mathrm{UEG}} = (3/10)(3\pi^2)^{\frac{2}{3}} n^{\frac{5}{3}}$.

The MGGA exchange enhancement factor we therefore write $F_x(n, \nabla n, \tau) = F_x(s, \alpha)$, and expand it in products P of Legendre polynomials B depending on s and α through transformed quantities t_s and t_α:

$$t_s(s) = \frac{2s^2}{q + s^2} - 1, \tag{2}$$

$$t_\alpha(\alpha) = \frac{(1-\alpha^2)^3}{1 + \alpha^3 + \alpha^6}, \tag{3}$$

$$P_{mn} = B_m(t_s) B_n(t_\alpha), \tag{4}$$

$$F_x(s, \alpha) = \sum_{m=0}^{M} \sum_{n=0}^{N} a_{mn} P_{mn}. \tag{5}$$

For the mBEEF fit we chose values of $M = N = 7$, giving $Z = (M+1) \times (N+1) = 64$ exchange basis functions with expansion coefficients a_{mn}, which more than exhaust the present exchange model space. Both t_s and t_α are confined to $[-1, +1]$. With $q = \kappa/\mu = 0.804/(10/81) = 6.5124$, transformation t_s is a Padé approximant to the PBEsol $F_x(s)$, while t_α is inspired by the MS0 exchange.[44]

Denoting by E_x^{mn} the exchange energy corresponding to P_{mn}, the full exchange-correlation energy is written

$$\begin{aligned} E_{xc} &= \sum_{m,n}^{M,N} a_{mn} E_x^{mn} + E_c^{\mathrm{PBEsol}} \\ &= \boldsymbol{x}\boldsymbol{a}^T + E_c^{\mathrm{PBEsol}}, \end{aligned} \tag{6}$$

where $\boldsymbol{x}$ is the vector of exchange basis function energy contributions for the system in question and the vector $\boldsymbol{a}$ contains the exchange model expansion coefficients in Eq. (5). The training data in $\boldsymbol{x}$ were obtained from PBEsol ground-state electron densities and single-particle eigenstates.

IV. TRAINING DATA SETS

Five significantly different sets of target chemical and materials properties are used in exchange model training. Most of the sets were also applied in Ref. 25, but are here updated or slightly modified.

The G3/99 molecular formation energies[53] and the related RE42 reaction energies[25] represent gas-phase chemistry. Both data sets are normalized in model training such as to approximately bring all data within each set on an equal footing, see Ref. 54 for details. Surface chemistry we represent by the CE27a chemisorption energies of simple adsorbates on late transition metal surfaces adapted from previous work.[25,55] Solid bulk energetics is represented by cohesive energies in the Sol54Ec set, and bulk structures by the derivatives of cohesive energies with respect to crystal volume around equilibrium.[82] Note that solid Pb is excluded from both data sets in model training, see Ref. 55. Experimental lattice constants are from the Sol58LC set.[55]

Density functional calculations are performed using GPAW,[56,57] an open-source DFT code implementing the projector augmented-wave method,[58] and the open-source ASE[59] package. GPAW can represent the Kohn-Sham equations on a real-space uniform grid as well as in a plane-wave expansion. Structural relaxations follow the prescriptions in Ref. 25 and use grid-point spacings of 0.15–0.16 Å. Chemisorption energies are calculated using a $(10 \times 10 \times 1)$ Monkhorst-Pack[60] $\boldsymbol{k}$-point mesh. Bulk calculations are done in plane-wave mode using a 1000 eV plane-wave energy cutoff and a $(16 \times 16 \times 16)$ $\boldsymbol{k}$-point mesh. Lattice constants and bulk moduli are computed by fitting the SJEOS equation of state[61] to 9 electronic total energies sampled at lattice constants spanning $\pm 1\%$ around the apparent equilibrium one.

V. EXCHANGE MODEL SELECTION

We seek a general-purpose density functional for surface chemistry studies with built-in error estimates. With the

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

flexible exchange model space defined in Eq. (5), maximizing not only performance on the training data sets but also transferability to unseen data is essential. To this end we use ideas from machine learning[62,63] and extend on developments in Refs. 25 and 27. We formulate the optimization problem in terms of a regularized cost function to be minimized for the optimum (mBEEF) exchange coefficient vector $\hat{\boldsymbol{a}}_0$. We also generate a Bayesian error estimation ensemble in terms of model fluctuations around $\hat{\boldsymbol{a}}_0$. Several aspects of model selection are most conveniently introduced in terms of fitting a single set of data.

A. Cost function and BEE ensemble

We parametrize an exchange model on a single training set by minimizing a cost function C consisting of a squared-error loss function L and a regularizer R,

$$\begin{aligned} C(\boldsymbol{a};\omega) &= L(\boldsymbol{a}) + R(\boldsymbol{a};\omega) \\ &= \boldsymbol{q}^T\boldsymbol{q} + \omega^2\boldsymbol{b}^T\boldsymbol{b}, \end{aligned} \tag{7}$$

which depends parametrically on the regularization strength $\omega \geq 0$. The residual vector of training errors is $\boldsymbol{q} = \boldsymbol{X}\boldsymbol{a} - \boldsymbol{y}$, where matrix $\boldsymbol{X}$ contains all exchange basis function contributions and $\boldsymbol{y}$ is a vector of targets. The vector $\boldsymbol{b}$ is a suitable affine mapping of $\boldsymbol{a}$, which we shall define later in Eq. (15). The minimal cost solution $\boldsymbol{a}_0$ for a given choice of ω is easily found, see Ref. 25. In the language of Bayesian statistics, minimizing C over $\boldsymbol{a}$ given ω is equivalent to maximizing the posterior probability for the model parameters given a prior expectation.[62,63] The regularizer in Eq. (7) imposes the prior expectation for $\boldsymbol{a}_0$ as a penalty term of variable strength. The effect is parameter shrinkage, a standard machine learning method for dealing with ill-posed regression problems and avoiding over-fitting by controlling the model complexity.[62,63] Any ordinary least-squares (OLS) regression solution in a sufficiently large model space will contain a number of poorly determined parameters—parameters that vary wildly for small perturbations of the training data—a clear indication of over-fitting. Singular value decomposition of regularized cost functions of the form (7) shows how the regularizer adds curvature to such weak modes in L and essentially freezes them out of the fit.[27,63] Regularization is therefore used to tune the model complexity in order to enhance model generalization. It is then natural to introduce the notion of an effective number of model parameters θ,[62,63]

$$\theta(\omega) = \sum_m \frac{\nu_m}{\nu_m + \omega^2}, \tag{8}$$

where ν_m are the eigenvalues of $\boldsymbol{X}^T\boldsymbol{X}$. Note that $\theta(0) = Z$ recovers the OLS solution while $\theta(\infty) = 0$. We may think of θ as the number of cost function eigenmodes that are not significantly affected by regularization.[27,62]

