Letter | Editors' Suggestion

Neural density functionals: Local learning and pair-correlation matching

Florian Sammüller [*] and Matthias Schmidt [†]
Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, 95447 Bayreuth, Germany

(Received 5 June 2024; accepted 30 July 2024; published 12 September 2024)

Recently, Dijkman *et al.* [arXiv:2403.15007] proposed training classical neural density functionals via bulk pair-correlation matching. We show their method to be an efficient regularizer for neural functionals based on local learning of inhomogeneous one-body direct correlations [Sammüller *et al.*, Proc. Natl. Acad. Sci. USA **120**, e2312484120 (2023)]. While Dijkman *et al.* demonstrated pair-correlation matching of a global neural free-energy functional, we argue in favor of local one-body learning for flexible neural modeling of the full Mermin-Evans density-functional map. Using spatial localization gives access to accurate neural free-energy functionals, including convolutional neural networks, that transcend the training box.

DOI: 10.1103/PhysRevE.110.L032601

As machine learning and density-functional theory [1–4] share high computational efficiency, it is natural to use neural networks to construct workable and accurate approximations for the required density-functional relationships in the classical [5–16] and quantum realms [17–24]. This common goal to make progress is nevertheless approached from quite different directions and a considerable range of different machine learning strategies have been put forward [5–25]. The task is both important and challenging. Overcoming the limited availability of flexible classical density-functional approximations would open up the study of a much broader class of soft matter systems than is currently accessible via, e.g., the hard-sphere perturbation paradigm of fundamental-measure theory [26–28] combined with mean-field attraction, as used in recent work [29–31].

In the classical context, typically the excess free-energy functional $F_{\rm exc}[\rho]$ is the object chosen to be approximated [1,2]. We recall that $F_{\rm exc}[\rho]$ is the nontrivial contribution to the total grand potential functional $\Omega[\rho] = F_{\rm id}[\rho] + F_{\rm exc}[\rho] + \int d\mathbf{r}\, \rho(\mathbf{r})[V_{\rm ext}(\mathbf{r}) - \mu]$, where the ideal gas free-energy functional $F_{\rm id}[\rho]$ is known explicitly, $V_{\rm ext}(\mathbf{r})$ is the external potential that generates spatial inhomogeneity, and μ is the chemical potential that together with absolute temperature T determines the thermodynamic conditions. The Mermin-Evans variational principle [1,32] ascertains that $\Omega[\rho]$ is minimized by the true equilibrium density profile $\rho(\mathbf{r})$, as a function of position $\mathbf{r}$ across the system, and that the minimum gives the equilibrium value of the grand potential.

The effects of the interparticle interactions are contained in $F_{\rm exc}[\rho]$, which *per se* is not easily accessible and requires the development of careful modeling strategies. Neural networks suit this task very well, due to the clear-cut correspondence between the spatially resolved input density profile $\rho(\mathbf{r})$ and the excess free energy $F_{\rm exc}[\rho]$ as a global output value. Reference data for supervised machine learning are provided by many-body simulations [33–35], which hence offer the perspective of making theoretical predictions with simulation precision.

Two very recent implementations use neural networks to represent the functional relationship of either the global excess free energy $F_{\rm exc}[\rho]$ [10] or the closely related one-body direct correlation functional $c_1(\mathbf{r};[\rho])$ [12,13]. The training strategy in Ref. [10] employs bulk pair-correlation matching, while Refs. [12,13] utilize a local learning approach involving inhomogeneous one-body profiles. An appeal of the former [10] is the sole requirement of simulation input in the form of bulk radial distribution functions $g(r)$ at different densities for the particular model under investigation. Local learning [12,13] rather requires inhomogeneous density profiles as training data. Going beyond mere interpolation tasks, the resulting neural functionals [12,13] were however shown to be fit for carrying out deep functional calculus based on exact sum rules and on functional identities. Both methods were argued to be numerically highly accurate [10,12,13].

In this Letter we contrast and cross-fertilize the underlying concepts of pair-correlation matching [10] and of inhomogeneous one-body learning [12,13]. We show how to incorporate pair-correlation matching in a local learning strategy, which then retains the ability of transcending the box size of the training simulations. In keeping with statistical mechanics, pair-correlation matching constitutes a physics-based regularizer for the training loss, thereby avoiding potential problems of more generic regularization methods [36].

We also demonstrate, based on comparison to reference simulation data and on violation of internal sum rule consistency, that training the density-functional dependence solely with bulk densities and pair-correlation matching can yield unreliable neural performance in general inhomogeneous situations. We rationalize this behavior in terms of the structure of the underlying exact functional dependences. We show that local learning does not suffer from such deficiencies and that the concept also generalizes beyond constructing neural one-body direct correlations, laying out two explicit strategies that make $F_{\rm exc}[\rho]$ directly accessible.

