OPEN ACCESS
IOP Publishing
Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **37** (2025) 083001 (23pp)
https://doi.org/10.1088/1361-648X/ad98da

Topical Review

Why hyperdensity functionals describe any equilibrium observable

Florian Sammüller* **and Matthias Schmidt***

Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, D-95447 Bayreuth, Germany

E-mail: Florian.Sammueller@uni-bayreuth.de and Matthias.Schmidt@uni-bayreuth.de

Received 18 October 2024
Accepted for publication 29 November 2024
Published 12 December 2024

Abstract

We give an introductory account of the recent hyperdensity functional theory for the equilibrium statistical mechanics of soft matter systems (Sammüller *et al* 2024 *Phys. Rev. Lett.* **133** 098201). Hyperdensity functionals give access to the behaviour of arbitrary thermal observables in spatially inhomogeneous equilibrium many-body systems. The approach is based on classical density functional theory applied to an extended ensemble using standard functional techniques. The associated formally exact generalized Mermin-Evans functional relationships can be represented accurately by neural functionals. These neural networks are trained via simulation-based supervised machine learning and they allow one to carry out efficient functional calculus using automatic differentiation and numerical functional line integration. Exact sum rules, including hard wall contact theorems and hyperfluctuation Ornstein–Zernike equations, interrelate the different correlation functions. We lay out close connections to hyperforce correlation sum rules (Robitschko *et al* 2024 *Commun. Phys.* **7** 103) that arise from statistical mechanical gauge invariance (Müller *et al* 2024 *Phys. Rev. Lett.* **133** 217101). Further quantitative measures of collective self-organization are provided by hyperdirect correlation functionals and spatially resolved hyperfluctuation profiles. The theory facilitates to gain deep insight into the inherent structuring mechanisms that govern the behaviour of both simple and complex order parameters in coupled many-body systems.

Keywords: classical density functional theory, liquid state theory, fluctuation profiles, force sampling, Ornstein–Zernike relation, hyperforce correlations, neural functionals

1. Introduction

The Mermin-Evans theorem [1–6] of density functional theory provides the finite temperature and classical mechanical generalization of the pivotal Hohenberg-Kohn proof [7, 8]. These theorems enable a full representation of the equilibrium statistical mechanics of particle-based systems via systematically constructed functional dependencies. In particular the eponymous one-body density profile $\rho(\mathbf{r})$, where $\mathbf{r}$ denotes spatial position, plays a leading role as a variational quantity. Higher-body correlation functions are accessible both via the Ornstein–Zernike relationship [5, 9] and Percus' test particle route [5, 10] to represent the full correlated physics that emerges from the underlying coupling of the individual particles. Full thermodynamic information, including bulk and interfacial contributions, is thereby accessible.

* Authors to whom any correspondence should be addressed.

In typical applications one is interested in situations where spatial inhomogeneity is induced by the action of an external potential $V_{\rm ext}(\mathbf{r})$. Appropriate choices of the form of $V_{\rm ext}(\mathbf{r})$ allow one to model a broad class of different types of external influence that is exerted onto a system, including walls, confinement, the behaviour around solutes, *etc*. The microscopic degrees of freedom of the system are thereby coupled through an interparticle interaction potential $u(\mathbf{r}^N)$ that depends on all coordinates $\mathbf{r}^N = \mathbf{r}_1, \ldots, \mathbf{r}_N$ of the N particles. Typically the thermodynamic state point is given via prescribing values of the chemical potential μ and of the absolute temperature T.

From an elementary statistical mechanical point of view it is in principle a standard task to predict the density profile $\rho(\mathbf{r})$, which in the classical realm is straightforward to accomplish in computer simulations. One needs to average the density operator $\hat{\rho}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$, where the sum over i runs over all N particles, and $\delta(\cdot)$ indicates the Dirac distribution. The thermal ensemble average is indicated by $\langle\cdot\rangle$ and it can readily be realized via importance sampling on the basis of e.g. grand canonical Monte Carlo simulations [11–13]. Thereby powerful histogram [12], force sampling [11, 14–21] and mapped averaging [22–34] techniques are available.

Given this quite pedestrian status it can seem surprising that the density profile $\rho(\mathbf{r}) = \langle\hat{\rho}(\mathbf{r})\rangle$ is of Nobel-prize winning format in the quantum realm [8]. The situation can seem even more perplexing as $\rho(\mathbf{r})$ is the only relevant variational variable in density functional theory, which hence seemingly lacks any explicit occurrence of two- and higher-body correlation functions. Using any further problem-specific, tailor-made order parameters is also quite alien to the framework.

That these apparent deficiencies can all be remedied by the mere inversion of the functional map, i.e. realizing and establishing $\rho(\mathbf{r}) \to V_{\rm ext}(\mathbf{r})$, can seem mysterious. Possibly at the centre of the mystery is the density functionalists' credo that 'everything is a density functional', which poignantly expresses that, for given interparticle interaction potential $u(\mathbf{r}^N)$, from knowledge of $\rho(\mathbf{r})$ the Hamiltonian itself can be reconstructed. Once the Hamiltonian is known along with the thermodynamic conditions, *any* equilibrium property, no matter how complex or intricate, is known in principle and has thus become a density functional. This formal structure can certainly seem surprising and we here wish to lay out its concrete consequences and route to practical implementation.

Functional relationships have acquired new and compelling relevance in light of the recent neural functional theory [35–44]. This hybrid approach utilizes many-body simulations to generate data for supervised training of an artificial neural network, which then acts as a neural functional. A number of features set this approach apart from more generic machine-learning methods [45–48], from physics-informed machine learning in liquid state integral equation theory [49–51], as well as from other uses of machine learning in classical [52–62] and in quantum density functional theory [63–70]. As is argued in [35–44], the neural functional theory constitutes a genuine theoretical framework that permits one to carry out deep functional calculus and to obtain a very complete picture of the physics under investigation.

We summarize several key features of neural functionals. (i) Learning of the relationship between input and output data pairs is based on a rigorous mathematical relationship, which is known from first principles to exist and to be unique, see the discussion given in [38]. (ii) Local learning [35–38] facilitates data-efficient training and subsequent 'beyond-the-box' application of the resulting neural functional to challenging multi-scale problems. (iii) Efficient implementation of functional calculus is provided by automatic functional differentiation [71] and fast numerical functional integration. (iv) The neural representations of both direct correlation functionals and of free energy functionals are accurate and performant. (v) The internal consistencies of a neural functional can be tested via numerical evaluation of a wide variety of exact statistical mechanical sum rules.

The body of literature addressing statistical mechanical sum rules, i.e. exact identities that hold universally, is both large and diverse; see e.g. [72–74]. The recent thermal Noether invariance theory provides a systematic approach for both the derivation and the classification of sum rules [75–84]. The thermal Noether invariance is thereby inherent in the very foundations of the statistical mechanics and constitutes a gauge transformation for statistical mechanical microstates [83, 84]; see the very recent Viewpoint given by Rotenberg [85]. The approach is free of simplifying assumptions and approximations. Technically the symmetry is an invariance of the phase space integral under specific transformation operations of the phase space variables. The Noether framework not only generates exact identities, but it also acts as a construction device to generate the specific correlation functions for which these identities hold. The correlation functions range from standard density-based statistical mechanical correlations to force-based observables, which were shown to shine new light on liquid structure even in bulk [80, 81].

In a remarkable contribution Hirschfelder [86] generalized the standard virial theorem [5], which dates back to Clausius and the very origins of thermodynamics, to include an arbitrary observable $\hat{A}$, as represented by a phase space function in the present classical context. The hyperforce theory of [82] performs a similar generalization of the force-balance relationship, as expressed in local form via a hierarchy of equations due to Yvon [87] and Born and Green [88]. Hyperforce sum rules hold both globally as well as locally resolved in position and they couple in specific ways the general observable $\hat{A}$ with the fundamental degrees of freedom of the system via force- and density-based correlation functions.

We recall that the one-body force balance equation integrates itself very naturally into density functional theory, where it is recovered as the spatial gradient of the fundamental Euler–Lagrange equation [2, 6, 79]. An analogous correspondence exists for the hyperforce framework [82–84], as this is mirrored and complemented by the corresponding hyperdensity functional theory [40]. This approach facilitates the explicit

construction of the mean of the observable $\hat{A}$ as a density functional, i.e. $A[\rho]$, where the brackets indicate the functional relationship. Here we give an introductory and extended account of the hyperdensity functional theory [40], including a detailed discussion of its relationship with the hyperforce sum rules of [82] and the underlying gauge invariance concept [83–85].

