Data-driven and constrained optimization of semi-local exchange and nonlocal correlation functionals for materials and surface chemistry

Kai Trepte | Johannes Voss

SUNCAT Center for Interface Science and Catalysis, SLAC National Accelerator Laboratory, Menlo Park, California, USA

Correspondence
Kai Trepte and Johannes Voss, SUNCAT Center for Interface Science and Catalysis, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA.
Email: ktrepte@slac.stanford.edu; vossj@slac.stanford.edu

Funding information
U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, Catalysis Science Program to the SUNCAT Center for Interface Science and Catalysis

Abstract
Reliable predictions of surface chemical reaction energetics require an accurate description of both chemisorption and physisorption. Here, we present an empirical approach to simultaneously optimize semi-local exchange and nonlocal correlation of a density functional approximation to improve these energetics. A combination of reference data for solid bulk, surface, and gas-phase chemistry and physical exchange-correlation model constraints leads to the VCML-rVV10 exchange-correlation functional. Owing to the variety of training data, the applicability of VCML-rVV10 extends beyond surface chemistry simulations. It provides optimized gas phase reaction energetics and an accurate description of bulk lattice constants and elastic properties.

1 | INTRODUCTION

Kohn-Sham (KS) density functional theory (DFT) is one of the most widely used electronic structure theories. It provides reasonable accuracy for many physical properties of atoms, molecules, surfaces, and solids. For the practical use of KS DFT, approximations to electronic exchange and correlation (XC) are needed. These so-called XC functionals have been developed for decades. There is a variety of both empirical and nonempirical functionals. In the latter approaches it was established that fulfilling analytical constraints leads to functionals with improved predictions of molecular and bulk properties. Typically, empirical as well as nonempirical functionals describe some physical or chemical properties with improved accuracy at the price of worse predictions of other properties.

To obtain an XC functional that accurately describes a wide range of physical properties, strategies from the empirical and nonempirical XC functional development approaches can be combined. Constraints from the latter can be imposed while fitting a multi-parameter empirical XC functional form to experimental and quantum chemistry reference data.

There are several approaches combining constraints and data for XC functionals. Recently, our group has applied such a combined approach to construct a multi-purpose, constrained and machine learned (MCML) XC functional. MCML shows improved predictions of surface and gas phase reactions without sacrificing the good description of bulk properties. The MCML functional is a so-called semi-local functional, lacking an explicit description of nonlocal, that is, van der Waals (vdW) type correlation. Supplementing a semi-local functional with vdW correlation does not only affect the performance for vdW-dominated interactions. For example, some semi-local approximations tend to overbind chemisorbed systems. Adding attractive dispersion forces can increase this tendency even further. The semi-local part of a functional can be optimized to compensate for effects of overbinding or lattice spacing contraction from the attractive vdW terms. On the other hand, the vdW part can be optimized using, e.g., the parameterized, semi-empirical rVV10 nonlocal vdW term. With that, performance on dispersion-dominated benchmark data is optimized while fixing the underlying semi-local functional.

In the present work, we optimize the semi-local exchange and nonlocal correlation parts of a vdW XC functional simultaneously for several bulk, surface, and gas phase property predictions. This simultaneous optimization of the semi-local exchange functional form and an rVV10 nonlocal term enables us to construct a functional that is accurate for a range of materials properties governed by different types of
chemical bonds. There is no compromise between the accuracy for the description of ionic, covalent, or metallic bonds and dispersive interactions, or vice versa. Incorporating physical constraints into this empirical approach leads to the VCML-rVV10 functional (vdW functional using constraints and machine learning, employing the rVV10 formalism), see Figure 1. This name shall clearly separate our functional from MCML. The semi-local exchange functional parts of MCML and VCML-rVV10 are different, and the latter is not simply a vdW-supplemented form of MCML. While we optimize the semi-local exchange as well as the nonlocal correlation part of our functional, for semi-local correlation we employ REGTPSS, in analogy to Brown et al.