[*] Contact author: Florian.Sammueller@uni-bayreuth.de
[†] Contact author: Matthias.Schmidt@uni-bayreuth.de

We first describe key ideas of pair-correlation matching as proposed by Dijkman *et al.* [10] to construct a neural network that represents the functional $F_{\rm exc}[\rho]$. Their method exploits that $F_{\rm exc}[\rho]$ is a generating functional for the hierarchy of direct correlation functionals. Specifically, at second order the two-body direct correlation functional is obtained [1,2],

$$c_2(\mathbf{r},\mathbf{r}';[\rho]) = -\frac{\delta^2\beta F_{\rm exc}[\rho]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}, \tag{1}$$

where $\delta/\delta\rho(\mathbf{r})$ indicates the functional derivative with respect to $\rho(\mathbf{r})$ and $\beta = 1/(k_BT)$ with Boltzmann constant k_B. When specializing to the important, yet restricted, case of bulk fluids, the density profile is constant, $\rho(\mathbf{r}) = \rho_b = \text{const}$, and the bulk pair direct correlation function $c_2(r;\rho_b)$ of liquid state theory is recovered [2]. This constitutes a functional reduction

$$c_2(r;\rho_b) = c_2(\mathbf{r},\mathbf{r}';[\rho=\rho_b]), \tag{2}$$

where the spatial dependence on $\mathbf{r}$ and $\mathbf{r}'$ simplifies to the dependence on radial distance $r = |\mathbf{r}-\mathbf{r}'|$ on the left-hand side. Also, the general functional dependence on the entirety of the density profile $\rho(\mathbf{r})$ in Eq. (1) is reduced to a mere parametric dependence on the value of the bulk density ρ_b in Eq. (2).

Pair-correlation matching [10] exploits that results for pair-correlation (or radial distribution) functions $g(r)$ are readily accessible via simulations. Using the total correlation function $h(r) = g(r) - 1$, the bulk Ornstein-Zernike equation ascertains that $h(r) = c_2^{\rm ref}(r) + \rho_b c_2^{\rm ref}(r) \star h(r)$, where the star indicates spatial convolution. We have denoted the bulk two-body direct correlation function from this route by $c_2^{\rm ref}(r)$, as originating from the simulation data for $g(r)$. Numerical solution of the Ornstein-Zernike equation yields quasiexact reference results for $c_2^{\rm ref}(r)$; details are given in our Supplemental Material [37]. Pair-correlation matching [10] trains $F_{\rm exc}[\rho]$ as a convolutional neural network by minimizing the difference of $c_2(r;\rho_b)$, as obtained via Eq. (1) and functional reduction (2), against the quasiexact reference data for $c_2^{\rm ref}(r)$ from simulations and Ornstein-Zernike inversion. Thereby, the functional derivative in Eq. (1) is performed via automatic differentiation [10,12–14,39]. The aim is to achieve equality,

$$c_2(r;\rho_b) = c_2^{\rm ref}(r;\rho_b). \tag{3}$$

Crucially, during training the neural functional is only evaluated via its Hessian with respect to the input in Eq. (2) at constant density profiles $\rho(\mathbf{r}) = \rho_b$ with varying values of ρ_b. No inhomogeneous density profiles are encountered during training.

By contrast, local direct correlation learning [12,13] starts with the standard liquid state relationship

$$c_1^{\rm ref}(\mathbf{r}) = \ln\rho(\mathbf{r}) + \beta V_{\rm ext}(\mathbf{r}) - \beta\mu, \tag{4}$$

where $c_1^{\rm ref}(\mathbf{r})$ is the one-body direct correlation function that forms the reference. One exploits that all quantities on the right-hand side of Eq. (4) are either known $[\beta,\mu,V_{\rm ext}(\mathbf{r})]$ or are directly accessible in simulations $[\rho(\mathbf{r})$, potentially via force sampling [40]]. Hence one can construct data for $c_1^{\rm ref}(\mathbf{r})$ to act as a reference. The one-body direct correlation function is then transcended to a density functional; recall $c_1(\mathbf{r};[\rho]) = -\delta\beta F_{\rm exc}[\rho]/\delta\rho(\mathbf{r})$. The direct correlation learning [12,13] represents $c_1(\mathbf{r};[\rho])$ at a considered position $\mathbf{r}$ directly as a neural network, with a simple yet very general multilayer perceptron architecture. Supervised training is used to approach

$$c_1(\mathbf{r};[\rho]) = c_1^{\rm ref}(\mathbf{r};[\rho]), \tag{5}$$

where the left-hand side is the output of the neural functional and the right-hand side is the reference obtained via Eq. (4) with simulation input for $\rho(\mathbf{r})$. Training to optimize the matching condition (5) is performed across a range of (several hundred) training systems and hence differing shapes of the density profile $\rho(\mathbf{r})$. The resulting functional can then be applied independently of the system size of the original simulations, thus enabling one to make predictions "beyond the box" [12,13].