The cost function is quadratic in $\boldsymbol{a}$ and can therefore around its minimum $C_0 = C(\boldsymbol{a}_0;\omega_0)$ be expressed in terms of the Hessian matrix $\boldsymbol{H} = \partial^2 C/\partial\boldsymbol{a}^T\partial\boldsymbol{a}$ and model perturbations $\delta\boldsymbol{a} = \boldsymbol{a} - \boldsymbol{a}_0$:

$$C(\boldsymbol{a}) = C_0 + \frac{1}{2}\delta\boldsymbol{a}^T\boldsymbol{H}\delta\boldsymbol{a}. \tag{9}$$

As in previous work[25–27,64] we define a probability distribution $\mathcal{P}$ for fluctuations $\delta\boldsymbol{a}$ around $\boldsymbol{a}_0$. From $\mathcal{P}$ we draw ensembles of perturbed density functionals used for estimating errors on DFT predictions. That the cost function can be assumed to represent the probability of a model given the data is intrinsically a Bayesian idea with no analog in frequentist statistics. We require the mean expectation value of predictions by ensemble models $\boldsymbol{a}' = \boldsymbol{a}_0 + \delta\boldsymbol{a}$ to reproduce the mean prediction error of $\boldsymbol{a}_0$:

$$\sum_j \langle(\delta q_j)^2\rangle_k = \sum_j (\Delta q_j)^2, \tag{10}$$

where $\langle\ldots\rangle_k$ indicates the average over $k \gg 1$ ensemble models. The sums are over j training data while δq_j and Δq_j are prediction errors by $\boldsymbol{a}'$ and $\boldsymbol{a}_0$, respectively. Following Refs. 27 and 64 the probability $\mathcal{P}$ is written

$$\mathcal{P} \propto \exp(-C/T), \tag{11a}$$

$$T = 2C_0/\theta, \tag{11b}$$

where the ensemble temperature T scales the model fluctuations such that Eq. (10) is satisfied. The temperature is in Eq. (11b) expressed in terms of the minimized cost and θ, the effective number of model parameters.

In practical applications of Bayesian error estimation we sample the distribution $\mathcal{P}$. An ensemble matrix $\boldsymbol{\Omega}$ is generated by scaling the inverse Hessian with the ensemble temperature:

$$\boldsymbol{\Omega} = T\boldsymbol{H}^{-1}. \tag{12}$$

Ensemble perturbations $\delta\boldsymbol{a}_k$ are then computed as

$$\delta\boldsymbol{a}_k = \boldsymbol{V}\cdot\boldsymbol{D}\cdot\boldsymbol{u}_k, \tag{13}$$

where matrix $\boldsymbol{V}$ contains the eigenvectors of $\boldsymbol{\Omega}$, matrix $\boldsymbol{D}$ is diagonal and contains the square root of the corresponding eigenvalues, and $\boldsymbol{u}_k$ is a random vector with normal distributed components of zero mean and a spread of 1. The Bayesian error estimate on any DFT prediction from total-energy differences, σ_{BEE}, is then simply related to the variance of $k \gg 1$ non-self-consistent predictions p_k by $\delta\boldsymbol{a}_k$:

$$\sigma_{\mathrm{BEE}} = \sqrt{\mathrm{Var}(\boldsymbol{p})} = \langle\boldsymbol{p}^T\boldsymbol{p}\rangle^{\frac{1}{2}}, \tag{14}$$

where $\boldsymbol{p}$ is a vector of ensemble predictions and the last equality is strictly true for $k \to \infty$, where $\langle\boldsymbol{p}\rangle^2 = 0$.

B. Tikhonov regularization

Smooth exchange enhancement factors are aesthetically pleasing and computationally convenient. Indeed, it was observed in Ref. 27 that smoothness of the enhancement factor is of key importance to obtaining a exchange functional that is transferable to systems not included in the training data. As in Ref. 25 we apply a Tikhonov regularizer R in the cost function Eq. (7) to impose our preference for smooth parametrizations of the MGGA $F_x(s,\alpha)$. This was in Ref. 25 achieved by defining R such that the cost function penalty on a GGA exchange model space was proportional to the integral of the squared second derivative of the exchange basis over its domain. Such

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

regularizers shrink model coefficients in a "smooth" space $\boldsymbol{b}$ defined by the Tikhonov matrix $\boldsymbol{\Gamma}$,

$$\boldsymbol{b} = \boldsymbol{\Gamma}(\boldsymbol{a} - \boldsymbol{a}_p), \tag{15}$$

where the prior vector $\boldsymbol{a}_p$ is an origo in model space. The prior is thus the resulting fitted model at infinite regularization strength, where all model deviations away from $\boldsymbol{a}_p$ are quenched by the regularizer. We here generalize the exchange regularizer of Ref. 25 to the MGGA exchange model space. The Tikhonov matrix we therefore define from the overlaps of a scaled Laplacian $\widetilde{\nabla}^2$ acting on the exchange basis functions $P(t_s, t_\alpha)$,

$$\begin{aligned} \widetilde{\nabla}^2 &= \frac{\partial^2}{\partial t_s^2} + \lambda \frac{\partial^2}{\partial t_\alpha^2}, \\ \boldsymbol{\Gamma}^2_{mnkl} &= \int_{-1}^{1}\int_{-1}^{1} dt_s\, dt_\alpha \widetilde{\nabla}^2 P_{mn} \widetilde{\nabla}^2 P_{kl}, \end{aligned} \tag{16}$$

where λ scales the regularization penalty between polynomials in t_s and t_α. We choose $\lambda = 10^2$, which in numerical tests seems to give a reasonable trade-off between smoothness along s and α. The elements of $\boldsymbol{\Gamma}^2$ grow as the polynomial order of the basis increases, and model regularization therefore preferentially shrinks the more oscillatory components in $F_x(s, \alpha)$. The prior vector $\boldsymbol{a}_p$ in Eq. (15) is chosen such that infinite regularization strength yields $F_x(s, 1) = 1$ for all s and half the MS0 exchange along the $F_x(0, \alpha)$ model space direction.