This article is structured as follows. In Section 2, a brief introduction of the theoretical background including the rVV10 formalism is given. The used computational parameters are shown in Section 3. All employed data sets are described in Section 4. Thereafter, our functional optimization approach is presented in Section 5. The main results are discussed in Section 6. A conclusion is presented in Section 7.

2 | THEORETICAL BACKGROUND

Atomic units are being used throughout this section. In the KS formulation of DFT, the total energy of a system with total electron density

\[ n(r) = n^1(r) + n^\uparrow(r), \]  

where \( n^\uparrow \) and \( n^1 \) are the spin densities of a system with collinear spins, can be expressed as

\[ E^{KS} = T_S[n^\uparrow, n^\uparrow] + E_{\text{ext}}[n] + E_{\text{H}}[n] + E_{\text{xc}}[n^\uparrow, n^\uparrow]. \]  

Here, \( T_S \) is the noninteracting kinetic energy, \( E_{\text{ext}} \) is the energy of the interaction with an external potential, and \( E_{\text{H}} \) is the Hartree energy of the electrons. Exchange and correlation effects are treated via the exchange-correlation functional \( E_{\text{xc}} \).

The electron density is the fundamental building block of DFT. It determines the energy as well as any other property. While DFT is in principle exact, the \( E_{\text{xc}} \) needs to be approximated. Such approximations can be categorized according to their ingredients in the functional form. This is known as the Jacob’s ladder of DFT. The lowest rung is given by the local (spin) density approximation (LSDA). Only the density itself is taken as a parameter. Additionally including the density gradient characterizes generalized gradient approximations (GGAs). Meta-GGAs also include a dependence on the kinetic energy density of the occupied orbitals.

The exchange energy for meta-GGAs is defined as an integral of the exchange energy density per electron of the homogeneous electron gas (HEG) at density \( n(r) \)

\[ E_X[n(r)] = \frac{3}{4}(3/\pi)^{1/3}n(r)^{5/3}. \]  

multiplied by an exchange enhancement factor \( F_X \)

\[ E_X[n(r)] = \int d^3r n(r)E_X[n(r)]F_X(s(r), a(r)). \]  

Here,

\[ s(r) = \left| \nabla n(r) \right| / \left( 2k_F(r)n(r) \right) \]  

is the reduced density gradient with the Fermi wave vector

\[ k_F(r) = \sqrt{3\pi^2n(r)}. \]  

The dependence on the kinetic energy density can be described as

\[ a(r) = \frac{\varepsilon_{\text{KS}}(r) - \varepsilon_{\text{Weizsäcker}}(r)}{\varepsilon_{\text{HEG}}(r)}, \]  

with

\[ \varepsilon_{\text{KS}}(r) = \frac{1}{2} \sum f_i |\nabla \phi_i|^2, \]  

\[ \varepsilon_{\text{Weizsäcker}}(r) = |\nabla n(r)|^2 / (8n(r)), \]  

\[ \varepsilon_{\text{HEG}}(r) = \frac{3}{10}(3\pi^2)^{2/3}n(r)^{5/3} \]  

being the kinetic energy density based on the KS orbitals \( \phi_i(r) \) with occupation \( f_i \), the von Weizsäcker kinetic energy density describing the single orbital limit, and the noninteracting kinetic energy density of a HEG at density \( n(r) \), respectively. The term \( \varepsilon_{\text{Weizsäcker}}(r) \) vanishes for zero charge density gradient and equals \( \varepsilon_{\text{KS}}(r) \) for a single orbital.

\[ F_X(s, a) = \frac{1}{2} \left( 1 + \frac{s}{a} \right)^{1/2}. \]  

This expression is fit to the accurate \( F_X \) of Knowles et al. to allow the easy implementation of a \( F_X \) functional.
density. Accordingly, Equation (7) approaches zero in the single orbital limit and one in the homogeneous limit where $\nabla n(r) = 0$ and $\varepsilon_{KS}(r) = \varepsilon_{HEG}(r)$. With that, the parameter $\alpha$ can characterize the type of bonding in a system.