Using automatic differentiation to functionally differentiate the one-body direct correlation functional $c_1(\mathbf{r};[\rho])$ yields the two-body direct correlation functional

$$c_2(\mathbf{r},\mathbf{r}';[\rho]) = \frac{\delta c_1(\mathbf{r};[\rho])}{\delta\rho(\mathbf{r}')}, \tag{6}$$

which is now expressed as a first derivative rather than (and formally consistent with) the second functional derivative in Eq. (1). This implies a significant reduction in terms of computational complexity for the case of short-ranged (truncated) interparticle interactions. Performing the functional reduction (2) then gives results for $c_2(r;\rho_b)$ with genuine predictive status as no information about $c_2^{\rm ref}(r;\rho_b)$ has entered the local learning scheme. Hence the bulk direct correlation matching (3) can be used as an *a posteriori* quality check of the neural functional $c_1(\mathbf{r};[\rho])$. Similarly, the exchange symmetry $c_2(\mathbf{r},\mathbf{r}';[\rho]) = c_2(\mathbf{r}',\mathbf{r};[\rho])$, which follows from interchanging the order of derivatives in Eq. (1), is a nontrivial consistency test.

We have kept the presentation general, but in line with Refs. [10–13], we train all neural functionals in planar geometry. We use the supercritical Lennard-Jones fluid at reduced temperature $k_BT/\epsilon = 1.5$ as a generic model system. Planar geometry offers computational benefits and it constitutes arguably the most important type of inhomogeneity due to a multitude of relevant interfacial and substrate applications (including phase transitions). The technical details of the conversion from radial to planar geometry are described in our Supplemental Material [37] and all data are openly available [41].

To combine the virtues of the different approaches, we first incorporate the pair-correlation matching (3) into the local learning method by using Eq. (6) to express $c_2(\mathbf{r},\mathbf{r}';[\rho])$ and then proceed to the functional reduction (2). The loss that is optimized in the supervised machine learning remains based on the one-body direct correlation matching (5) but it is supplemented by a regularizer based on Eq. (3). We find that using the regularizer improves the training dynamics and the overall quality of the results achieved, as is desired and expected when supplying additional information about bulk correlations; details are given in our Supplemental Material [37]. Note the inhomogeneous training data is reused from Ref. [12]. As a test, we display results for the predicted bulk structure from both unregularized one-body learning and pair-matching-regularized one-body learning in Fig. 1. The results with regularization are smoother and the predicted