C. Exchange model compromise

The XC model compromises illustrated in Fig. 1 indicate the existence of significant constraints on the performance of semi-local general-purpose density functional approximations: A gain in accuracy on one chemical or materials property is typically associated with a loss of accuracy on a different property. Simultaneously minimizing the prediction errors on several different properties in a transferable manner is therefore a multi-objective (or Pareto) optimization problem. In such Pareto-optimizations, where one cannot *a priori* infer a strict measure of the relative importance of the individual objectives, there is still one set of solutions that are superior to all other. This is the Pareto set of non-dominated solutions, or the set for which one cannot improve one quality without making another quality worse. Among the Pareto-optimal set of solutions one still has a choice in what importance is given to the different qualities. In Ref. 25 a simple but effective approach to this type of problem was developed in the context of density functional fitting, based on minimizing the product of cost functions for the individual training sets including their individual regularizations. The logic underlying this choice is to find a solution among all the Pareto-optimal solutions where the relative improvement of one property leads to a similar relative deterioration of the other properties. The product of cost functions achieves exactly this, if the cost represents the qualities to be optimized. Here we refine that approach by considering a fully regularized cost function for all training data. This corresponds to considering the squared residuals a better measure of quality than the individually minimized (and regularized) cost functions. It is our impression that this improvement offers slightly better fits, and it has the added benefit that the Bayesian interpretation of the statistics is significantly more straightforward, since the functional results from one fit to all data rather than separate fits to each chemical or materials property.

The new starting point for dealing with the exchange model compromise can then be stated as an objective function Φ:

$$\Phi(\boldsymbol{a}; \omega) = \Pi_i L_i(\boldsymbol{a}) \times e^{R(\boldsymbol{a};\omega)}, \tag{17}$$

where L_i is the squared-residual loss function for data set i and the exponential a functional form for the prior expectation for the model parameters. Since minimizing Φ is equivalent to minimizing $\ln\{\Phi\}$ we define the following regularized cost function K for the exchange model compromise:

$$K(\boldsymbol{a}; \omega) = \ln\{\Phi\} = \sum_i \ln\{L_i(\boldsymbol{a})\} + R(\boldsymbol{a}; \omega). \tag{18}$$

The minimizing argument vector $\boldsymbol{a}_0(\omega)$, minimizing the objective function K given ω, is a vector that fulfills the zero-gradient condition

$$\frac{\partial K}{\partial \boldsymbol{a}} = 0 = \sum_i \frac{\partial \ln L_i}{\partial \boldsymbol{a}} + \frac{\partial R}{\partial \boldsymbol{a}} = \sum_i \frac{1}{L_i}\frac{\partial L_i}{\partial \boldsymbol{a}} + \frac{\partial R}{\partial \boldsymbol{a}}. \tag{19}$$

If K had been quadratic and positive definite, the existence of only a single solution vector $\boldsymbol{a}_0(\omega)$ would have been certain. This, however, does not appear to be a significant problem, at least not with the data sets we have fitted in the present study. Since the loss function, $L_i(\boldsymbol{a})$, and the regularizer, $R(\boldsymbol{a}; \omega)$, are both quadratic in $\boldsymbol{a}$, the zero-gradient condition above is very close to representing a traditional least-squares minimization problem, and we solve it by iterative least-squares minimization by casting it on the form of Eq. (7):

$$\begin{aligned} \widetilde{K}(\boldsymbol{a}; \omega, \boldsymbol{a}_*) &= \widetilde{L}(\boldsymbol{a}; \boldsymbol{a}_*) + R(\boldsymbol{a}; \omega) \\ &= \sum_i \frac{L_i(\boldsymbol{a})}{L_i(\boldsymbol{a}_*)} + R(\boldsymbol{a}; \omega) \\ &= \tilde{\boldsymbol{q}}^T \tilde{\boldsymbol{q}} + \omega^2 \boldsymbol{b}^T \boldsymbol{b}, \end{aligned} \tag{20}$$

with least-squares solution $\tilde{\boldsymbol{a}}$. This solution is then inserted for $\boldsymbol{a}_*$ iteratively, and convergence is reached in very few steps, when $\tilde{\boldsymbol{a}} = \boldsymbol{a}_*$. In that case, the model-compromise cost Eq. (20) reduces to $\widetilde{K} = N_D + \omega^2 \boldsymbol{b}^T \boldsymbol{b}$, with N_D the number of training data sets.

The concept of an effective number of model parameters, as defined in Eq. (8), applies equally well to $\widetilde{K}$, as does the definition of the Bayesian ensemble matrix in Eq. (12). Only the model complexity $\hat{\theta}$, corresponding to the globally optimum exchange model $\hat{\boldsymbol{a}}_0$, remains to be determined. This model should constitute a suitable trade-off between model bias and variance such that it generalizes well to properties outside the training sets.[63] We here apply a clustered leave-one-out cross validation estimator of the generalization error, Δ^2:

$$\Delta^2(\omega) = \frac{1}{N_D}\sum_{i=1}^{N_D} L_i(\tilde{\boldsymbol{a}}_i(\omega)), \tag{21}$$

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

where training set i has been excluded from $\widetilde{K}$ when determining $\tilde{\boldsymbol{a}}_i$.

In summary, we thus determine the optimal simultaneous fit to all training data in the protocol, $\hat{\boldsymbol{a}}_0$, by identifying the regularization strength $\hat{\omega}_0$ that minimizes the generalization error Δ^2. The corresponding exchange model complexity is $\hat{\theta}$, and Bayesian error estimates on materials property predictions by $\hat{\boldsymbol{a}}_0$ are obtained following Eqs. (11b)–(14).

VI. RESULTS

A. mBEEF density functional and BEE ensemble

Figure 2 shows a range of meta-GGA exchange enhancement factors obtained by minimizing Eq. (20) for increasing model complexities, i.e., for decreasing ω. The enhancement factors are neatly smooth along s (top panel) and α (bottom panel) for small θ, but develop increasingly non-smooth features when the exchange models are allowed to become more complex, particularly for $\theta > 12$. The optimum trade-off between performance and transferability, as determined by minimizing Δ^2, we find at $\hat{\theta} = 8.8$. This model we henceforth denote mBEEF exchange. It is indicated by solid black lines in Fig. 2. Note that the full mBEEF exchange-correlation functional uses PBEsol correlation, see Eq. (6), and that mBEEF exchange does not conform to the formal UEG limit. This appears to be a quite general feature of semi-local DFAs optimized for chemistry.[25,27,45] Consequently, the mBEEF enhancement of local exchange for a UEG-like electronic structure is $F_x(0, 1) = 1.037$, while for rapidly varying densities $F_x(\infty, 1) = 1.145$. The latter is a significantly lower exchange enhancement in the large gradient/small density regime than for most semi-local functionals.