Briefly, hyperdensity functional theory [40] allows one to address the equilibrium behaviour of complex order parameters $\hat{A}$ in a tightly integrated and concrete statistical mechanical framework. The equilibrium average of $\hat{A}$, expressed as a density functional $A[\rho]$, is associated with a hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ and a hyperfluctuation profile $\chi_A(\mathbf{r})$, both of which are specific to the form of $\hat{A}$. The one-body hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ plays a role similar to that of the one-body direct correlation functional $c_1(\mathbf{r};[\rho])$ of standard density functional theory. The hyperfluctuation profile $\chi_A(\mathbf{r})$ can be viewed as a generalization of thermodynamic fluctuation profiles [89–96], such as the well-studied local compressibility $\chi_\mu(\mathbf{r})$ [89–91, 96].

Together with the standard two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$, the correlation functions $c_A(\mathbf{r};[\rho])$ and $\chi_A(\mathbf{r})$ are connected via an exact hyper-Ornstein–Zernike relation [40], which generalizes the standard inhomogeneous two-body Ornstein–Zernike relation [2, 5, 6] to general observables $\hat{A}$. The hyper-Ornstein–Zernike relation is of relatively simple one-body form (two-body functions feature only inside of a spatial integral) and the mathematical structure is akin to the one-body fluctuation-Ornstein–Zernike relationships [94, 95]. The fluctuation-Ornstein–Zernike equation for the local compressibility $\chi_\mu(\mathbf{r})$ [94, 95] was shown to deliver efficiently accurate results in demanding drying situations using neural density functional methods together with automatic differentiation [39]. An illustration of the relationship between the key quantities is shown in figure 1.

In exemplary applications to concrete systems, the hyperdensity functional was used for the investigation of clustering properties of standard model fluids. The behaviour of one-dimensional hard rods was compared to that of square-well attractive rods and of hard sphere fluids in three dimensions. Our definition of the clustering follows a standard procedure (see, e.g. [97]). One starts with a bonding criterion that declares two particles as bonded provided that their mutual distance is smaller than some cut-off (taken to be 1.2σ with σ indicating the particle size). Although the bonding criterion itself is a two-body function, the resulting graph structure of bonded particles is non-trivial with naive implementations requiring iterated passes over all particles, but more efficient algorithms can be used [40, 97]. In a second step the number of particles in each cluster is counted and the largest such number is searched for, which then is taken as the observable $\hat{A}$. Each microstate gives a unique value of the size of the largest cluster $\hat{A}$, hence $\hat{A}$ is indeed a phase space function, as is required in the hyperdensity functional framework. However, in contrast to standard phase space functions such as the Hamiltonian *etc*, no closed expression for $\hat{A}$ is available and hence $\hat{A}$ constitutes an *algorithmically defined* observable. Nevertheless, as

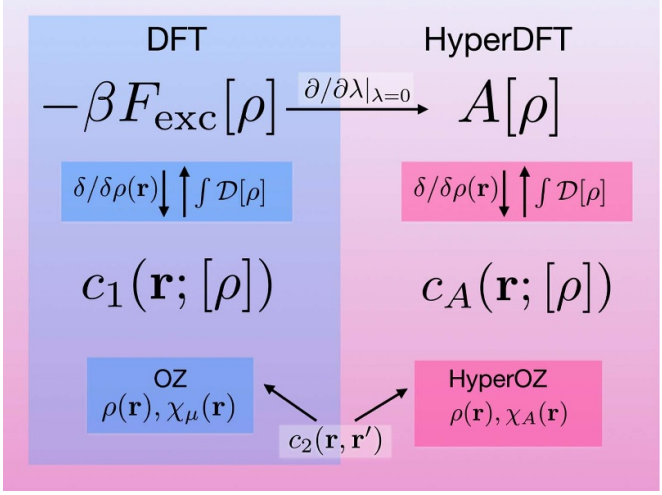

Figure 1. Functional relationships of classical density functional theory (left column) and the hyperdensity functional generalization (right column). The scaled excess free energy functional $-\beta F_{\rm exc}[\rho]$ is related to the one-body direct correlation functional $c_1(\mathbf{r};[\rho])$ via functional differentiation, $\delta/\delta\rho(\mathbf{r})$, and, inversely, by functional integration, $\int\mathcal{D}[\rho]$. Analogous relationships connect the hyperdensity functional $A[\rho]$ that expresses the mean of any given observable $\hat{A}$ and the hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$. The local compressibility $\chi_\mu(\mathbf{r})$ and the hyperfluctuation profile $\chi_A(\mathbf{r})$ play analogous roles and they are respectively connected via Ornstein–Zernike and hyper-Ornstein–Zernike relations, which both feature the two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$. An extended ensemble allows differentiating with respect to the coupling parameter, $\partial/\partial\lambda|_{\lambda=0}$.

we lay out, the thermal mean $A=\langle\hat{A}\rangle$ is a unique hyperdensity functional.

The paper is structured as follows. In section 2 we describe the hyperdensity functional theory of [40]. Specifically, the extended ensemble for the treatment of general observables $\hat{A}$ is described in section 2.1. The average $\langle\hat{A}\rangle$ and the local hyperfluctuation profile $\chi_A(\mathbf{r})$ are introduced in section 2.2. The Mermin-Evans minimization principle of classical density functional theory, as is at the heart of the hyperdensity generalization, is described in section 2.3. A brief account of Levy's constrained search method for the construction of the intrinsic free energy density density functional is given in section 2.4. The hyper-Ornstein–Zernike relation is derived in section 2.5. General observables are expressed as hyperdensity functionals in section 2.6. The wall hypercontact theorem is presented in section 2.7.

In section 3 we relate the hyperdensity functional theory to the hyperforce theory of [82] as it arises from statistical mechanical gauge invariance [83, 84]. In particular we derive in section 3.1 the exact one-body hyperforce balance relationship [82] from the extended ensemble, which is an alternative both to the Noether functional invariance method [82] and to the phase space operator approach [83, 84]. In section 3.2 we present exact sum rules that connect the hyperforce correlation functions with the hyperdensity functionals. In section 4

we treat specific simple cases of choice of the observable $\hat{A}$ including one-body (section 4.1) and two-body forms (section 4.2). We describe applications in section 5, including a description of the workflow that is required in the hyperdensity functional studies in section 5.1, the training of neural hyperdensity functionals in section 5.2, and a model application to clustering of hard spheres in section 5.3. We give our conclusions and an outlook in section 6.

2. Hyperdensity functional theory

2.1. Extended statistical ensemble

We start by introducing the generalized equilibrium grand ensemble used in [40] as a basis for the statistical mechanics. The extended ensemble facilitates to incorporate the statistical behaviour of a given observable $\hat{A}$, which can be arbitrarily chosen, into the density functional framework. The use of an extended ensemble is a standard means of generalization, see e.g. [98, 99]. In their remarkable contribution Anero *et al* [99] formulated a functional thermodynamics as a generalization of dynamic density functional theory [2, 6, 41, 100, 101] to non-isothermal situations.

Here we are ultimately interested in the equilibrium properties of systems with Hamiltonians of the following standard form:

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m} + u\left(\mathbf{r}^N\right) + \sum_i V_{\mathrm{ext}}\left(\mathbf{r}_i\right), \tag{1}$$

where $\mathbf{p}_i$ denotes the momentum of particle i, the variable m indicates the particle mass, $u(\mathbf{r}^N)$ is the interparticle interaction potential as a function of all position coordinates $\mathbf{r}^N = \mathbf{r}_1, \ldots, \mathbf{r}_N$, the external one-body potential is $V_{\mathrm{ext}}(\mathbf{r})$, and the sums over i run over all N particles, such that $i = 1, \ldots, N$.

We wish to address the properties of a given phase space function $\hat{A}(\mathbf{r}^N)$ which represents a physically relevant observable, such as an order parameter that characterizes the behaviour of the system specified by the Hamiltonian (1). We consider the system to be coupled to both a heat bath with absolute temperature T and to a particle reservoir that sets the chemical potential μ. Hence the corresponding standard grand canonical Boltzmann factor is $\mathrm{e}^{-\beta(H-\mu N)}$, where $\beta = 1/(k_{\mathrm{B}}T)$ with k_{B} denoting the Boltzmann constant. According to standard procedure the partition sum is $\mathrm{Tr}\mathrm{e}^{-\beta(H-\mu N)}$ and the grand potential is $-k_{\mathrm{B}}T\ln\mathrm{Tr}\mathrm{e}^{-\beta(H-\mu N)}$. Here the classical 'trace' operation is defined in the standard way as $\mathrm{Tr} = \sum_{N=0}^{\infty}(N!h^{dN})^{-1}\int \mathrm{d}\mathbf{r}^N\mathrm{d}\mathbf{p}^N$, where h denotes the Planck constant, d is the space dimension, and $\int \mathrm{d}\mathbf{r}^N\mathrm{d}\mathbf{p}^N$ indicates the phase space integral over all position and momentum variables.