For spin-polarized systems, $E_X$ is obtained via the spin scaling relation

$$E_X[n^+(r),n^-(r)] = \frac{E_X[n^2(r)] + E_X[n^1(r)]}{2}. \quad (11)$$

The exchange enhancement factor $F_X$ is optimized in this work by training against reference data, see Section 5 for more details.

In the rVV10 methodology, a nonlocal correlation energy is introduced as

$$E_{C_{\text{nl}}} = \frac{1}{2} \int \left[ q^2 d^3 r \Theta(n(r),n(r')) n(r') \right]. \quad (12)$$

The nonlocal kernel as described by Sabatini et al. is given by

$$\Theta(n(r),n(r'))^\text{rVV10} = \frac{3}{2} \frac{1}{(q(r)R^2 + 1)(q'(r')R^2 + 1)(q(r)R^2 + q'(r')R^2 + 2)}. \quad (13)$$

in which

$$q(r) = \frac{\omega_0(n(r),|\nabla n(r)|)}{k(r)} \quad (14)$$

and $R = |r - r'|$. The same expression is obtained for $q'$ by replacing $r$ with $r'$. The term $\omega_0$ depends on the local band gap as well as the plasma frequency. For more information, see Sabatini et al. as well as the Supporting Information, Section 1. Importantly for further discussion, the term

$$k(r) = 3ab \left( \frac{n(r)}{9\pi} \right)^{\frac{3}{2}} \quad (15)$$

determines the short-range damping of the $R^{-6}$ divergence of the nonlocal kernel. The parameter $b$ needs to be optimized for any functional $E_{XC}$ that $E_{C_{\text{nl}}}$ is added to. Simply speaking, the larger the $b$ parameter, the weaker the vdw interaction. The optimization of this parameter is discussed in Section 5. For additional details we refer to the original works of Vydrov and Van Voorhis and Sabatini et al.

3 | COMPUTATIONAL DETAILS

In this work, the Vienna Ab Initio Simulation Package (VASP) was employed. The DFT calculations were carried out using projector-augmented wave [1] pseudopotentials which are based on PBE all-electron atomic calculations, and a plane wave basis set. An energy cutoff of 1000 eV (73.50 Ry or 36.75 $E_h$) was used together with an electronic SCF tolerance of $10^{-7}$ eV. For geometry optimizations, forces were converged such that the maximum force per atom is at most $10^{-2}$ eV/Å. Note that in the ADS41 set (see Supporting Information, Tab. ST25), $[\text{H}_2\text{O,CH}_3\text{OH}]@\text{Pt111}$ using SCAN required a looser criterion of $2 \cdot 10^{-2}$ eV/Å because of numerical instabilities, which are absent in $r^2$SCAN. The k-point spacing for the surface (in two-dimensions) and bulk (in three-dimensions) calculations was at most 0.018 Å$^{-1}$. A Gaussian smearing of KS occupation numbers with a width of $2 \cdot 10^{-2}$ eV was used for surfaces and bulk systems.

For data sets involving molecules, the unit cell has been prepared as follows. Around the largest system in the set, a box was constructed such that there is at least 20 Å in between periodic images in all directions. This box was employed for all molecules in the corresponding data set. For atoms, a cell of 23 Å was used.

To compare the performance of the VCML-rVV10 functional, calculations were also performed with PBE, PBE-D3, MS2, SCAN, r^2SCAN, SCAN-rVV10, MCMC, and MCML-rVV10 (introduced in this work). For the error analysis, the mean error (ME) and mean absolute error (MAE) are computed

$$\text{ME} = \frac{1}{N} \sum_{i=1}^{N} x_{i}^{\text{calc}} - x_{i}^{\text{ref}} \quad (16)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |x_{i}^{\text{calc}} - x_{i}^{\text{ref}}| \quad (17)$$

where $x_{i}^{\text{calc}}$ is the calculated value and $x_{i}^{\text{ref}}$ the corresponding reference value. These values are provided in the Supporting Information, Tab. ST7, 9, 11–20, 22–25, and 27. All individual errors for any system of a data set are computed as $x_{i}^{\text{calc}} - x_{i}^{\text{ref}}$.