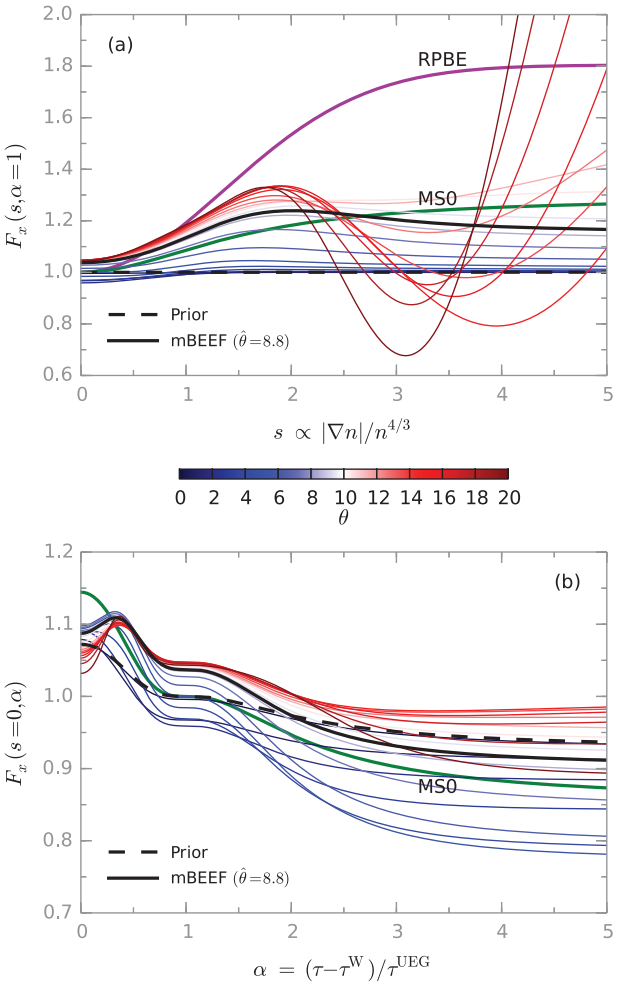

FIG. 2. Model-compromise optimized mBEEF type exchange enhancement factors for increasing number of effective parameters $\theta \in [0, 20]$. Blue lines indicate $\theta < 10$ and red lines $\theta > 10$. Solid black lines illustrate the chosen mBEEF $F_x(s, \alpha)$, dashed black lines the prior model. (a) Projections along s for $\alpha = 1$. Note that $\alpha = 1$ for a uniform electron gas, and that for this value of the reduced KED the MGGA $F_x(s, \alpha)$ is equivalent to a GGA exchange enhancement factor. (b) Projections along α for $s = 0$. All but the most constrained mBEEF type exchange functionals have a curved feature between the single-electron limit ($\alpha = 0$) and the UEG region ($\alpha \approx 1$).

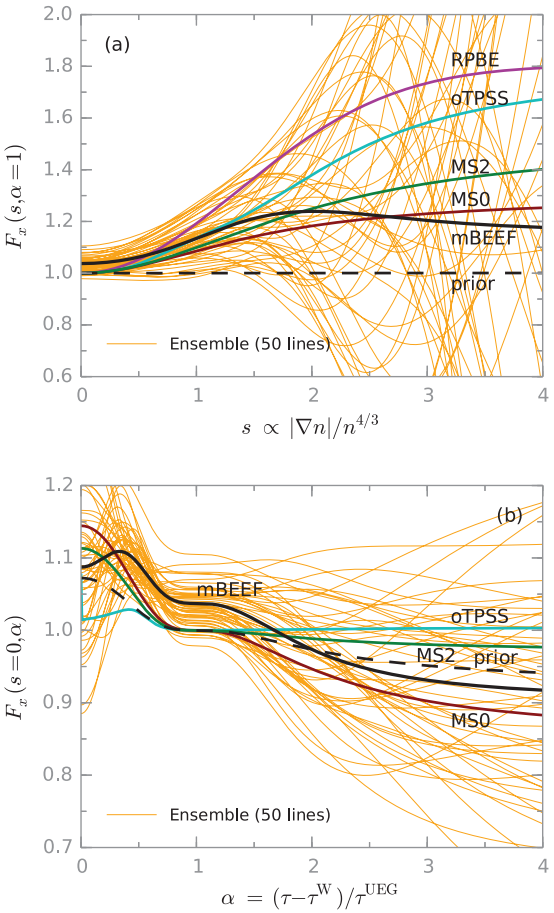

FIG. 3. Bayesian ensemble of exchange models (yellow) around the mBEEF (solid black). Standard GGA and MGGA exchange functionals are illustrated by colored lines along with the prior model. (a) Projections along s for $\alpha = 1$. (b) Projections along α for $s = 0$. The ensemble is rather constrained around the UEG limit $(s, \alpha) = (0, 1)$ in both panels, but spreads out significantly in other regions of exchange model space.

The mBEEF error estimation exchange ensemble is illustrated in Fig. 3. Note how constrained the ensemble is around $(s, \alpha) = (0, 1)$, and that it clearly straddles the UEG limit in this point. The ensemble models spread out significantly for $(s, \alpha) > (2, 2)$, indicating that the functional form of the mBEEF $F_x(s, \alpha)$ is less constrained in this region of the MGGA electronic-structure parameter space. We reported similar findings in Ref. 25 for the large-s regime of the BEEF-vdW ensemble. The fact that the ensemble is very broad for

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

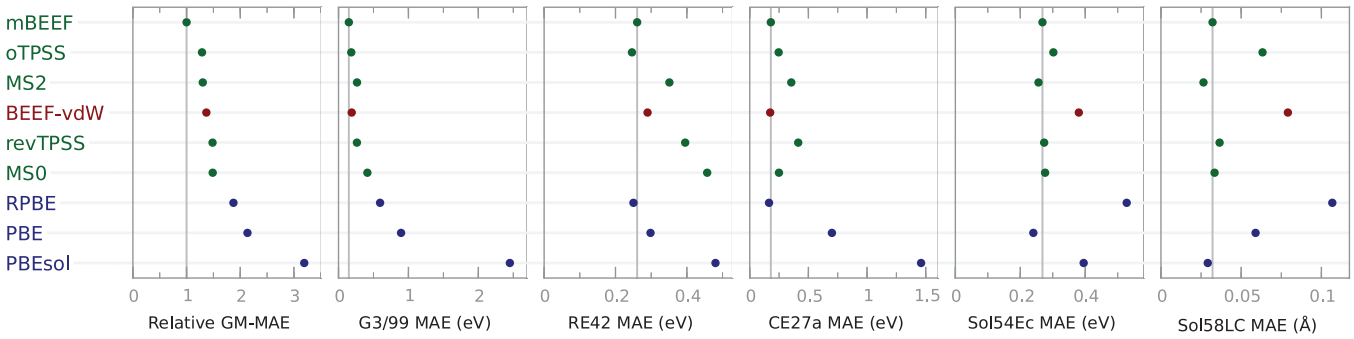

FIG. 4. Benchmark of mBEEF against popular or recent GGA (blue) and MGGA (green) density functionals in terms of mean absolute errors (MAEs) on predicting the chemical and materials properties represented by the 5 data sets applied in mBEEF training. BEEF-vdW is also included (red). The first panel ranks the tested density functionals according to the geometric mean of the MAEs, see Eq. (22). All data were obtained from self-consistent DFT calculations.