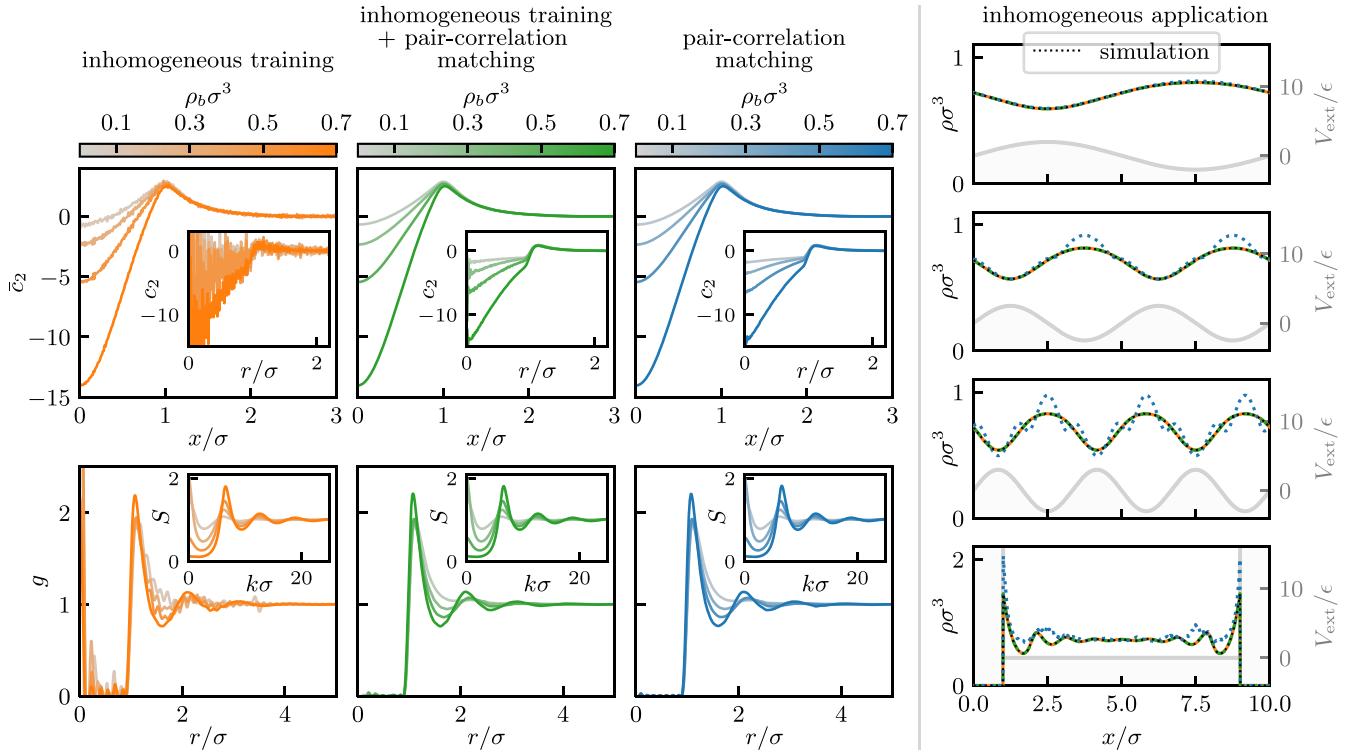

FIG. 1. Neural functional results for bulk pair structure and planar inhomogeneous density profiles. We show bulk results from neural functionals obtained via one-body inhomogeneous training (first column), adding the bulk pair-correlation regularization (second column), and from pure pair-correlation matching (third column). The supercritical Lennard-Jones fluid is investigated at reduced temperature $k_BT/\epsilon = 1.5$ and different scaled bulk densities $\rho_b\sigma^3$ (color bar). Shown is the bulk pair direct correlation function $\bar{c}_2(x;\rho_b)$ in planar geometry as a function of scaled planar distance x/σ obtained via automatic differentiation (6) (top panels) and $c_2(r;\rho_b)$ as a function of radial distance r/σ (insets in top panels). The Ornstein-Zernike relation yields the corresponding pair correlation function $g(r)$ (bottom panels) and bulk structure factor $S(k)$ (insets in bottom panels). Also shown are predictions from the same three neural functionals for planar inhomogeneities (fourth column) as induced by an external potential $V_{\rm ext}(x)$ (gray lines) at $\mu/\epsilon = 0$. Results from the pure pair-correlation matched functional deviate increasingly from those with inhomogeneous training, which remain almost identical to the simulation reference (thin dotted lines) upon increasing inhomogeneity (from top to bottom) even for confinement between parallel hard walls (bottom panel).

radial distribution functions $g(r)$ are closer to zero inside the core, consistent with the expectation based on the supply of bulk information into the scheme.

We have also trained a one-body correlation functional based on pair-correlation matching alone with no information about inhomogeneous systems entering. This is potentially important as it would remove the need to generate inhomogeneous training data, as is argued in Ref. [10]. Within our scheme the method is straightforward to implement by basing the loss solely on the bulk direct correlation matching (3). Similar to Ref. [10], we provide the required integration constants by noting from Eq. (4) that $\beta\mu_{\rm exc} \equiv \beta\mu - \ln\rho_b = -c_1(\mathbf{r};[\rho])$, which the neural functional delivers with constant density input, $\rho(\mathbf{r}) = \rho_b$. Results for the excess chemical potential $\mu_{\rm exc}$ are obtained from simulating bulk systems with $V_{\rm ext}(\mathbf{r}) = 0$. We find, despite apparently successful performance in training metrics, that the pair-correlation-matched neural functional $c_1(\mathbf{r};[\rho])$ is of significantly poorer quality than those from the two local one-body learning methods, which rely on inhomogeneous training data. This becomes clear in predictions for inhomogeneous systems, where the results of the former neural functional fall short of those from the inhomogeneously trained networks, as shown in Fig. 1, column 4. Also the exchange symmetry $c_2(\mathbf{r},\mathbf{r}';[\rho]) = c_2(\mathbf{r}',\mathbf{r};[\rho])$, which is a necessary Cauchy-Riemann-like integrability condition for the existence of a generating functional $F_{\rm exc}[\rho]$ according to Eq. (1), is violated, as shown in the Supplemental Material [37]. The method of Ref. [10] satisfies the exchange symmetry by construction via the neural representation of $F_{\rm exc}[\rho]$ and the interchangeability of the order of the two derivatives in Eq. (1).