4 | BENCHMARK DATA SETS

To obtain a multi-purpose XC functional from the proposed methodology, see Section 5, a variety of training data is required. Our training data consists of molecular, surface as well as bulk properties, as described below. The systems of each data set and all calculated properties are provided in the Supporting Information, Tab. ST6–27.

a. The DBH24 set consists of 12 forward and 12 reverse reaction barrier heights of small molecules. Reference values are the best estimates, consisting of quantum chemical and experimental data, according to Zheng et al. Total energies were computed on the reference structures; no structural optimizations were carried out.

b. The RE42 set contains 42 reaction energies involving 45 molecules of the G2/97 test set. Experimental reference values are taken from Wellendorf et al. The geometries of all molecules were fully optimized before their total energy was used to calculate the reaction energies.
c. The $S66x8$ set\(^{20}\) has 66 molecular complexes at eight distances, the interaction of which is dominated by noncovalent bonding. Reference values are the recommended dissociation energies, calculated by MP2-F12 energies at the basis set limit combined with CCSD(T)$_{E}$- F12 correlation, according to Brauer et al.\(^{39}\) Total energies were computed at the reference geometries, thus no geometry optimizations were carried out.

d. The W4-11 set\(^{40}\) contains 140 atomization energies of small to medium sized molecules. Reference values are taken from Karton et al.\(^{41}\) which are based on the Weizmann-4 (W4) computational thermochemistry method. No structural optimizations were carried out, and total energies were computed on the reference structures.

e. The SOL62 sets based on a set of 64 solids proposed by Zhang et al.\(^{17}\) Of that set we removed all elements heavier than Au, as for those systems (Pb, Th) the predicted errors will typically be very large due to strong relativistic effects.\(^{23}\) The reference values are taken from the supplemental information of Zhang et al.,\(^{17}\) where zero-point energy corrections of PBE were subtracted from the experimental values. All unit cells were fully optimized. Afterwards, the lattice constants $a_{lat}$ and bulk moduli $B$ were computed using an equation of state (EOS) fit\(^{41}\) from five energies around and at the minimum volume. Further, cohesive energies $E_{coh}$ were computed with respect to the isolated atoms in their energetically preferred magnetic ground state. This data set is split into the three subset $a_{lat}@SOL62$, $E_{coh}@SOL62$, and $B@SOL62$.

f. The ADS41 set\(^{22,42}\) consists of 41 surface reaction energies. Experimental reference values are taken from the collection in Sharada et al.\(^{22}\) The atomic positions of the isolated molecules, the isolated surfaces as well as the combined systems were optimized. For the clean surfaces and the combined systems, the cell was constrained to the bulk lattice dimensions, while introducing at least 15 Å of vacuum in the out-of-plane direction. The vacuum layer is the same for the isolated surface and the combined system, ensuring consistent results. The surfaces are constructed to have four layers. The lower two are fixed at the bulk lattice positions, while the top two layers (and the adsorbed molecule) are optimized. Adsorption energies are computed per adsorbate, thus the total adsorption energy is divided by the number of adsorbates in the computational cell. Because ADS41 contains Co and Ru surfaces, we additionally optimized the unit cells of either hcp solid to obtain the corresponding lattice constants. These solids are not part of the SOL62 set, and were only optimized for ADS41. The ADS41 data set is subdivided into physisorption-dominated reactions (15 reaction energies, denoted as $E_{pH}@ADS41$) and chemisorption-dominated reactions (26 reaction energies, denoted as $E_{ch}@ADS41$).