large reduced density gradients suggests that the decay of the mBEEF $F_x(s, 1)$ towards 1.145 for $s \to \infty$ is not imposed by the training data sets. Rather, the training data offer very little electronic-structure information for $(s, \alpha) > (2, 2)$, and the exchange model in this region therefore becomes strongly dominated by the prior model $\boldsymbol{a}_p$.

The mBEEF exchange expansion vector and error estimation ensemble matrix are available online[65] for easy implementation into DFT codes already implementing meta-GGA functionals.

B. Benchmark

Figure 4 illustrates a broad benchmark of some popular or recent GGA and MGGA density functionals in addition to mBEEF and BEEF-vdW. The first panel summarizes the benchmark by ranking the density functionals in terms of the geometric mean (GM) of mean absolute errors on the $N = 5$ training data sets applied in the mBEEF fitting,

$$\text{GM-MAE} = \left(\Pi_i^N MAE_i\right)^{1/N}, \tag{22}$$

relative to that of mBEEF. The remaining panels in the figure illustrate the actual MAEs for all considered functionals on each training set. The mBEEF exchange model compromise appears quite reasonable, as one would expect, since the functional was trained on this data: The MAE is among the three lowest for all five properties and presents a considerable improvement over the BEEF-vdW in predicting the lattice constants and cohesive energies of bulk solids, while not compromising the good description of the adsorbate–surface bond strengths in CE27a, which is almost on the level of the RPBE functional. In total, mBEEF simultaneously achieve very acceptable predictions within the five classes of chemical and materials properties. The relative GM-MAEs in the first panel of Fig. 4 have values of 1.3 or more for all other tested functionals. Note also the clustering of the density functionals into GGAs and MGGAs+BEEF-vdW when ranked according to this measure of overall performance. This is a direct consequence of improved possibilities for the XC model compromise.

C. Transferability

We shall now assess the mBEEF transferability by considering quantities outside the training data. Table I compares error statistics on the MB08-165 decomposition energies of artificial molecules, the BM32 bulk moduli, the SE30 surface energies, and 26 of the 27 binding energies of neutral and charged water clusters in the WATER27 benchmark set. The mBEEF functional appears to generalize reasonably well to prediction of properties not explicitly included in the training sets used to generate it. The decomposition energies and bulk moduli are on average predicted with only a limited systematic bias. The surface energies are on average underestimated. For this property mBEEF performs better than PBE and is nearly on par with MS0, but does not attain the accuracy of the TPSS-class functionals and MS2. As observed in Fig. 1, this may be due to mBEEF's focus on performing well for chemisorption energies. Interestingly, the water cluster binding energies are surprisingly well captured by mBEEF even though systems with significant noncovalent interactions were not included in the training data. Contrary to the two TPSS-type MGGAs, PBE, MS0, and MS2 also appear to describe this sort of hydrogen bonding well. Similar findings were reported in Refs. 44, 48, and 66. This suggests that the high accuracy of mBEEF for hydrogen bonding may to some extent be due to the use of a MS0-based form of the α-dependence in the exchange model space. We therefore concur with the hypothesis of Perdew *et al.*[66] that future high-performance van der Waals (vdW) density functionals might benefit greatly from optimized MS-based exchange.

We note in passing that mBEEF also correctly predicts the sequence of relative stabilities of the 4 isomers of the water hexamer included in the WATER26 set. Moreover, the MAE over the 4 isomers is less than 1 kcal/mol. Most semi-local DFAs agree much worse with benchmark quantum chemical calculations on these systems. According to literature, it usually takes highly specialized exchange-correlation functionals optimized for hydrogen bonding[67] or dedicated vdW functionals[49] to get the energetic ordering of water hexamers right. MS0 and MS2 correctly predict the ordering when the benchmark (B3LYP) structures are used, but fail upon structural relaxation, which destabilizes the "prism" isomer by 1.3 and 1.4 kcal/mol, respectively. This is not the case

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

TABLE I. Error statistics for different density functionals in predicting various chemical and materials properties not included in mBEEF training. Computed statistics are mean error (ME) or mean relative error (MRE) and their absolute counterparts.

	LSDA	PBEsol	PBE	RPBE	revTPSS	oTPSS	MS0	MS2	mBEEF	BEEF-vdW
					MB08-165 decomposition energies of artificial molecules[a] *(kcal/mol)*					
ME	15.4	7.6	1.4	− 4.9	− 7.3	− 2.3	− 11.0	− 9.3	0.1	− 2.0
MAE	19.9	12.7	9.0	11.3	13.2	6.8	18.4	14.6	8.1	12.2
					BM32 bulk moduli[b] *(%)*					
MRE	5.7	− 3.0	− 10.7	− 17.9	− 3.2	− 7.3	− 0.8	0.7	− 0.7	− 12.8
MARE	7.9	5.1	10.9	17.9	6.8	8.6	5.5	4.6	7.1	14.7
					SE30 surface energies[c] *(%)*					
MRE	− 7	− 13	− 26	− 35	− 6	− 11	− 18	− 12	− 22	− 21
MARE	14	17	26	35	12	15	20	17	23	23
					WATER26 binding energies of neutral and charged water clusters[d] *(%)*					
MRE	47.5	17.3	2.7	− 18.8	− 7.8	− 13.2	− 2.5	− 5.8	2.3	− 12.5
MARE	47.5	17.3	3.6	18.9	7.8	13.6	2.7	5.8	2.7	12.5

[a]Quantum chemical benchmark from Ref. 71.
[b]Thirty-two experimental bulk moduli from Refs. 36, and 72, all corrected for thermal contributions and zero-point phonon effects.
[c]Thirty experimental surface energies from Ref. 73.
[d]Quantum chemical benchmark from Ref. 74. This set was in Ref. 47 named WATER27, but we exclude here the last benchmark data point since it is a conformational energy difference rather than a binding energy.

with mBEEF, where relaxation leads to near-isoenergetics for the "prism" and "cage" isomers.