In order to address the statistical behaviour of a given observable $\hat{A}(\mathbf{r}^N)$ we consider an extended setup, which can be described in two equivalent ways. First, we consider the Hamiltonian (1) to be unchanged but extend the statistical ensemble itself, such that its generalized Boltzmann factor is $\mathrm{e}^{-\beta(H-\mu N)+\lambda\hat{A}}$, where the coupling parameter λ regulates the degree of influence of $\hat{A}$ on the probability distribution function. The corresponding normalization factor is the extended partition sum given by $\Xi = \mathrm{Tr}\mathrm{e}^{-\beta(H-\mu N)+\lambda\hat{A}}$. Averages in the extended ensemble are then obtained as $\langle\cdot\rangle = \mathrm{Tr}\cdot\mathrm{e}^{-\beta(H-\mu N)+\lambda\hat{A}}/\Xi$ and the extended grand potential (or extended grand canonical free energy) is $\Omega = -k_{\mathrm{B}}T\ln\Xi$ and this depends per construction parametrically on λ. Despite the generalization, we remain ultimately interested in the limit $\lambda \to 0$, which however is taken typically only after carrying out the appropriate derivatives.

The second route leads to the identical statistical physics and it is based on modifying the Hamiltonian itself rather than merely the statistical ensemble used for the description. Here one defines an extended Hamiltonian

$$H_A = H - \lambda\hat{A}/\beta, \tag{2}$$

where H is the original Hamiltonian according to equation (1) and the observable $\hat{A}$ is hence considered to actually contribute to the interactions that define the system. Again the parameter λ tunes the strength of these now extended interactions and as before the physical units of λ are those of the inverse to $\hat{A}(\mathbf{r}^N)$ such that the second term in equation (2) has units of energy, which arise from $1/\beta = k_{\mathrm{B}}T$.

In this second route the extended Hamiltonian (2) is fed in a straightforward way into the standard grand ensemble machinery. Hence the Boltzmann factor is $\mathrm{e}^{-\beta(H_A-\mu N)} = \mathrm{e}^{-\beta(H-\lambda\hat{A}/\beta-\mu N)} = \mathrm{e}^{-\beta(H-\mu N)+\lambda\hat{A}}$, where in the first step we have replaced H_A according to equation (2) and in the second step have re-ordered the terms. The resulting expression is identical to the Boltzmann factor of the extended ensemble according to the above first route. Following the standard procedure the resulting partition sum is then $\Xi = \mathrm{Tr}\mathrm{e}^{-\beta(H_A-\mu N)}$, the thermal equilibrium average is defined as $\langle\cdot\rangle = \mathrm{Tr}\mathrm{e}^{-\beta(H_A-\mu N)}/\Xi$, and the grand potential is $\Omega = -k_{\mathrm{B}}T\ln\Xi$. These expressions are all equivalent to their respective above counterparts from the first route via the ensemble extension.

2.2. Observables beyond the density profile

As laid out in the introduction, density functional theory assigns a special role to the equilibrium one-body density distribution $\rho(\mathbf{r})$. When expressed as a thermal average the density profile is simply given as

$$\rho\left(\mathbf{r}\right) = \langle\hat{\rho}\left(\mathbf{r}\right)\rangle. \tag{3}$$

The one-body density 'operator' (phase space function) $\hat{\rho}(\mathbf{r})$ has the standard form

$$\hat{\rho}\left(\mathbf{r}\right) = \sum_i \delta\left(\mathbf{r} - \mathbf{r}_i\right), \tag{4}$$

with $\delta(\cdot)$ denoting the Dirac distribution in d dimensions. We consider the average in equation (3) to be taken over the extended ensemble. Hence this is specific to the form of $\hat{A}(\mathbf{r}^N)$ and it

depends parametrically on the value of the coupling parameter λ. Taking $\lambda \to 0$ restores the original Hamiltonian, $H_A \to H$, and thus $\rho(\mathbf{r})$ reduces to the density profile that corresponds to the 'real' Hamiltonian H.

Besides thermodynamic quantities, such as the free energy and the pressure, density functional theory encompasses in principle two- and higher-body correlation functions, as defined in generalization of equation (3) via the Ornstein–Zernike route; we sketch the mathematical structure below in section 2.5. For systems that interact only with pairwise interparticle forces this allows one to express the locally resolved force density within a force-based formulation of density functional theory [79, 102]. Furthermore, hyperforce sum rules can be exploited to obtain a range of additional correlation functions from standard density functional calculations; we refer the Reader to the conclusions of [82] for a discussion of these opportunities.

Here we generalize further and hence are interested in the behaviour of a given observable $\hat{A}(\mathbf{r}^N)$. We first consider its mean value

$$A = \langle \hat{A} \rangle, \tag{5}$$

which is a global quantity provided that $\hat{A}(\mathbf{r}^N)$ itself carries no further dependence on position. Alternatively, in case $\hat{A}(\mathbf{r}^N;\mathbf{r},\mathbf{r}',\ldots)$ carries further dependence on generic position variables $\mathbf{r},\mathbf{r}',\ldots$ then the mean $A(\mathbf{r},\mathbf{r}',\ldots)$ will be a spatially resolved (correlation) function.

In order to address the average (5) via the extended ensemble of section 2.1, we first revert to the standard mechanism of generating averages via parametric derivatives. In the present case we have

$$A = -\frac{\partial \beta\Omega}{\partial \lambda}, \tag{6}$$

where we recall the definition of the extended grand potential $\Omega = -k_\mathrm{B}T\ln\Xi$ with the extended partition sum $\Xi = \mathrm{Tr}\mathrm{e}^{-\beta(H-\mu N)+\lambda\hat{A}}$. Both the form of the external potential $V_\mathrm{ext}(\mathbf{r})$ as well as the state point μ, T are kept fixed when differentiating with respect to λ in equation (6). The validity of this relationship can be seen from standard parametric differentiation [5]: $A = \partial \ln\Xi/\partial\lambda = \Xi^{-1}\mathrm{Tr}\,\partial \mathrm{e}^{-\beta(H-\mu N)+\lambda\hat{A}}/\partial\lambda = \Xi^{-1}\mathrm{Tr}\mathrm{e}^{-\beta(H-\mu N)+\lambda\hat{A}}\hat{A} = \langle\hat{A}\rangle$.

In order to also systematically incorporate locally resolved fluctuations of $\hat{A}$ we construct a one-body fluctuation profile $\chi_A(\mathbf{r})$. The inspiration stems from the local compressibility [89–96] and the local thermal susceptibility [92, 94, 95], which respectively correlate the fluctuations of the local density with the total number of particles and the entropy.

In the present case of a general observable $\hat{A}$, the hyperfluctuation profile is obtained as the covariance of the considered observable with the position-dependent density operator (4) according to

$$\chi_A(\mathbf{r}) = \mathrm{cov}\left(\hat{\rho}(\mathbf{r}),\hat{A}\right) \tag{7}$$

$$= \langle \hat{\rho}(\mathbf{r})\hat{A}\rangle - \rho(\mathbf{r})A, \tag{8}$$

where equation (8) is the explicit form of the covariance in equation (7). The covariance of two general phase space functions $\hat{X}$ and $\hat{Y}$ is $\mathrm{cov}(\hat{X},\hat{Y}) = \langle\hat{X}\hat{Y}\rangle - \langle\hat{X}\rangle\langle\hat{Y}\rangle$. We recall that by construction for cases where the mean product factorizes according to $\langle\hat{X}\hat{Y}\rangle = \langle\hat{X}\rangle\langle\hat{Y}\rangle$, then by its very definition the covariance vanishes, $\mathrm{cov}(\hat{X},\hat{Y}) = 0$. Hence nonvanishing covariance of two observables is a measure of the degree of deviation from an idealized factorization behaviour. Using the generic definition of the covariance in equation (7) gives the hyperfluctuation profile $\chi_A(\mathbf{r})$ in the more explicit form (8), where the density profile $\rho(\mathbf{r})$ is given by equation (3) and the mean A is given by equation (5) and, equivalently, by equation (6).

While defining the mean A via equation (5) is entirely standard, the definition of the hyperfluctuation profile $\chi_A(\mathbf{r})$ via equation (7) introduces position-dependence by correlating the presence of a particle at a specific position $\mathbf{r}$ with the overall value of $\hat{A}(\mathbf{r}^N)$. The hyperfluctuation profile is hence not a merely 'local' version of the order parameter $\hat{A}(\mathbf{r}^N)$ in the sense that a space integral of the local version gives the global version. Rather the position integral gives $\int \mathrm{d}\mathbf{r}\chi_A(\mathbf{r}) = \int \mathrm{d}\mathbf{r}\mathrm{cov}(\hat{\rho}(\mathbf{r}),\hat{A}) = \mathrm{cov}(N,\hat{A})$, where the latter equality holds due to $\int \mathrm{d}\mathbf{r}\hat{\rho}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}\sum_i \delta(\mathbf{r}-\mathbf{r}_i) = \sum_i 1 = N$. Hence one obtains the (non-trivial) covariance of the value of $\hat{A}$ and the total number N of particles in the system, where we recall that the latter is a fluctuating quantity in the grand ensemble that we use as a foundation.