We next describe two variants of inhomogeneous one-body learning for modeling directly the generating functional $F_{\rm exc}[\rho]$. We work with a localized excess free-energy integrand $f_{\rm exc}(\mathbf{r};[\rho])$ such that

$$F_{\rm exc}[\rho] = \int d\mathbf{r}\,\rho(\mathbf{r}) f_{\rm exc}(\mathbf{r};[\rho]), \tag{7}$$

where the spatial integration domain is the entire system volume. The general functional integral $\beta F_{\rm exc}[\rho] = -\int \mathcal{D}[\rho] c_1(\mathbf{r};[\rho])$ is parametrized by Eq. (7) when choosing [3,13]

$$-\beta f_{\rm exc}(\mathbf{r};[\rho]) = \int_0^1 da\, c_1(\mathbf{r};[a\rho]), \tag{8}$$

where $a\rho(\mathbf{r})$ is a scaled version of the density profile. The integrals in Eqs. (7) and (8) can be performed numerically

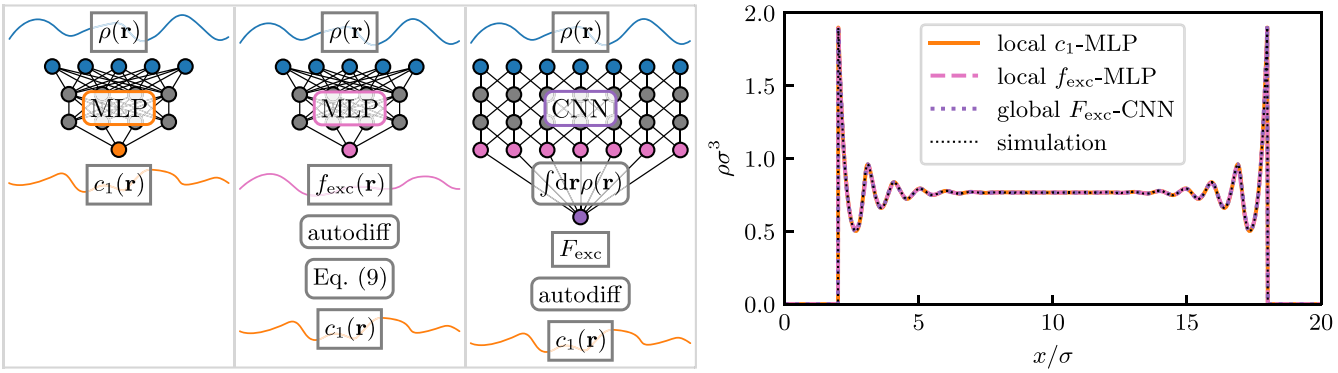

FIG. 2. Illustration of the three different network architectures for neural functionals. Shown are the multilayer perceptrons (MLP) for local learning of $c_1(\mathbf{r};[\rho])$ (first column) and for $f_{\rm exc}(\mathbf{r};[\rho])$ (second column) and the convolutional neural network (CNN) for $F_{\rm exc}[\rho]$. Corresponding predictions for the scaled density profiles $\rho(x)\sigma^3$ (fourth column) are shown for the supercritical Lennard-Jones fluid with $\mu/\epsilon = 1$ and $k_BT/\epsilon = 1.5$ confined between two hard walls at $x/\sigma = 2$ and 18. Results for $\rho(x)$ are shown from density-functional minimization using $c_1(x;[\rho])$ obtained from local learning according to Eq. (5), from local learning of $f_{\rm exc}(x;[\rho])$ via Eq. (9), and from inhomogeneous one-body training of the convolutional neural network representation of $F_{\rm exc}[\rho]$. The predictions from the three different neural functionals are identical to the simulation reference (thin dotted line) on the scale of the plot.

with high efficiency to obtain the value of the excess free energy [12,13].

We obtain the corresponding one-body direct correlation functional by recalling that $c_1(\mathbf{r};[\rho]) = -\delta\beta F_{\rm exc}[\rho]/\delta\rho(\mathbf{r})$ and inserting $F_{\rm exc}[\rho]$ in the form (7), which gives

$$c_1(\mathbf{r};[\rho]) = -\beta f_{\rm exc}(\mathbf{r};[\rho]) - \int d\mathbf{r}'\rho(\mathbf{r}')\frac{\delta\beta f_{\rm exc}(\mathbf{r};[\rho])}{\delta\rho(\mathbf{r}')}. \tag{9}$$

We have exchanged $\mathbf{r}$ and $\mathbf{r}'$ according to the symmetry $\delta f_{\rm exc}(\mathbf{r}';[\rho])/\delta\rho(\mathbf{r}) = \delta f_{\rm exc}(\mathbf{r};[\rho])/\delta\rho(\mathbf{r}')$; this identity follows from differentiating Eq. (8) and from the exchange symmetry of $c_2(\mathbf{r},\mathbf{r}';[\rho])$. The form (9) is suitable for the application to local learning.