We further highlight the interesting finding above that mBEEF performs surprisingly well for non-covalently bonded systems by considering the S22 quantum chemical benchmark set[68,69] for non-covalently bonded complexes. This data set exhibits hydrogen bonding as well as van der Waals dispersion. Figure 5 shows error statistics for several GGA, MGGA, and vdW-DF type density functionals in reproducing the S22 binding energies. Semi-local DFAs do not contain the physics needed to reliably capture long-ranged dispersion interactions. It is therefore no surprise that the largest prediction errors in Fig. 5 are found for GGAs and MGGAs, while vdW functionals with explicitly non-local correlation are better suited for this. Though mBEEF is a semi-local functional and is not explicitly designed to capture dispersion interactions, its performance on S22 is good. Notably, mBEEF for example seems to on average outperform the significantly more expensive vdW-DF functional. We would expect even better performance on the S22 benchmark if a suitable non-local correlation term[46,70] was added to the mBEEF model space.

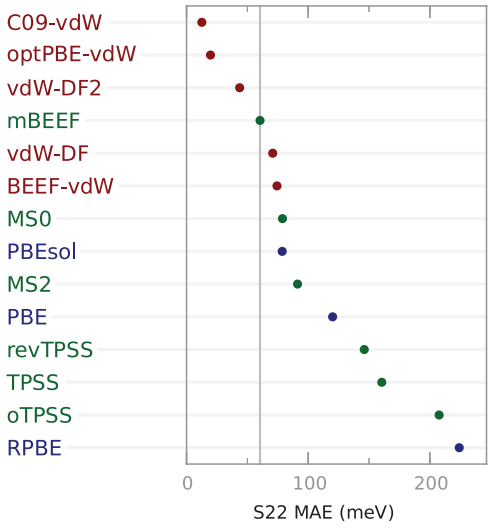

FIG. 5. Comparison of mean absolute density functional prediction errors for the S22 non-covalent benchmark set. Data are adapted from Ref. 25 except for mBEEF, MS0, and MS2.

D. Bayesian error estimates: The CO puzzle

Finally, let us consider an example of applying the BEE approach to error estimation in DFT. We choose a prototypical surface chemical problem: Predicting the site preference of molecular CO adsorption on close-packed surfaces of late transition metals. Most semi-local density functionals fail to correctly predict the most stable adsorption site over several such metals. This "CO puzzle" is a standing issue in computational surface chemistry, and a large number of studies have been devoted to elucidating its origin and its possible solutions, see, for example, Refs. 75–81.

Figure 6 shows calculated adsorption energy differences ΔE between the experimentally most stable CO adsorption site and less stable sites among the hollow and atop sites on close-packed facets of Rh, Pd, Pt, Cu, Co, and Ru, such that $\Delta E < 0$ eV corresponds to a correct theoretical prediction of the most stable of the two sites. Predictions made using a range of DFAs are indicated with different colors in the figure. The experimentally observed preference at low temperature and coverage is for the 1-fold coordinated atop site on all of the considered surfaces except Pd(111), on which the 3-fold coordinated fcc site is found to be energetically most favorable.

Bayesian error estimates σ_{BEE} are shown for mBEEF and BEEF-vdW calculations. Most GGAs and MGGAs correctly predict $\Delta E < 0$ eV on Rh(111), Pd(111), and Ru(0001), while on Pt(111), Cu(111), and Co(0001) the theoretical predictions are scattered around or just above $\Delta E = 0$ eV. The mBEEF σ_{BEE} values provide very reasonable estimates of the spread of predictions by different GGA or MGGA density functionals.

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

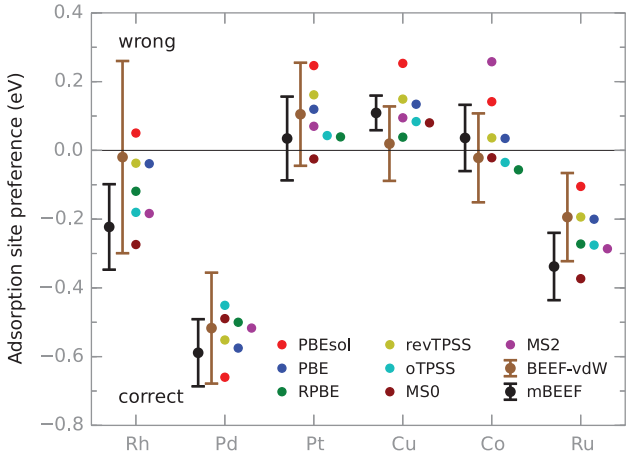

FIG. 6. Site preference ΔE for CO adsorption on (111) surfaces of Rh, Pd, Pt, and Cu and (0001) surfaces of Co and Ru at 0.25 monolayer coverage. Error bars on mBEEF and BEEF-vdW predictions indicate Bayesian error estimates.

In particular, the BEEs indicate that calculated adsorption site preferences for CO on Pt(111) and Co(0001) should not be considered indisputable, but may well change depending on the choice of exchange-correlation functional. In some sense Fig. 6 shows that such sensitivities of scientific conclusions are also found if we meticulously compute each ΔE using a wide range of different DFAs. However, Bayesian error estimation ensembles provide a quantitative and computationally inexpensive approach to such analysis.

VII. CONCLUSIONS

Broadly applicable semi-local density functionals must somehow be designed with the exchange-correlation model compromise in mind. The XC model selection procedure in the Bayesian error estimation functional framework effectively addresses this multi-objective optimization problem. We here used it to develop the mBEEF exchange-correlation functional, and argue that this can be considered a very reasonable general-purpose meta-GGA density functional. It delivers highly accurate predictions of a wide range of different properties in materials physics and chemistry, and we expect mBEEF to be particularly well suited for computational studies in surface science, including catalysis. A Bayesian ensemble for error estimation in DFT is an intrinsic feature of the BEEF-class of density functionals. The ensemble is defined in terms of XC model fluctuations, and we have illustrated the application of error estimation by considering a prototypical surface chemical problem. The mBEEF ensemble error estimates correctly indicate that one cannot conclusively determine the site-preference of CO adsorption on a number of late transition metal surfaces. A DFT user may traditionally try to get some idea about the sensitivity of calculated quantities on the choice of density functional approximation by tediously applying various established functionals to the same problem. The BEE provides a more structured approach to such analysis via systematic but computationally inexpensive computations of non-self-consistent XC energy perturbations. We expect this approach to quantitative error estimation of correlated errors to become a useful and very general tool for validating scientific conclusions based on DFT in computational materials science and chemistry. Finally, we find that the mBEEF functional captures the strength of hydrogen bonding and medium-range van der Waals bonding surprisingly well, even though it is a semi-local approximation and was not explicitly designed for this. This suggests that mBEEF may be a very appropriate starting point for a meta-GGA exchange-correlation functional explicitly including non-local van der Waals correlation to accurately account for long-range dispersion interactions.