That the specific covariance form (7) of $\chi_A(\mathbf{r})$ is a physically meaningful measure of local fluctuations of $\hat{A}(\mathbf{r}^N)$ is suggested by several recent theoretical developments that independently point to the relevance of this specific type of correlation function. The arguably most prominent example is the local compressibility $\chi_\mu(\mathbf{r}) = \beta\mathrm{cov}(\hat{\rho}(\mathbf{r}),N)$, which is much advocated by Evans and coworkers as a highly useful indicator for drying phenomena that occur at a substrate [89–91, 93]. In the present general framework we recover $\chi_\mu(\mathbf{r})$ by the simple choice $\hat{A} = \beta N$ in equation (7). The local compressibility $\chi_\mu(\mathbf{r})$ attains further significance as the parametric derivative $\chi_\mu(\mathbf{r}) = \partial\rho(\mathbf{r})/\partial\mu$ [89–91, 93], which hence measures the changes in the equilibrium density profile upon changing the chemical potential (while keeping both the shape of the external potential and temperature constant).

The motivation to treat the dependence on T similarly to the dependence on μ led Eckert *et al* [94, 95] to correspondingly consider the thermal susceptibility $\chi_T(\mathbf{r}) = \partial\rho(\mathbf{r})/\partial T$, whereby μ and again the shape of the external potential are kept fixed upon differentiating. An equivalent covariance expression is $\chi_T(\mathbf{r}) = \mathrm{cov}(\hat{\rho}(\mathbf{r}),\hat{S})$, where the entropy operator is $\hat{S} = -k_\mathrm{B}\ln\Psi_\mathrm{eq}$ with the standard equilibrium probability distribution $\Psi_\mathrm{eq} = \mathrm{e}^{-\beta(H-\mu N)}/\Xi$; see [94, 95] for further considerations that make $\chi_T(\mathbf{r})$ accessible in simulations.

While both of the above similarities are based on concrete choices for $\hat{A}$, in the recent hyperforce theory by Robitschko *et al* [82] the explicit form (7) of $\chi_A(\mathbf{r})$ features in the formulation of exact Noether sum rules that emerge from thermal gauge invariance [83]. We lay out in detail the relationship of the present hyperdensity functional approach to the hyperforce theory below in section 3.

It is straightforward to show that the covariance form (7) of the hyperfluctuation profile $\chi_A(\mathbf{r})$ is generated in analogy to the mechanism used in [89–92, 94, 95] from parametrically differentiating the density profile (3). In the present case we have

$$\chi_A(\mathbf{r}) = \frac{\partial \rho(\mathbf{r})}{\partial \lambda}, \tag{9}$$

where μ, T and the form of the external potential $V_{\text{ext}}(\mathbf{r})$ are all kept fixed upon building the derivative. That equation (9) holds can be seen by explicit calculation of the parameter derivative of the density profile. To do so, we explicitly spell out the average (3) to obtain the density profile in the form $\rho(\mathbf{r}) = \text{Tr}\,\hat{\rho}(\mathbf{r})\text{e}^{-\beta(H-\mu N)+\lambda\hat{A}}/\Xi$. The dependence on λ occurs both directly in the extended Boltzmann factor $\text{e}^{-\beta(H-\mu N)+\lambda\hat{A}}$ as well as in the extended partition sum $\Xi = \text{Tr}\text{e}^{-\beta(H-\mu N)+\lambda\hat{A}}$. Upon differentiating, the product rule gives two terms that constitute the covariance (7). As before, we remain thereby interested in the case $\lambda \to 0$ after having taken the derivative in equation (9). Hence the local fluctuations of $\hat{A}$ are captured as they are generated from interparticle coupling that the original Hamiltonian H generates in the system.

A further important mechanism to generate the hyperfluctuation profile was identified by Eckert *et al* [95] in the context of their investigation into the local compressibility and thermal susceptibility. Upon investigating this specific case, they have identified the following general mechanism:

$$\chi_A(\mathbf{r}) = -\frac{\delta A}{\delta \beta V_{\text{ext}}(\mathbf{r})}. \tag{10}$$

Equation (10) is significant as it reveals $\chi_A(\mathbf{r})$ as the response function of the average A against changes in the (negative and thermally scaled) external potential.

Proving equation (10) is straightforward upon starting with $\chi_A(\mathbf{r})$ in the parametric derivative form (9) and interchanging the order of differentiation according to

$$\begin{aligned}\chi_A(\mathbf{r}) &= \frac{\partial \rho(\mathbf{r})}{\partial \lambda} = \frac{\partial}{\partial \lambda}\frac{\delta \Omega}{\delta V_{\text{ext}}(\mathbf{r})} = \frac{\delta}{\delta V_{\text{ext}}(\mathbf{r})}\frac{\partial \Omega}{\partial \lambda} \\ &= -\frac{\delta}{\delta \beta V_{\text{ext}}(\mathbf{r})}\left(-\frac{\partial \beta \Omega}{\partial \lambda}\right) = -\frac{\delta A}{\delta \beta V_{\text{ext}}(\mathbf{r})}.\end{aligned} \tag{11}$$

We have first re-written the density profile via the standard functional response formula [2, 5]

$$\rho(\mathbf{r}) = \frac{\delta \Omega}{\delta V_{\text{ext}}(\mathbf{r})}, \tag{12}$$

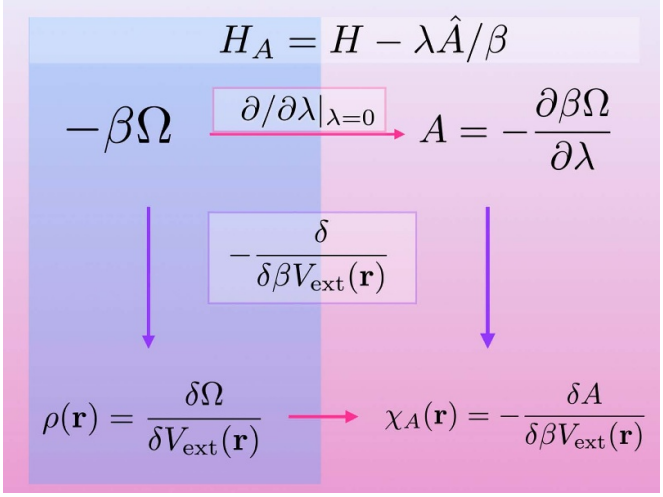

Figure 2. Illustration of parametric and functional relationships in the extended ensemble. The generalized Hamiltonian $H_A = H - \lambda\hat{A}/\beta$ is given by equation (2) and it has the associated scaled grand potential $-\beta\Omega$ and density profile $\rho(\mathbf{r})$. The mean A is obtained via parametric differentiation according to $\partial/\partial\lambda|_{\lambda=0}$. Similarly the hyperfluctuation profile $\chi_A(\mathbf{r})$ follows according to equation (10).

and then identified the average $A = -\partial\beta\Omega/\partial\lambda$ via equation (6). Note that the standard mechanism (12) to generate the density profile $\rho(\mathbf{r})$ bears strong similarities with the generation (10) of the hyperfluctuation profile $\chi_A(\mathbf{r})$ via functional differentiation. Figure 2 depicts an illustration of the mutual relationships.

2.3. Mermin-Evans minimization principle in the extended ensemble

We base the construction of the density functional formulation on the extended Hamiltonian (2) and hence consider the system as an interacting many-body system that is exposed to an unchanged external potential $V_{\text{ext}}(\mathbf{r})$ as it occurs in equation (1). Crucially, we attribute the differences between the two Hamiltonian H and H_A solely to a change in interparticle interactions. This implies going from the original interparticle interaction potential $u(\mathbf{r}^N)$ to the virtual interparticle interaction potential $u_A(\mathbf{r}^N)$ of the extended system, as it is given by

$$u_A(\mathbf{r}^N) = u(\mathbf{r}^N) - \lambda\hat{A}(\mathbf{r}^N)/\beta. \tag{13}$$

Using equation (13) to eliminate the explicit occurrence of $u(\mathbf{r}^N)$ in the extended Hamiltonian (2) then yields H_A in the standard form $H_A = \sum_i \mathbf{p}_i^2/(2m) + u_A(\mathbf{r}^N) + \sum_i V_{\text{ext}}(\mathbf{r}_i)$. The same result is obtained from simply replacing $u(\mathbf{r}^N)$ by $u_A(\mathbf{r}^N)$ in equation (1).