5 | FUNCTIONAL FITTING APPROACH

5.1 | General outline

In the following, the meta-GGA as well as the vdW fitting procedure is presented. The procedure to obtain the meta-GGA functional form has been outlined in Brown et al.\(^{18}\) for the MCML functional. We will briefly discuss its main components here. Our exchange enhancement factor is expanded in a series of Legendre polynomials $P$ for $s$ and $a$ as

$$F_X(s(r),a(r)) = \sum_{l=0}^{7} \sum_{j=0}^{7} c_lP_l(s(r))P_j(a(r)).$$  \(\text{(18)}\)

where

$$\tilde{s}(r) = \frac{(s(r))^2}{\eta + (s(r))^2} - 1$$  \(\text{(19)}\)

maps the semi-infinite interval of reduced density gradients $s$ to the interval spanned by the Legendre polynomials, $[-1,1]$. Here, $\eta = \kappa/\mu_{GE}$ is used to represent the exchange enhancement of PBEsol\(^{11}\) by the first two Legendre polynomials $P_0(s(r))$ and $P_1(s(r))$. Further, $1 + \kappa = 1.804$ is the local Lieb-Oxford bound,\(^{43,44}\) and $\mu_{GE} = 10/81$ is the lowest-order coefficient of the gradient correction to the free electron gas exchange energy.\(^{45}\) Similarly,

$$\tilde{a}(r) = \frac{(1 - a(r))^2}{1 + a(r)^2 + 4a(r)^4}$$  \(\text{(20)}\)

maps the semi-infinite interval of $a$ to the interval $[-0.25,1]$. This form of $\tilde{a}$ coincides with the GGA-weighting function of the MS2 functional.\(^{34}\) The transformations of $s$ and $a$ are visualized in the SI, Fig. SF1. Optimizing the 64 coefficients $c_l$ in Equation \((18)\), as described in the next sections, delivers an optimal semi-local meta-GGA functional. The physical constraints introduced for MCML are also applied to our functional, that is, the LDA limit with $F_X = 1$ for $s = 0$ and $a = 1$, the exchange gradient expansion for $s \approx 0$ at $a = 1$, and the cancelation of the spurious Hartree energy in the H atom for $a = 0$. The LDA limit is enforced via

$$\sum_{l=0}^{7} \sum_{k=0}^{3} (-1)^{l+k} \frac{(2k)!}{2^{2k}k!} c_{2l2k} = 1,$$  \(\text{(21)}\)

while the exchange gradient expansion is described with a curvature of $2\tilde{a}_{GE}$ for $s \approx 0$ at $a = 1$

$$\sum_{l=0}^{7} \sum_{k=0}^{3} (-1)^{l+k} \frac{(2k)!}{2^{2k}k!} \frac{2(i + 1)}{\eta} c_{2l2k} = 2\tilde{a}_{GE}.$$  \(\text{(22)}\)

The spurious Hartree energy of the hydrogen atom of $\frac{1}{2}E_h$ is canceled with

$$\left(\frac{162}{\pi^2}\right)^{\frac{5}{2}} \sum_{l=0}^{7} \sum_{j=0}^\infty \int dr^2 y^2 P_l(s^H) c_j = \frac{5}{16}.$$  \(\text{(23)}\)

where $y = \exp(-4r/3)$, $s^H = 2p/(y + p) - 1$, and $p = s^2 = (6x)^{-2/3}/y$ is the square of the reduced density gradient of the 1s hydrogen atom ground state.
Besides PBE, the functionals considered in this study fulfill all three constraints Equations (21), (22), and (23).

5.2 Regularized optimization of $F_X$

We extend the approach introduced in Brown et al.\textsuperscript{18} by including the vdW description via rVV10. The vdW description is optimized based on the functional form, after which our exchange functional is optimized based on a given vdW parameterization. MCML is used as a starting point for the optimization. We took the product of the non-self-consistent (nonSCF) MAEs for all data sets as the cost function for the optimization. These nonSCF predictions are explained in Brown et al.,\textsuperscript{18} and are briefly outlined in Supporting Information, section 3. The cost function

$$\theta = \prod_i (\text{MAE}_i)^{w_i}$$

(24)

was minimized using the Nelder–Mead simplex algorithm,\textsuperscript{46–48} raising each MAE to a weight $w_i$. Adjusting the weights results in different fits, see Section 5.3. This enables tuning trade-offs in performance for different data sets against each other.