ACKNOWLEDGMENTS

We acknowledge financial support to SUNCAT from the Chemical Sciences, Geosciences and Biosciences Division of the U.S. Department of Energy's Office of Basic Energy Science. We also wish to thank the Danish Center for Scientific Computing for computational resources.

[1]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[2]W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[3]K. Burke, J. Chem. Phys. **136**, 150901 (2012).
[4]G. H. Jóhannesson, T. Bligaard, A. V. Ruban, H. L. Skriver, K. W. Jacobsen, and J. K. Nørskov, Phys. Rev. Lett. **88**, 255506 (2002).
[5]T. Bligaard, G. H. Jóhannesson, A. V. Ruban, H. L. Skriver, K. W. Jacobsen, and J. K. Nørskov, Appl. Phys. Lett. **83**, 4527 (2003).
[6]D. Morgan, G. Ceder, and S. Curtarolo, Meas. Sci. Technol. **16**, 296 (2005).
[7]W. Setyawan, R. M. Gaume, S. Lam, R. S. Feigelson, and S. Curtarolo, ACS Comb. Sci. **13**, 382 (2011).
[8]A. Jain, G. Hautier, C. J. Moore, S. P. Ong, C. C. Fischer, T. Mueller, K. A. Persson, and G. Ceder, Comput. Mater. Sci. **50**, 2295 (2011).
[9]I. E. Castelli, T. Olsen, S. Datta, D. D. Landis, S. Dahl, K. S. Thygesen, and K. W. Jacobsen, Energy Environ. Sci. **5**, 5814 (2012).
[10]G. Hautier, A. Miglio, G. Ceder, G.-M. Rignanese, and X. Gonze, Nat. Commun. **4**, 2292 (2013).
[11]P. Strasser, Q. Fan, M. Devenney, W. H. Weinberg, P. Liu, and J. K. Nørskov, J. Phys. Chem. B **107**, 11013 (2003).
[12]J. Greeley, T. F. Jaramillo, J. Bonde, I. Chorkendorff, and J. K. Nørskov, Nat. Mater. **5**, 909 (2006).
[13]M. P. Andersson, T. Bligaard, A. Kustov, K. E. Larsen, J. Greeley, T. Johannessen, C. H. Christensen, and J. K. Nørskov, J. Catal. **239**, 501 (2006).
[14]F. Studt, F. Abild-Pedersen, T. Bligaard, R. Z. Sørensen, C. H. Christensen, and J. K. Nørskov, Science **320**, 1320 (2008).
[15]J. K. Nørskov, T. Bligaard, J. Rossmeisl, and C. H. Christensen, Nat. Chem. **1**, 37 (2009).
[16]G. Hautier, A. Jain, H. Chen, C. Moore, S. P. Ong, and G. Ceder, J. Mater. Chem. **21**, 17147 (2011).
[17]W. Kohn, A. D. Becke, and R. G. Parr, J. Phys. Chem. **100**, 12974 (1996).
[18]J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, J. Chem. Phys. **123**, 062201 (2005).
[19]Y. Zhao and D. G. Truhlar, J. Chem. Phys. **128**, 184109 (2008).
[20]S. Kurth, J. P. Perdew, and P. Blaha, Int. J. Quantum Chem. **75**, 889 (1999).
[21]R. Peverati and D. G. Truhlar, J. Chem. Phys. **136**, 134704 (2012).
[22]A. Ruzsinszky, G. I. Csonka, and G. E. Scuseria, J. Chem. Theory Comput. **5**, 763 (2009).
[23]J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, Phys. Rev. Lett. **108**, 253002 (2012).
[24]J. C. Snyder, M. Rupp, K. Hansen, L. Blooston, K.-R. Müller, and K. Burke, J. Chem. Phys. **139**, 224104 (2013).
[25]J. Wellendorff, K. T. Lundgaard, A. Møgelhøj, V. Petzold, D. D. Landis, J. K. Nørskov, T. Bligaard, and K. W. Jacobsen, Phys. Rev. B **85**, 235149 (2012).
[26]J. J. Mortensen, K. Kaasbjerg, S. L. Frederiksen, J. K. Nørskov, J. P. Sethna, and K. W. Jacobsen, Phys. Rev. Lett. **95**, 216401 (2005).
[27]V. Petzold, T. Bligaard, and K. W. Jacobsen, Top. Catal. **55**, 402 (2012).
[28]R. Y. Brogaard, P. G. Moses, and J. K. Nørskov, Catal. Lett. **142**, 1057 (2012).