In typical systems $u(\mathbf{r}^N)$ can be of very specific form, say being composed of solely pair interaction contributions. Even if this is the case, then the form of $u_A(\mathbf{r}^N)$ can in general be

much more complex and in particular it can possess many-body contributions, depending on the specific form of the dependence of $\hat{A}(\mathbf{r}^N)$. Furthermore additional one-body contributions, on top of the bare external potential $V_{\text{ext}}(\mathbf{r})$, could be present. In principle one can split off such terms and combine them with $V_{\text{ext}}(\mathbf{r})$ into a modified external potential. We do not perform this modification though and both for simplicity and conceptual clarity keep $u_A(\mathbf{r}^N)$ in its full form (13). We hence rather exploit that the general density functional framework poses no restrictions on the form of the interparticle potential other than that the resulting thermal ensemble needs to be well-defined. We return to the cases of specific one- and two-body forms of the observable $\hat{A}(\mathbf{r}^N)$ in section 4 after the development of the general framework.

Although we remain ultimately interested only in the limit $\lambda \to 0$, such that the extended Hamiltonian (2) reduces to the Hamiltonian (1) of the original system, we start by formulating the grand potential density functional $\Omega[\rho]$ for the generalized Hamiltonian H_A. The general variational principle [1–3] ascertains that

$$\frac{\delta\Omega[\rho]}{\delta\rho(\mathbf{r})} = 0 \quad (\text{min}), \tag{14}$$

where equality holds at the minimum and the minimizer is the true equilibrium density distribution $\rho(\mathbf{r})$ given as the thermal average (3). In our present formulation $\rho(\mathbf{r})$ is the density profile of the extended system with generalized interparticle potential $u_A(\mathbf{r}^N)$ and in the presence of the fixed external potential $V_{\text{ext}}(\mathbf{r})$ and at the fixed thermodynamic state point μ, T.

The standard form of the grand potential density functional $\Omega[\rho]$ consists of a sum of ideal gas, excess (over ideal gas), and external contributions according to [2, 5]

$$\Omega[\rho] = F_{\text{id}}[\rho] + F_{\text{exc}}[\rho] + \int \text{d}\mathbf{r}\rho(\mathbf{r})\left[V_{\text{ext}}(\mathbf{r}) - \mu\right]. \tag{15}$$

The intrinsic Helmholtz free energy functional of the ideal gas, $F_{\text{id}}[\rho]$, is exactly known in the form $F_{\text{id}}[\rho] = k_{\text{B}}T\int \text{d}\mathbf{r}\rho(\mathbf{r})[\ln(\rho(\mathbf{r})\Lambda^d) - 1]$, where Λ denotes the thermal de Broglie wavelength. The excess free energy functional $F_{\text{exc}}[\rho]$ is unknown in general and due to all remaining interactions that are not accounted for by $V_{\text{ext}}(\mathbf{r})$. Hence in the present setup $F_{\text{exc}}[\rho]$ depends on and is generated by the extended interparticle interaction potential $u_A(\mathbf{r}^N)$, as is given by equation (13). In particular, the dependence of $u_A(\mathbf{r}^N)$ on the coupling parameter λ renders the excess free energy functional $F_{\text{exc}}[\rho]$ parametrically dependent on λ. In the limit $\lambda \to 0$ the excess free energy functional of the original system, with bare interparticle interaction potential $u(\mathbf{r}^N)$, is thereby restored by construction.

Inserting the grand potential density functional splitting (15) into the minimization principle (14) and calculating the functional derivative yields the Euler–Lagrange equation of classical density functional theory in the following standard form [2, 5, 6]:

$$c_1(\mathbf{r};[\rho]) = \ln\left[\rho(\mathbf{r})\Lambda^d\right] + \beta\left[V_{\text{ext}}(\mathbf{r}) - \mu\right]. \tag{16}$$

Thereby the one-body direct correlation functional $c_1(\mathbf{r};[\rho])$ is given as the density functional derivative of the scaled excess free energy functional $-\beta F_{\text{exc}}[\rho]$, such that

$$c_1(\mathbf{r};[\rho]) = -\frac{\delta\beta F_{\text{exc}}[\rho]}{\delta\rho(\mathbf{r})}. \tag{17}$$

Clearly the ensemble generalization is imprinted into $c_1(\mathbf{r};[\rho])$ through the generalized interparticle potential $u_A(\mathbf{r}^N)$ given via equation (13). As a result $c_1(\mathbf{r};[\rho])$ depends both on the value of the coupling parameter λ as well as on the specific form of the observable $\hat{A}(\mathbf{r}^N)$. Performing the limit $\lambda \to 0$ restores the one-body direct correlation functional of the original system, with bare interparticle interaction potential $u(\mathbf{r}^N)$.

Before we take the parametric limit though, we differentiate equation (17) with respect to λ and hence define the resulting hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ as

$$c_A(\mathbf{r};[\rho]) = \left.\frac{\partial c_1(\mathbf{r};[\rho])}{\partial\lambda}\right|_\rho, \tag{18}$$

where the density profile is kept fixed upon building the parametric derivative, which hence acts exclusively on the implied Hamiltonian H_A. In general $c_A(\mathbf{r};[\rho])$ will remain nonvanishing when the limit $\lambda \to 0$ is performed in equation (18) after the parametric derivative is taken. This procedure both restores the statistical mechanics of the original system, but as we will demonstrate in the following, it also allows one to capture the relevant statistical information of the behaviour of $\hat{A}(\mathbf{r}^N)$ that the original system displays via the one-body hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$.

The functional derivative relationship (17) can be inverted straightforwardly by functional integration. The general functional integral $-\beta F_{\text{exc}}[\rho] = \int \mathcal{D}[\rho]c_1(\mathbf{r};[\rho])$ can be parametrized efficiently according to [2, 3, 36]:

$$-\beta F_{\text{exc}}[\rho] = \int \text{d}\mathbf{r}\rho(\mathbf{r})\int_0^1 \text{d}a c_1(\mathbf{r};[a\rho]). \tag{19}$$

The functional argument $a\rho(\mathbf{r})$ of the one-body direct correlation functional is thereby merely a scaled version of the 'target' density profile $\rho(\mathbf{r})$ [2, 3]. Equation (19) is suitable for numerical evaluation upon providing a specific form of $c_1(\mathbf{r};[\rho])$ for the system under consideration [35–40]. Analogously, functional integration is relevant for evaluating the density functional $A[\rho]$ given a hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$, as described in more detail in section 2.6.

2.4. Levy's constrained search in the extended ensemble

The existence and uniqueness of the grand potential density functional (15) is typically proven by contradiction [2, 5].

Levy's constrained search method [8, 103] provides an arguably more constructive route and the method applies classically as well [6, 104]. Briefly, one starts from the standard Mermin-Evans many-body functional [1, 2] $\Omega_{\rm M}[\Psi] = {\rm Tr}\,\Psi\left(H_A - \mu N + k_{\rm B}T\ln\Psi\right)$ where $\Psi(\mathbf{r}^N,\mathbf{p}^N)$ is a many-body distribution function which is normalized according to ${\rm Tr}\,\Psi = 1$ but is otherwise general. Splitting off from $\Omega_{\rm M}[\Psi]$ the contributions from chemical potential and external potential leaves over an intrinsic many-body functional $F_{\rm M}[\Psi]$ given by

$$F_{\rm M}\left[\Psi\right] = {\rm Tr}\left(\sum_i \frac{\mathbf{p}_i^2}{2m} + u_A\left(\mathbf{r}^N\right) + k_{\rm B}T\ln\Psi\right)\Psi. \tag{20}$$

The Levy method generates from equation (20) the corresponding free energy density functional $F[\rho] = F_{\rm id}[\rho] + F_{\rm exc}[\rho]$ via the following constrained search:

$$F\left[\rho\right] = \min_{\Psi\to\rho} F_{\rm M}\left[\Psi\right]. \tag{21}$$

Here $\Psi(\mathbf{r}^N,\mathbf{p}^N)$ is constrained to generate the prescribed 'target' density profile $\rho(\mathbf{r})$, which is the functional argument on the left hand side of equation (21). The constraint, indicated as $\Psi\to\rho$ in equation (21), enforces the following identity

$$\rho\left(\mathbf{r}\right) = {\rm Tr}\,\hat{\rho}\left(\mathbf{r}\right)\Psi, \tag{22}$$

where the density operator $\hat{\rho}(\mathbf{r})$ is defined in equation (4).

It is interesting to note that the mean of the considered observable, for given form of $\Psi(\mathbf{r}^N,\mathbf{p}^N)$ is obtained as the following partial derivative:

$$A = -\frac{\partial F_{\rm M}\left[\Psi\right]}{\partial\lambda}\Big|_{\Psi,\mu,T}, \tag{23}$$

where the value A is the intended one provided that $\Psi(\mathbf{r}^N,\mathbf{p}^N)$ has the correct equilibrium form. That no further terms become relevant when parametrically differentiating is not immediately obvious and we postpone the construction of the density functional $A[\rho]$ to section 2.6 below.