5.3 Choosing weights

The weights were chosen according to Figure 2. As a starting point, all weights were set to 1. The functional form is optimized, and the non-SCF predictions of the MAE for each data set are computed. If an MAE is beyond a predefined threshold, the weight for the data set is increased (see Supporting Information, Tab. ST1 for values). After this modification, the functional is optimized once again and the process starts anew. This optimization route was performed 100 times. The best possible weights, that is, the best compromise between all predicted MAEs, were adjusted to minimize the errors further. The final weights (see Supporting Information, Tab. ST2) were then taken to obtain the optimal $c_{ij}$ (see Supporting Information, Tab. ST3).

5.4 Enforce smoothness

To avoid oscillatory behavior of high-order polynomial fits, the $F_X$ is enforced to be smooth. The number of sign changes beyond given thresholds in the first and second derivatives of $F_X$ with respect to $s$ and $\alpha$ is penalized. A penalty is added if there is more than one sign change in the first derivatives. This allows the exchange enhancement factor to have one extremum in $s$ for a given $\alpha$. Further, a penalty is applied if there are more than two sign changes in the second derivatives. With that, the functional should go smoothly toward and away from the extremum. Derivatives of the Legendre polynomials used to expand $F_X$ are efficiently computed via recursion. The zeros of the resulting polynomials corresponding to the first and second derivatives are obtained as the eigenvalues of the companion matrices. This sign-change penalty technique enables us to enforce smoothness of $F_X$ in a computationally straightforward way. As an example, the first and second derivative of $F_X$ with respect to $s$ at $\alpha = 0$ (single orbital limit) for VCML, the semi-local exchange part of VCML-rVV10, are shown in Figure 3. As one can see, the proposed regularizations work as intended in making the functional smooth. Note that there are different mathematical forms of $F_X$ not based on polynomial expansions allowing for other techniques to enforce constraints and smoothness.\textsuperscript{19}
5.5 | Summary of entire procedure to obtain VCML

Figure 4 summarizes the fitting procedure. It starts by taking the MCML functional. An optimal $b$ parameter for $rV_{10}$ is determined. The resulting MCML-$rV_{10}$ functional is used to compute all data sets self-consistently. Based on these results, the functional form was adjusted according to the minimization of Equation (24). The adjustment of the weights is illustrated in Figure 2. This results in an intermediate functional form, called Fit1-$rV_{10}$. For this functional, the $b$ parameter was re-optimized, and all data sets were re-calculated self-consistently. Finally, the functional form was re-optimized once again, resulting in VCML-$rV_{10}$. Due to careful consideration of the errors in S66x8 as described in the Data S1, section 3, the $b$ parameter needed no further optimization.

6 | RESULTS AND DISCUSSION

6.1 | Comparing $F_X$ to other functionals

The $F_X$ of MS2, SCAN, MCML, and VCML are plotted in Figure 5. In certain regions of $s$ and $\alpha$, VCML is similar to MS2, for
For chemisorption-dominated systems, VCML-rVV10 performs significantly better than all other vdW-supplemented functionals considered here; the resulting MAE is a good as MS2. As such, VCML-rVV10 is the second best functional (together with MS2) for these systems. Only MCML gives smaller errors, at the price of not explicitly accounting for dispersion forces.

For all data sets it is clear that only adding rVV10 on-top of MCML, that is, MCML-rVV10, does not provide a good vdW-meta-GGA functional. Reshaping the functional form based on a given vdW correction drastically improves the performance of the functional, making it applicable to a wide range of physical properties while also treating nonlocal interactions via the rVV10 methodology.

6.3 | Graphene on Ni

To test VCML-rVV10 outside of the training data shown in Figure 6, we calculated the interaction energy of graphene on a Ni(111) surface. Previous theoretical investigations found that there likely are two distinct minima. These minima can be characterized as a chemisorption minimum at a smaller distance between graphene and Ni, and a physisorption minimum at a larger distance. In experiments, usually only the
In the literature, reference calculations have been carried out using the random phase approximation (RPA). It has been noted that the RPA is likely to underestimate the chemisorption minimum. This can be seen when comparing to experimental estimates, which are shown as the gray cross in Figure 7. The second minimum, on the other hand, is likely captured well by the RPA.