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20

[29]M. Dell'Angela, T. Anniyev, M. Beye, R. Coffee, A. Föhlisch, J. Gladh, T. Katayama, S. Kaya, O. Krupin, J. LaRue, A. Møgelhøj, D. Nordlund, J. K. Nørskov, H. Öberg, H. Ogasawara, H. Öström, L. G. M. Pettersson, W. F. Schlotter, J. A. Sellberg, F. Sorgenfrei, J. J. Turner, M. Wolf, W. Wurth, and A. Nilsson, Science **339**, 1302 (2013).
[30]F. Studt, F. Abild-Pedersen, J. B. Varley, and J. K. Nørskov, Catal. Lett. **143**, 71 (2013).
[31]R. Y. Brogaard, B. M. Weckhuysen, and J. K. Nørskov, J. Catal. **300**, 235 (2013).
[32]B. Hammer, L. B. Hansen, and J. K. Nørskov, Phys. Rev. B **59**, 7413 (1999).
[33]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[34]J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Phys. Rev. Lett. **100**, 136406 (2008).
[35]P. Haas, F. Tran, P. Blaha, and K. Schwarz, Phys. Rev. B **83**, 205117 (2011).
[36]L. Schimka, J. Harl, and G. Kresse, J. Chem. Phys. **134**, 024116 (2011).
[37]R. Peverati and D. G. Truhlar, Phys. Chem. Chem. Phys. **14**, 16187 (2012).
[38]A. D. Becke, J. Chem. Phys. **109**, 2092 (1998).
[39]J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, Phys. Rev. Lett. **82**, 2544 (1999).
[40]A. D. Boese and N. C. Handy, J. Chem. Phys. **116**, 9559 (2002).
[41]J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).
[42]J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin, and J. Sun, Phys. Rev. Lett. **103**, 026403 (2009).
[43]R. Peverati and D. G. Truhlar, J. Phys. Chem. Lett. **3**, 117 (2012).
[44]J. Sun, B. Xiao, and A. Ruzsinszky, J. Chem. Phys. **137**, 051101 (2012).
[45]A. D. Boese and N. C. Handy, J. Chem. Phys. **114**, 5497 (2001).
[46]M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004).
[47]L. Goerigk and S. Grimme, J. Chem. Theory Comput. **6**, 107 (2010).
[48]J. Sun, R. Haunschild, B. Xiao, I. W. Bulik, G. E. Scuseria, and J. P. Perdew, J. Chem. Phys. **138**, 044113 (2013).
[49]J. Klimes, D. R. Bowler, and A. Michaelides, J. Phys.: Condens. Matter **22**, 022201 (2010).
[50]V. R. Cooper, Phys. Rev. B **81**, 161104(R) (2010).
[51]Linear relationships between DFT-predicted chemisorption energies and surface energies have appeared in recent literature for the particular cases of CO on Pt(111) and Rh(111), see Refs. 78 and 80. However, bivariate analyses of DFT prediction errors for surface chemistry and stability has, to the authors' knowledge, not previously been considered on such firm statistical footing as in Fig. 1.
[52]B. G. Janesko, V. Barone, and E. N. Brothers, J. Chem. Theory Comput. **9**, 4853 (2013).
[53]L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **112**, 7374 (2000).
[54]The G3/99 standardization factor is $1/(N_a - 1)$, where N_a is the number of atoms in each molecule. For RE42 the factor is $1/(N_r - N_p + 1)$, where N_r and N_p are the number of reactants and products in each reaction, respectively.
[55]The CE27a experimental chemisorption energies are essentially from Ref. 25, but are here referenced to free atoms rather than gas-phase adsorbates. The Sol54Ec and Sol58LC datasets both derive from related sets applied in Ref. 25, containing experimental cohesive energies and lattice constants of monoatomic fcc, bcc, and diamond structured crystals corrected for thermal and vibrational contributions. We here also include diatomic rocksalt, cesiumchloride, and zincblende crystals, and update the lattice constant phonon corrections with recent *ab initio* values from Ref. 82. Note that Sol54Ec and Sol58LC contain data for Pb(fcc) which is excluded in model training. The cohesive energy of lead appears a very significant outlier in model training, and is known to be severely overestimated in DFT, see the supplementary material for Ref. 25.
[56]J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, Phys. Rev. B **71**, 035109 (2005).
[57]J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Duak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen, H. H. Kristoffersen, M. Kuisma, A. H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P. G. Moses, J. Ojanen, T. Olsen, V. Petzold, N. A. Romero, J. Stausholm-Møller, M. Strange, G. A. Tritsaris, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G. K. H. Madsen, R. M. Nieminen, J. K. Nørskov, M. Puska, T. T. Rantala, J. Schiøtz, K. S. Thygesen, and K. W. Jacobsen, J. Phys.: Condens. Matter **22**, 253202 (2010).
[58]P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[59]S. R. Bahn and K. W. Jacobsen, Comput. Sci. Eng. **4**, 56 (2002).
[60]H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[61]A. B. Alchagirov, J. P. Perdew, J. C. Boettger, R. C. Albers, and C. Fiolhais, Phys. Rev. B **63**, 224115 (2001).
[62]C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st ed. (Springer, 2006).
[63]T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. (Springer, 2009).
[64]S. L. Frederiksen, K. W. Jacobsen, K. S. Brown, and J. P. Sethna, Phys. Rev. Lett. **93**, 165501 (2004).
[65]See supplementary material at http://dx.doi.org/10.1063/1.4870397 for the 64 mBEEF exchange expansion coefficients and the 64 × 64 mBEEF error estimation ensemble matrix.
[66]J. Sun, B. Xiao, Y. Fang, R. Haunschild, P. Hao, A. Ruzsinszky, G. I. Csonka, G. E. Scuseria, and J. P. Perdew, Phys. Rev. Lett. **111**, 106401 (2013).
[67]E. E. Dahlke, R. M. Olson, H. R. Leverentz, and D. G. Truhlar, J. Phys. Chem. A **112**, 3976 (2008).
[68]P. Jurečka, J. Šponer, J. Černy, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[69]T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, and C. D. Sherrill, J. Chem. Phys. **132**, 144104 (2010).
[70]O. A. Vydrov and T. Van Voorhis, J. Chem. Phys. **133**, 244103 (2010).
[71]M. Korth and S. Grimme, J. Chem. Theory Comput. **5**, 993 (2009).
[72]G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán, Phys. Rev. B **79**, 155107 (2009).
[73]L. Vitos, A. V. Ruban, H. L. Skriver, and J. Kollár, Surf. Sci. **411**, 186 (1998).
[74]V. S. Bryantsev, M. S. Diallo, A. C. T. van Duin, and W. A. Goddard III, J. Chem. Theory Comput. **5**, 1016 (2009).
[75]P. J. Feibelman, B. Hammer, J. K. Nørskov, F. Wagner, M. Scheffler, R. Stumpf, R. Watwe, and R. Dumesic, J. Phys. Chem. B **105**, 4018 (2001).
[76]M. Gajdos, A. Eichler, and J. Hafner, J. Phys.: Condens. Matter **16**, 1141 (2004).
[77]F. Abild-Pedersen and M. P. Andersson, Surf. Sci. **601**, 1747 (2007).
[78]A. Stroppa and G. Kresse, New J. Phys. **10**, 063020 (2008).
[79]X. Ren, P. Rinke, and M. Scheffler, Phys. Rev. B **80**, 045402 (2009).
[80]L. Schimka, J. Harl, A. Stroppa, A. Grüneis, M. Marsman, F. Mittendorfer, and G. Kresse, Nat. Mater. **9**, 741 (2010).
[81]P. Lazić, M. Alaei, N. Atodiresei, V. Caciuc, R. Brako, and S. Blügel, Phys. Rev. B **81**, 045401 (2010).
[82]P. Hao, Y. Fang, J. Sun, G. I. Csonka, P. H. T. Philipsen, and J. P. Perdew, Phys. Rev. B **85**, 014111 (2012).

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 192.38.67.112 On: Wed, 23 Jul 2014 08:42:20