2.5. Hyper-Ornstein–Zernike relation

While we have introduced microscopic expressions for the hyperfluctuation profile $\chi_A(\mathbf{r})$ both via the covariance (7) and, equivalently, by the parametric derivative (9), this particular correlation function emerges also very naturally within the present hyperdensity functional treatment based on a hyper-Ornstein–Zernike relation, as we will demonstrate in the following. We recall that the two-body Ornstein–Zernike relationship is a staple of liquid state theory and that its origin lies, together with the introduction of pair direct correlation functions, in the treatment of critical opalescence by Ornstein and Zernike in 1914 [5, 9].

The two-body direct correlation functional is given as the following second density functional derivative:

$$c_2\left(\mathbf{r},\mathbf{r}';[\rho]\right) = -\frac{\delta^2\beta F_{\rm exc}\left[\rho\right]}{\delta\rho\left(\mathbf{r}\right)\delta\rho\left(\mathbf{r}'\right)}. \tag{24}$$

Expressing one of the two chained functional derivatives in equation (24) via equation (17) leads to the following alternative and equivalent form:

$$c_2\left(\mathbf{r},\mathbf{r}';[\rho]\right) = \frac{\delta c_1\left(\mathbf{r};[\rho]\right)}{\delta\rho\left(\mathbf{r}'\right)}. \tag{25}$$

A complementary hierarchy of total correlation functions is obtained from functionally differentiating the elementary grand potential with respect to $V_{\rm ext}(\mathbf{r})$ [2, 5], such that $\delta^2\Omega/[\delta V_{\rm ext}(\mathbf{r})\delta V_{\rm ext}(\mathbf{r}')] = \beta{\rm cov}(\hat{\rho}(\mathbf{r}),\hat{\rho}(\mathbf{r}')) = \beta H_2(\mathbf{r},\mathbf{r}')$, where $H_2(\mathbf{r},\mathbf{r}')$ is the standard two-body correlation function of density fluctuations [2, 5]. The inhomogeneous Ornstein–Zernike equation then relates the total and direct correlation functions on the two-body level according to [2, 5, 6]:

$$H_2\left(\mathbf{r},\mathbf{r}'\right) = \rho\left(\mathbf{r}\right)\delta\left(\mathbf{r}-\mathbf{r}'\right) + \rho\left(\mathbf{r}\right)\int {\rm d}\mathbf{r}''c_2\left(\mathbf{r},\mathbf{r}'';[\rho]\right)H_2\left(\mathbf{r}'',\mathbf{r}'\right). \tag{26}$$

To derive the hyper-Ornstein–Zernike relation, we follow a formal and very direct strategy based on the Euler–Lagrange equation [6, 94, 95, 105, 106]. The method is based on exploiting the generality of the minimization condition (16) and the concept that this remains a valid identity upon suitable differentiation. Carrying out functional derivatives thereby requires to take account of both direct changes as well as mediated changes that are generated via changes in the density profile itself. We refer the Reader to [6] for the corresponding derivation of the two-body Ornstein–Zernike relation (26) from functionally differentiating the standard Euler–Lagrange equation (16) with respect to $V_{\rm ext}(\mathbf{r}')$.

Here we turn to the extended ensemble and differentiate the Euler–Lagrange equation (16) with respect to the coupling parameter λ. We consider both the one-body direct correlation functional $c_1(\mathbf{r};[\rho])$ and the equilibrium density profile $\rho(\mathbf{r})$ to be those corresponding to the extended Hamiltonian H_A. The external potential $V_{\rm ext}(\mathbf{r})$ and the state point μ,T remain prescribed. Hence the Euler–Lagrange equation (16) is satisfied for a range of values of λ. We hence retain a valid identity upon differentiating equation (16) with respect to λ. Respecting the involved dependencies yields the following hyper-Ornstein–Zernike equation:

$$c_A\left(\mathbf{r};[\rho]\right) = \frac{\chi_A\left(\mathbf{r}\right)}{\rho\left(\mathbf{r}\right)} - \int {\rm d}\mathbf{r}'c_2\left(\mathbf{r},\mathbf{r}';[\rho]\right)\chi_A\left(\mathbf{r}'\right), \tag{27}$$

where the one-body hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ remains being defined via the parametric derivative (18).

We provide a detailed derivation of equation (27) in the following. We first consider the left hand side of the Euler–Lagrange equation (16). Differentiating the occurring one-body direct correlation functional $c_1(\mathbf{r};[\rho])$ with respect to λ yields

$$\frac{\partial c_1(\mathbf{r};[\rho])}{\partial\lambda}\Big|_{V_{\rm ext}} = \frac{\partial c_1(\mathbf{r};[\rho])}{\partial\lambda}\Big|_{\rho} + \int \mathrm{d}\mathbf{r}' \frac{\delta c_1(\mathbf{r};[\rho])}{\delta\rho(\mathbf{r}')}\frac{\partial\rho(\mathbf{r}')}{\partial\lambda}\Big|_{V_{\rm ext}} \tag{28}$$

$$= c_A(\mathbf{r};[\rho]) + \int \mathrm{d}\mathbf{r}' c_2(\mathbf{r},\mathbf{r}';[\rho])\,\chi_A(\mathbf{r}')\,. \tag{29}$$

The two terms in equation (28) arise from the direct changes upon changing λ (first term) and from the changes that are mediated by alteration of the density profile and using the functional chain rule (second term). To obtain equation (29) we have identified the two parametric derivatives as the one-body hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ via equation (18) and the hyperfluctuation profile $\chi_A(\mathbf{r})$ via equation (9).

It remains to also differentiate the right hand side of the Euler–Lagrange equation (16) with respect to λ. We recall that the state point μ, T and the shape of the external potential $V_{\rm ext}(\mathbf{r})$ are kept fixed upon changing the parameter λ. Hence the corresponding terms vanish upon differentiating and the result is compact:

$$\frac{\partial\ln\left[\rho(\mathbf{r})\Lambda^d\right]}{\partial\lambda} + \frac{\partial\left[\beta V_{\rm ext}(\mathbf{r})-\beta\mu\right]}{\partial\lambda} = \frac{1}{\rho(\mathbf{r})}\frac{\partial\rho(\mathbf{r})}{\partial\lambda} = \frac{\chi_A(\mathbf{r})}{\rho(\mathbf{r})}, \tag{30}$$

where in the last step we have again identified $\chi_A(\mathbf{r})$ via equation (9). The exact sum rule then follows from equating the right hand sides of equations (29) and (30), which yields the hyper-Ornstein–Zernike equation in the form (27) after rearranging.

The hyper-Ornstein–Zernike equation (27) is exact and in a similar fashion as the inhomogeneous two-body Ornstein–Zernike equation of liquid state theory [5] it can assume a multitude of different roles. That equation (27) only depends on a single free position argument, rather than two space points, as are occurring in equation (26), is in keeping with the fluctuation Ornstein–Zernike relations for $\chi_\mu(\mathbf{r})$ and for $\chi_{\rm T}(\mathbf{r})$ [94, 95]. Furthermore, similarly to equation (27) these exact relationships also feature the mediation of fluctuations via spatial integration of the standard two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$; we refer the Reader to [95] for further details.

In the following we describe two application scenarios of the hyper-Ornstein–Zernike relation (27) with a specific choice of the observable $\hat{A}(\mathbf{r}^N)$ being under consideration. We assume that the density profile $\rho(\mathbf{r})$ for the bare Hamiltonian H has been obtained from either solution of the Euler–Lagrange equation (16) or from the equilibrium average (3) as carried out e.g. in grand canonical Monte Carlo simulations. In both cases the form of $V_{\rm ext}(\mathbf{r})$ has been given and the thermodynamic state point μ, T has been prescribed. The situation is hence similar to application of the standard inhomogeneous two-body Ornstein–Zernike equation (26) where knowledge of the density profile is typically also required. Furthermore we assume that the two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$ is available. This can come either from the standard second functional derivative (24) of a (typically approximate) analytical model form of the excess free energy functional $F_{\rm exc}[\rho]$. Alternatively, the neural functional calculus of [35–37] provides direct representation of $c_1(\mathbf{r};[\rho])$ as a trained neural network and performing automatic differentiation [35–37, 71] then creates the neural two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$ via the first functional derivative (25).

We first assume that both the hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ and the two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$ are known. The particles interact thereby solely via the bare interaction potential $u(\mathbf{r}^N)$ and hence $\lambda\to 0$ has been taken. Then evaluating $c_A(\mathbf{r};[\rho])$ at the equilibrium density $\rho(\mathbf{r})$ turns the left hand side of the hyper-Ornstein–Zernike equation (27) into a fixed spatial inhomogeneity $c_A(\mathbf{r})$. The remaining terms on the right hand side need to accommodate this inhomogeneity and hence equation (27) constitutes an integral equation for the hyperfluctuation profile $\chi_A(\mathbf{r})$. Here $c_2(\mathbf{r},\mathbf{r}';[\rho])$, when evaluated at the equilibrium density profile, yields a generalized, i.e. fully position-dependent, convolution kernel $c_2(\mathbf{r},\mathbf{r}')$. Besides the (standard) input of the density profile and $c_2(\mathbf{r},\mathbf{r}';[\rho])$, this requires mere availability of the hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$, which points to the prowess of this object and we recall its definition via the parametric derivative (18) of the extended one-body direct correlation functional. See [39] for the application of this route to the calculation of the local compressibility.