In Figure 7, we compare the interaction energy per C atom calculated with all functionals considered in this article. As a note, the optimal lattice constant for Ni for each functional was employed. Of all the functionals, only MCML-rVV10 and VCML-rVV10 describe the energy of the first minimum accurately, with VCML-rVV10 being within 6 meV of the average estimate of the experimental chemisorption energy. All other functionals underestimate this minimum or predict it to be as strongly bound as the physisorbed case. Further, VCML-rVV10 is energetically closest to RPA for the second minimum. As a note, VCML-rVV10 and MCML-rVV10 are the only functionals that are very close, within 5 meV, to the RPA values at larger distances. As such, VCML-rVV10 describes both chemisorption and physisorption accurately, indicating a well-balanced functional.

The decomposition of $E_X$ into contributions from the different Legendre polynomial products in the expansion of the exchange-enhancement factor $F_X$ allows for efficient non-self-consistent estimates of changes in $E_X$ due to perturbations of $F_X$. Following the strategy outlined in References 18 and 53, we use such perturbations to estimate the uncertainty in the interaction energy of graphene with Ni(111) in the example above using Bayesian inference. Perturbations to $F_X$ are drawn randomly with a probability $\propto \exp(-\text{Loss}_2(\theta)/\tau)$. Loss$_2(\theta)$ is the squared error with respect to the reference fitting data as a function of the parameters $\theta$ defining $F_X$ under fulfillment of the constraints Equations (21), (22), and (23). For the computation of Loss$_2(\theta)$, the energies in each data set are normalized such that the MAE of each data set is one. The fictitious temperature $\tau$ is determined such that the residual fitting error of VCML-rVV10 is reproduced (see Reference 18 for further details on the implementation of the error estimates). The predicted uncertainty for the interaction of graphene with Ni(111) is shown as the shaded area in Figure 7. The Bayesian error estimate for the chemisorption energy is about 0.2 eV, which is consistent with the residual error on the chemisorption benchmark data, see Figure 6. The ensemble error covers all stable adsorption minima. The predicted error vanishes at larger distances, where VCML-rVV10 and RPA agree very well.

7 | CONCLUSION

In this work, a new meta-GGA van der Waals (vdW) exchange-correlation (XC) functional called VCML-rVV10 is introduced. This functional is obtained from a simultaneous optimization of both semilocal exchange via a multi-parameter model and nonlocal correlation of rVV10-type. VCML-rVV10 was trained on a number of data sets representing several chemical and physical properties. This enables the multi-purpose character of the functional. In addition, several constraints of the exact XC functional were enforced for physicality of the VCML-rVV10 model.

The newly introduced functional shows good performance in comparison to all other functionals tested, some of which also employing the rVV10 methodology. It also performs very well on the description of graphene on Ni(111), which was not included in the training data. This benchmark system requires a balanced, accurate description of surface physisorption as well as chemisorption.

In the future, the introduced methodology of simultaneous, regularized optimization of several parts of an XC functional could be extended to hybrid functionals, where a to-be-fitted amount of screened exact exchange would be admixed to the to-be-optimized functional.

ACKNOWLEDGEMENT

This research was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, Catalysis Science Program to the SUNCAT Center for Interface Science and Catalysis. We greatly appreciate the help of James Furness for providing a patch for VASP making $r^2$SCAN available.

DATA AVAILABILITY STATEMENT

All data, including the VASP inputs/outputs as well as all calculated values, is available at GitLab https://gitlab.com/kaitrepte/vcml_data. The AD$541$ adsorbate system structures can also be found on catalysis-hub, see https://www.catalysis-hub.org/publications/KaiData-driven2022. VCML-rVV10 is available in libXC, see https://github.com/ComputationalElectronicStructureLibrary/libxc. An alternative FORTRAN routine for evaluation of the VCML exchange energy can be found at https://github.com/vossgroup/CML.