In our first free-energy method we represent $f_{\rm exc}(\mathbf{r};[\rho])$ as a neural functional and perform one-body direct correlation matching via Eq. (5) on the basis of Eq. (9), where automatic differentiation is used to carry out the functional derivative. The second free-energy method uses a convolutional neural network architecture similar to that of Ref. [10] to implement via $f_{\rm exc}(\mathbf{r};[\rho])$ the excess free-energy functional $F_{\rm exc}[\rho]$ in the structure of Eq. (7). We thereby perform no pooling (coarse graining, as is common practice [36]) and design a final layer that represents $f_{\rm exc}(\mathbf{r};[\rho])$, which is then multiplied by $\rho(\mathbf{r})$ and integrated according to Eq. (7) for the global value of $F_{\rm exc}[\rho]$; details are given in the Supplemental Material [37]. This specific convolutional neural network design extends naturally to arbitrary system sizes, thus replicating the capability of the local correlation learning for making predictions beyond the box [12,13]. Illustrations of the different neural functionals together with representative results from the three methods for planar density profiles $\rho(x)$ of a Lennard-Jones fluid confined between parallel hard walls are shown Fig. 2. Having been trained with the same inhomogeneous one-body profiles from grand canonical Monte Carlo simulations [12], all methods give an excellent account of the highly inhomogeneous spatial fluid structure.

We conclude with several conceptual points related to the inherent functional reduction (2) of bulk pair-correlation matching [10]. The feeding of knowledge of $g(r)$ into the neural functional constitutes a very significant amount of information about the system under consideration. Henderson's uniqueness theorem [42] formally ensures that for systems interacting solely with a pair potential $\phi(r)$, perfect knowledge of $g(r)$, at a single state point, is sufficient to determine $\phi(r)$. Density-functional theory [1] then implies that $F_{\rm exc}[\rho]$ is uniquely determined in principle, given the information that no higher-body interparticle interactions are present [43].

However, it is difficult to see why a neural functional trained through the aperture of the functional reduction (2) would render Henderson's theorem operational. As the density-functional framework (1)–(3), utilized within the pair-correlation matching, applies equally to a system with many-body interparticle interactions, we see no mechanism in Ref. [10] that would intrinsically reduce $F_{\rm exc}[\rho]$ to pair interactions only (see Ref. [43] for a related discussion in the context of simulation work). This situation cannot be remedied by supplying $g(r)$ in more abundance and with higher precision. In contrast, the one-body matching condition (5) allows for full exploration of the entire density-functional dependence for any given form of the underlying Hamiltonian, which gives a formal mechanism to rationalize the high quality of our correspondingly obtained results. Our locally trained functional $c_1(\mathbf{r};[\rho])$ [41] does not suffer from the somewhat limited extrapolation capabilities reported for the inhomogeneous training method of Ref. [10]. We also find robust values when evaluating $c_1(\mathbf{r};[\rho])$ at positions $\mathbf{r}$ inside a hard wall, where the local density vanishes, which corroborates the extrapolation capability to unseen cases. Such regions, per construction, do not contribute directly to the free energy integral (7), due to the multiplication with $\rho(\mathbf{r}) = 0$.

In future work it would be interesting to further compare the different one-body learning strategies to each other. One could also attempt to utilize test particle concepts [44] to identify $\rho(\mathbf{r}) = \rho_b g(r)$ upon setting $V_{\rm ext}(\mathbf{r}) = \phi(r)$, as also explored dynamically [45–47] and quantum mechanically [48]. Despite the implied restriction to spherical symmetry, this would allow simulation results

for pair-correlation functions $g(r)$ to be used as genuinely inhomogeneous one-body density profiles $\rho(\mathbf{r})$. The three-body direct correlation functional follows from (Hessian) automatic differentiation with respect to the density profile according to $c_3(\mathbf{r}, \mathbf{r}', \mathbf{r}''; [\rho]) = \delta^2 c_1(\mathbf{r}; [\rho])/\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')$ [12]. Then one could exploit Noether invariance sum rules that relate to spatial gradients of $c_2(\mathbf{r}, \mathbf{r}'; [\rho])$ [13,49]. Furthermore, the use of the force density integral $\rho(\mathbf{r})\nabla c_1(\mathbf{r}; [\rho]) = -\int d\mathbf{r}' \rho_2(\mathbf{r}, \mathbf{r}')\nabla\beta\phi(|\mathbf{r} - \mathbf{r}'|)$ [1,2,50,51] could be beneficial for making progress despite an increase in complexity [52,53]. Localized forces are also relevant for functional treatments of general observables [14] and for nonequilibrium flow [11,15].

We thank J. Dijkman, M. Welling, J.-W. van de Meent, B. Ensing, K. Burke, M. Dijkstra, R. Evans, and S. M. Kampa for useful discussions.