The second use of the hyper-Ornstein–Zernike relation (27) targets the construction of $c_A(\mathbf{r};[\rho])$. The method is based on simulations, where the hyperfluctuation profile $\chi_A(\mathbf{r})$ is available via sampling the covariance (7). We assume that this has been accomplished together with sampling of $\rho(\mathbf{r})$, again for given $V_{\rm ext}(\mathbf{r}), \mu, T$. We also assume that the two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$ is known by one of the two methods described above (or alternatively by inhomogeneous liquid integral equation theory). Then all terms on the right hand side of equation (27) can be evaluated for the specific system under consideration, as specified by its Hamiltonian H, including the form of the external potential, and the thermodynamic parameters μ, T. For the specific chosen observable $\hat{A}(\mathbf{r}^N)$ the right hand side of equation (27) can hence be evaluated and the one-body hyperdirect correlation function $c_A(\mathbf{r})$ for the specific system under consideration has become available. While knowing the function in a specific case does not yet imply knowledge of the functional $c_A(\mathbf{r};[\rho])$, creating a set of such profiles from simulating the system under a range of different conditions is the basis for using supervised machine learning following [35–39] in the construction of neural hyperdirect correlation functionals $c_A(\mathbf{r};[\rho])$ [40].

Details and concrete applications of this technique are given in section 5.

Having demonstrated the intimate links between the hyper-Ornstein–Zernike equation (27), the hyperdirect correlation functional (18) and the hyperfluctuation profile (7), we next turn to addressing the mean A of the chosen observable, given via the elementary form (5), in the density functional context.

2.6. Any observable as a hyperdensity functional

We aim at a density functional representation of the average A. We start from the parametric differentiation according to equation (6). However, rather than using the elementary form of the grand potential, $\Omega = -k_{\mathrm{B}}T\ln\Xi$, we work on the basis of the grand potential functional $\Omega[\rho]$. The splitting of $\Omega[\rho]$ into its additive constituents via equation (15) enables one to carry out the required parametric derivative with respect to λ under identical conditions as in equation (6), i.e. at fixed form of $V_{\mathrm{ext}}(\mathbf{r})$ and fixed state point μ, T. The result (derived below) is compact:

$$A\left[\rho\right] = -\frac{\partial \beta F_{\mathrm{exc}}\left[\rho\right]}{\partial\lambda}\Big|_{\rho}, \tag{31}$$

where $F_{\mathrm{exc}}[\rho]$ is the excess free energy functional generated by the extended Hamiltonian H_A and the density profile $\rho(\mathbf{r})$ is kept fixed upon differentiating. As before we can take the limit $\lambda \to 0$ after the derivative is taken, in order to restore the statistical mechanics of the original Hamiltonian H. No other terms in equation (15) depend explicitly on λ, which leads to the simplicity of the right hand side of equation (31).

As equation (31) is a central and arguably counter-intuitive result, we provide more elementary details for its derivation. Taking into account the definition of A via equation (6), we start from the grand potential functional $\Omega[\rho]$, as expressed via the splitting (15), and differentiate with respect to the coupling parameter λ as follows:

$$\frac{\partial\Omega\left[\rho\right]}{\partial\lambda}\Big|_{V_{\mathrm{ext}}} = \frac{\partial\Omega\left[\rho\right]}{\partial\lambda}\Big|_{\rho} + \int \mathrm{d}\mathbf{r}\frac{\delta\Omega\left[\rho\right]}{\delta\rho\left(\mathbf{r}\right)}\Big|_{V_{\mathrm{ext}}}\chi_A\left(\mathbf{r}\right). \tag{32}$$

The first term on the right hand side of equation (32) accounts for the 'inherent' changes in $\Omega[\rho]$ that occur upon keeping the density profile fixed, as is indicated in the notation. This term can be pictured as λ controlling the extended Hamiltonian H_A, see equation (2), upon keeping the density profile fixed when changing H_A. By splitting $\Omega[\rho]$ according to equation (15) we can simplify as follows:

$$\begin{aligned}\frac{\partial\Omega\left[\rho\right]}{\partial\lambda}\Big|_{\rho} &= \frac{\partial F_{\mathrm{exc}}\left[\rho\right]}{\partial\lambda}\Big|_{\rho} \\ &\quad + \frac{\partial}{\partial\lambda}\Big|_{\rho}\left(F_{\mathrm{id}}\left[\rho\right] + \int \mathrm{d}\mathbf{r}\rho\left(\mathbf{r}\right)\left[V_{\mathrm{ext}}\left(\mathbf{r}\right) - \mu\right]\right) \qquad (33)\\ &= \frac{\partial F_{\mathrm{exc}}\left[\rho\right]}{\partial\lambda}\Big|_{\rho}, \qquad (34)\end{aligned}$$

where the ideal, external, and chemical contributions in equation (33) carry no dependence on λ and hence vanish under the parametric derivative, which leads to equation (34).

The second contribution on the right hand side of equation (32) arises from the chain rule and identifying the hyperfluctuation profile $\chi_A(\mathbf{r})$ via the definition (9) as the partial derivative $\partial\rho(\mathbf{r})/\partial\lambda$. However, it follows straightforwardly that

$$\int \mathrm{d}\mathbf{r}\frac{\delta\Omega\left[\rho\right]}{\delta\rho\left(\mathbf{r}\right)}\Big|_{V_{\mathrm{ext}}}\chi_A\left(\mathbf{r}\right) = 0. \tag{35}$$

This simplification is due to the first term inside of the integrand vanishing: $\delta\Omega[\rho]/\delta\rho(\mathbf{r})|_{V_{\mathrm{ext}}} = 0$, where, as indicated, the derivative is taken at fixed $V_{\mathrm{ext}}(\mathbf{r})$ and at constant value of the coupling parameter λ and fixed state point μ, T. Due to the fundamental variational principle (14) in the extended ensemble the result vanishes identically. Collecting the results (34) and (35) leaves over only the compact right hand side of equation (31).

Certainly the simplicity of equation (31) points towards reaffirming the central role of the excess free energy functional within the density functional framework as encapsulating the essence of the interparticle coupling. The extended ensemble thereby facilitates to both retain this encapsulation, but also to extend towards a general observable $\hat{A}$.

We can now establish a connection between $A[\rho]$ and $c_A(\mathbf{r};[\rho])$. Inserting equation (17) into equation (18), exchanging the orders of the functional density derivative and the parametric derivative with respect to λ, and identifying the density functional $A[\rho]$ via equation (31) yields:

$$c_A\left(\mathbf{r};\left[\rho\right]\right) = \frac{\delta A\left[\rho\right]}{\delta\rho\left(\mathbf{r}\right)}. \tag{36}$$

Equation (36) adds further significance to the hyperdirect correlation functional as reflecting the changes in the mean A upon changing the density profile. As laid out and explicitly used in the derivation, the changes in the density profile are monitored in the system at fixed thermodynamic conditions and the particles interact solely via the original interaction potential $u(\mathbf{r}^N)$.

As we have alluded to above, the practical access to $c_A(\mathbf{r};[\rho])$ via supervised machine learning is arguably more direct than attempting to construct $A[\rho]$ from first principles. Once the one-body hyperdirect correlation functional is known, one can straightforwardly carry out a functional integral, which we write first in a formal way in the form

$$A\left[\rho\right] = \int \mathcal{D}\left[\rho\right]c_A\left(\mathbf{r};\left[\rho\right]\right). \tag{37}$$

The functional integral operator $\int\mathcal{D}[\rho]$ performs a line integral in the space of density functions; we refer the Reader to the classical presentations by Evans [2, 3] and to [36] for the implications in the light of the neural functional theory. In practice carrying out the formal functional integral (37)

requires to choose a parametrizaion, e.g. according to a simple scaling:

$$A[\rho] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r}) \int_0^1 \mathrm{d}a c_A(\mathbf{r};[a\rho]). \tag{38}$$

Here the functional argument $a\rho(\mathbf{r})$, with parameter $0 \leqslant a \leqslant 1$, of the hyperdirect correlation functional is a scaled version of the 'target' density profile $\rho(\mathbf{r})$ that is the argument on the left hand side of equation (38). Equation (38) thereby mirrors closely the functional integration of $c_1(\mathbf{r};[\rho])$ to obtain the excess free energy via equation (19). When working with neural functionals, evaluating equation (38) numerically is a very fast operation, similar in performance to the corresponding functional integral (19) over $c_1(\mathbf{r};[\rho])$ [35–39, 43, 44].