[1] R. Evans, The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids, Adv. Phys. **28**, 143 (1979).
[2] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 4th ed. (Academic Press, London, 2013).
[3] R. Evans, in *Fundamentals of Inhomogeneous Fluids*, edited by D. Henderson (Dekker, New York, 1992), Chap. 3.
[4] R. Evans, M. Oettel, R. Roth, and G. Kahl, New developments in classical density functional theory, J. Phys.: Condens. Matter **28**, 240401 (2016).
[5] S.-C. Lin and M. Oettel, A classical density functional from machine learning and a convolutional neural network, SciPost Phys. **6**, 025 (2019).
[6] S.-C. Lin, G. Martius, and M. Oettel, Analytical classical density functionals from an equation learning network, J. Chem. Phys. **152**, 021102 (2020).
[7] P. Cats, S. Kuipers, S. de Wind, R. van Damme, G. M. Coli, M. Dijkstra, and R. van Roij, Machine-learning free-energy functionals using density profiles from simulations, APL Mater. **9**, 031109 (2021).
[8] P. Yatsyshin, S. Kalliadasis, and A. B. Duncan, Physics-constrained Bayesian inference of state functions in classical density-functional theory, J. Chem. Phys. **156**, 074105 (2022).
[9] A. Malpica-Morales, P. Yatsyshin, M. A. Duran-Olivencia, and S. Kalliadasis, Physics-informed Bayesian inference of external potentials in classical density functional theory, J. Chem. Phys. **159**, 104109 (2023).
[10] J. Dijkman, M. Dijkstra, R. van Roij, M. Welling, J.-W. van de Meent, and B. Ensing, Learning neural free-energy functionals with pair-correlation matching, arXiv:2403.15007.
[11] D. de las Heras, T. Zimmermann, F. Sammüller, S. Hermann, and M. Schmidt, Perspective: How to overcome dynamical density functional theory, J. Phys.: Condens. Matter **35**, 271501 (2023).
[12] F. Sammüller, S. Hermann, D. de las Heras, and M. Schmidt, Neural functional theory for inhomogeneous fluids: Fundamentals and applications, Proc. Natl. Acad. Sci. USA **120**, e2312484120 (2023).
[13] F. Sammüller, S. Hermann, and M. Schmidt, Why neural functionals suit statistical mechanics, J. Phys.: Condens. Matter **36**, 243002 (2024).
[14] F. Sammüller, S. Robitschko, S. Hermann, and M. Schmidt, Hyperdensity functional theory of soft matter, Phys. Rev. Lett. (2024).
[15] T. Zimmermann, F. Sammüller, S. Hermann, M. Schmidt, and D. de las Heras, Neural force functional for non-equilibrium many-body colloidal systems, arXiv:2406.03606.
[16] A. Simon, J. Weimar, G. Martius, and M. Oettel, Machine learning of a density functional for anisotropic patchy particles, J. Chem. Theory Comput. **20**, 1062 (2024).
[17] R. Nagai, R. Akashi, S. Sasaki, and S. Tsuneyuki, Neural-network Kohn-Sham exchange-correlation potential and its out-of-training transferability, J. Chem. Phys. **148**, 241737 (2018).
[18] J. Schmidt, C. L. Benavides-Riveros, and M. A. L. Marques, Machine learning the physical nonlocal exchange-correlation functional of density-functional theory, J. Phys. Chem. Lett. **10**, 6425 (2019).
[19] Y. Zhou, J. Wu, S. Chen, and G. Chen, Toward the exact exchange-correlation potential: A three-dimensional convolutional neural network construct, J. Phys. Chem. Lett. **10**, 7264 (2019).
[20] R. Nagai, R. Akashi, and O. Sugino, Completing density functional theory by machine learning hidden messages from molecules, npj Comput. Mater. **6**, 43 (2020).
[21] L. Li, S. Hoyer, R. Pederson, R. Sun, E. D. Cubuk, P. Riley, and K. Burke, Kohn-Sham equations as regularizer: Building prior knowledge into machine-learned physics, Phys. Rev. Lett. **126**, 036401 (2021).
[22] H. Li, N. Zou, M. Ye, R. Xu, X. Gong, and W. Duan, Deep-learning density functional theory Hamiltonian for efficient ab initio electronic-structure calculation, Nat. Comput. Sci. **2**, 367 (2022).
[23] R. Pederson, B. Kalita, and K. Burke, Machine learning and density functional theory, Nat. Rev. Phys. **4**, 357 (2022).
[24] J. Gedeon, J. Schmidt, M. J. P. Hodgson, J. Wetherell, C. L. Benavides-Riveros, and M. A. L. Marques, Machine learning the derivative discontinuity of density-functional theory, Mach. Learn.: Sci. Technol. **3**, 015011 (2022).
[25] M. M. Kelley, J. Quinton, K. Fazel, N. Karimitari, C. Sutton, and R. Sundararaman, Bridging electronic and classical density-functional theory using universal machine-learned functional approximations, arXiv:2405.20270.
[26] Y. Rosenfeld, Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing, Phys. Rev. Lett. **63**, 980 (1989).
[27] H. Hansen-Goos and R. Roth, Density functional theory for hard-sphere mixtures: The White Bear version mark II, J. Phys.: Condens. Matter **18**, 8413 (2006).
[28] R. Roth, Fundamental measure theory for hard-sphere mixtures: A review, J. Phys.: Condens. Matter **22**, 063102 (2010).
[29] R. Evans, M. C. Stewart, and N. B. Wilding, A unified description of hydrophilic and superhydrophobic surfaces in terms of the wetting and drying transitions of liquids, Proc. Natl. Acad. Sci. USA **116**, 23901 (2019).

[30] M. K. Coe, R. Evans, and N. B. Wilding, Density depletion and enhanced fluctuations in water near hydrophobic solutes: Identifying the underlying physics, Phys. Rev. Lett. **128**, 045501 (2022).

[31] M. K. Coe, R. Evans, and N. B. Wilding, Understanding the physics of hydrophobic solvation, J. Chem. Phys. **158**, 034508 (2023).

[32] N. D. Mermin, Thermal properties of the inhomogeneous electron gas, Phys. Rev. **137**, A1441 (1965).

[33] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 3rd ed. (Academic Press, London, 2023).

[34] N. B. Wilding, Computer simulation of fluid phase transitions, Am. J. Phys. **69**, 1147 (2001).

[35] A. V. Brukhno, J. Grant, T. L. Underwood, K. Stratford, S. C. Parker, J. A. Purton, and N. B. Wilding, DL_MONTE: A multipurpose code for Monte Carlo simulation, Mol. Simul. **47**, 131 (2021).

[36] F. Chollet, *Deep Learning with Python* (Manning, Shelter Island, 2017).

[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.110.L032601 for details of the Ornstein-Zernike route, neural networks, and training procedures, which contains Ref. [38].

[38] N. B. Wilding, Critical-point and coexistence-curve properties of the Lennard-Jones fluid: A finite-size scaling study, Phys. Rev. E **52**, 602 (1995).

[39] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Automatic differentiation in machine learning: A survey, J. Mach. Learn. Res. **18**, 1 (2018).

[40] B. Rotenberg, Use the force! Reduced variance estimators for densities, radial distribution functions, and local mobilities in molecular simulations, J. Chem. Phys. **153**, 150902 (2020).

[41] Code, simulation data, and neural functionals are available at https://github.com/sfalmo/local-pair-matching.

[42] R. L. Henderson, A uniqueness theorem for fluid pair correlation functions, Phys. Lett. A **49**, 197 (1974).

[43] R. Evans, Comment on reverse Monte Carlo simulation, Mol. Simul. **4**, 409 (1990).

[44] J. K. Percus, Approximation methods in classical statistical mechanics, Phys. Rev. Lett. **8**, 462 (1962).

[45] L. L. Treffenstädt and M. Schmidt, Universality in driven and equilibrium hard sphere liquid dynamics, Phys. Rev. Lett. **126**, 058002 (2021).

[46] L. L. Treffenstädt, T. Schindler, and M. Schmidt, Dynamic decay and superadiabatic forces in the van Hove dynamics of bulk hard sphere fluids, SciPost Phys. **12**, 133 (2022).

[47] M. Schmidt, Power functional theory for many-body dynamics, Rev. Mod. Phys. **94**, 015007 (2022).

[48] R. J. McCarty, D. Perchak, R. Pederson, R. Evans, Y. Qiu, S. R. White, and K. Burke, Bypassing the energy functional in density functional theory: Direct calculation of electronic energies from conditional probability densities, Phys. Rev. Lett. **125**, 266401 (2020).

[49] S. Hermann and M. Schmidt, Noether's theorem in statistical mechanics, Commun. Phys. **4**, 176 (2021).

[50] J. Yvon, La Théorie Statistique des Fluides et l'Équation d'État, Actualités Scientifiques et Industrielles Vol. 203 (Hermann & Cie, Paris, 1935).

[51] M. Born and H. S. Green, A general kinetic theory of liquids I. The molecular distribution functions, Proc. R. Soc. London Ser. A **188**, 10 (1946).

[52] S. M. Tschopp, F. Sammüller, S. Hermann, M. Schmidt, and J. M. Brader, Force density functional theory in- and out-of-equilibrium, Phys. Rev. E **106**, 014115 (2022).

[53] F. Sammüller, S. Hermann, and M. Schmidt, Comparative study of force-based classical density functional theory, Phys. Rev. E **107**, 034109 (2023).