2.7. Wall hypercontact theorem

As a specific situation, we consider a semi-infinite system, where a hard wall at $x = 0$ constrains all particle coordinates $\mathbf{r}_i$ such that all $x_i > 0$. For large distances from the wall, $x \to \infty$, the system approaches bulk with vanishing external potential. The bulk pressure p then characterizes the system and it is related to the density $\rho(0^+)$ at the wall by the contact theorem [5, 79],

$$\rho(0^+) = \beta p. \tag{39}$$

In generalization of the procedure by Evans and Stewart of changing μ [89], as followed up by Eckert *et al* who considered changing T [95], we here parametrically differentiate the hard wall contact theorem (39) with respect to the coupling parameter λ. The result is

$$\chi_A(0^+) = \frac{A_b}{V}, \tag{40}$$

where $\chi_A(0^+)$ is the contact value of the fluctuation profile $\chi_A(\mathbf{r})$, the bulk expectation value of $\hat{A}(\mathbf{r}^N)$ is indicated by A_b, and V is the bulk volume. The left hand side of the contact theorem (40) is obtained from using the parametric derivative form (9) of the hyperfluctuation profile. The right hand side of equation (40) follows from noting that in bulk $\Omega = -pV$ and expressing the mean A via the parametric derivative (6) of the grand potential, which gives $A_b = -\partial\beta\Omega/\partial\lambda = V\partial\beta p/\partial\lambda = V\partial\rho(0^+)/\partial\lambda = V\chi_A(0^+)$.

This concludes our presentation of the essentials of the hyperdensity functional theory of [40]. We proceed to demonstrating the close relationship with the hyperforce framework by Robitschko *et al* [82]. This theory has rather different roots, as it arises from the application of Noether's theorem [75–83] to the gauge invariance of statistical mechanics [83]. Similarly to the present hyperdensity functional theory, however, this approach also enables one to independently choose the Hamiltonian H and the observable $\hat{A}$ under investigation.

3. Hyperforce correlations

3.1. Local hyperforce balance

The hyperdensity functional theory described in section 2 is based on the standard density functional theory concepts of representing the correlated many-body physics in terms of generating functionals. Examples of their use are the generation of the direct correlation functionals from the excess free energy functional $F_{\mathrm{exc}}[\rho]$ via equation (17) and the locally resolved equilibrium balance condition in the form of the Euler–Lagrange equation (16) that expresses the spatial homogeneity of the sum of all contributions to the chemical potential. Although the underlying Hamiltonian generates forces and the classical mechanics crucially rests on the concept of forces, typically these do not feature prominently in accounts of density functional theory, although there are notable exceptions [6, 79, 102].

Here we aim to introduce forces into the hyperdensity functional framework and thereby work on the basis of the extended Hamiltonian H_A together with the standard grand canonical setting at chemical potential μ and temperature T, as described in section 2.1. The extended Hamiltonian H_A contains the extended interparticle interaction potential given by equation (13), which we reproduce as $u_A(\mathbf{r}^N) = u(\mathbf{r}^N) - \lambda\hat{A}(\mathbf{r}^N)/\beta$. The corresponding extended one-body interparticle force density operator is then defined as

$$\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r}) = -\sum_i \delta(\mathbf{r} - \mathbf{r}_i)\nabla_i u_A(\mathbf{r}^N), \tag{41}$$

where the spatial localization is provided by the Dirac distribution, similarly to the mechanism in the density operator (4). The mean equilibrium force density balance then attains the standard form [6]:

$$\beta\mathbf{F}_{\mathrm{int}}(\mathbf{r}) = \nabla\rho(\mathbf{r}) + \rho(\mathbf{r})\nabla\beta V_{\mathrm{ext}}(\mathbf{r}). \tag{42}$$

The mean one-body interparticle force density distribution is the average

$$\mathbf{F}_{\mathrm{int}}(\mathbf{r}) = \langle\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r})\rangle, \tag{43}$$

with $\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r})$ given by equation (41) and the density profile $\rho(\mathbf{r}) = \langle\hat{\rho}(\mathbf{r})\rangle$ according to equation (3), with all averages being those of the extended ensemble.

By construction the extended force density balance (42) holds for any value of λ. We proceed very similarly to our above strategy in section 2.5, where we differentiated the Euler–Lagrange equation (16) for the extended system with respect to λ to derive the hyper-Ornstein–Zernike relation (27).

Here we parametrically differentiate the force density balance (42) with respect to λ, which yields a non-trivial identity as we will demonstrate. We proceed stepwise and first address

the interparticle force density operator (41). Taking account of the form (13) of $u_A(\mathbf{r}^N)$ yields the result

$$\frac{\partial\beta\hat{\mathbf{F}}_{\rm int}(\mathbf{r})}{\partial\lambda} = \hat{\mathbf{S}}_A(\mathbf{r}), \tag{44}$$

where we follow Müller *et al* [83, 84] in defining the hyperforce density 'operator' (phase space function) as

$$\hat{\mathbf{S}}_A(\mathbf{r}) = \sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla_i \hat{A}(\mathbf{r}^N). \tag{45}$$

We next differentiate the entire average on the left hand side of equation (42), which gives via the product rule:

$$\frac{\partial\beta\mathbf{F}_{\rm int}(\mathbf{r})}{\partial\lambda} = {\rm cov}\left(\beta\hat{\mathbf{F}}_{\rm int}(\mathbf{r}),\hat{\rho}(\mathbf{r})\right) + \mathbf{S}_A(\mathbf{r}), \tag{46}$$

where the mean hyperforce density is $\mathbf{S}_A(\mathbf{r}) = \langle\hat{\mathbf{S}}_A(\mathbf{r})\rangle$ with $\hat{\mathbf{S}}_A(\mathbf{r})$ given by equation (45). The covariance on the right hand side of equation (46) arises from differentiating the mean $\langle\hat{\mathbf{F}}_{\rm int}(\mathbf{r})\rangle$ upon keeping the extended force density operator (41) itself fixed. The second term on the right hand side of equation (46), i.e. the hyperforce density $\mathbf{S}_A(\mathbf{r})$, stems from differentiating the operator $\hat{\mathbf{F}}_{\rm int}(\mathbf{r})$, according to equation (44), inside of the average.

It remains to also parametrically differentiate the two terms on the right hand side of equation (42), which respectively gives

$$\frac{\partial}{\partial\lambda}\nabla\rho(\mathbf{r}) = \nabla\chi_A(\mathbf{r}), \tag{47}$$

$$\frac{\partial}{\partial\lambda}\rho(\mathbf{r})\nabla\beta V_{\rm ext}(\mathbf{r}) = \chi_A(\mathbf{r})\nabla\beta V_{\rm ext}(\mathbf{r}). \tag{48}$$

Carrying out the derivatives on the left hand sides of equations (47) and (48) is performed as before with both β and $V_{\rm ext}(\mathbf{r})$ being kept fixed. In both cases the parametric derivative of $\rho(\mathbf{r})$ with respect to λ then generates the hyperforce fluctuation profile $\chi_A(\mathbf{r})$ via equation (9).

In summary, using the results (46)–(48) allows one to express the derivative with respect to λ of the extended force density balance (42) as

$$\mathbf{S}_A(\mathbf{r}) + {\rm cov}\left(\beta\hat{\mathbf{F}}_{\rm int}(\mathbf{r}),\hat{A}(\mathbf{r}^N)\right) - \nabla\chi_A(\mathbf{r}) - \chi_A(\mathbf{r})\nabla\beta V_{\rm ext}(\mathbf{r}) = 0. \tag{49}$$

Taking the limit $\lambda\to 0$ retains the form of equation (49) and reduces the extended force density operator therein to that of the original system, $\hat{\mathbf{F}}_{\rm int}(\mathbf{r}) = -\sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla_i u(\mathbf{r}^N)$. Hence equation (49) applies to general forms of interparticle interaction potentials $u(\mathbf{r}^N)$ and independently chosen forms of the observable $\hat{A}(\mathbf{r}^N)$.

The formally exact sum rule (49) constitutes the locally resolved hyperforce density balance obtained previously by Robitschko *et al* [82]. Their derivation of equation (49) rests on the both conceptually and practically very different method of exploiting thermal Noether invariance of the average A in the standard ensemble. Alternatively, the hyperforce density balance (49) can also be derived from suitable ad hoc partial integration procedures on phase space according to the Yvon theorem [82]. Its arguably most striking, as well as compact, derivation is that from the phase space operator methods proposed by Müller *et al* [83] based on the recently discovered statistical mechanical gauge invariance of microstates [83, 84].

By collecting the different force contributions one can put equation (49) into more compact form as

$$\mathbf{S}_A(\mathbf{r}) + {\rm cov}\left(\beta\hat{\mathbf{F}}(\mathbf{r}),\hat{A}(\mathbf{r}^N)\right) = 0. \tag{50}$$

Here we have introduced the total one-body force density operator $\hat{\mathbf{F}}(\mathbf{r})$ in the simple momentum-independent form

$$\hat{\mathbf{F}}(\mathbf{r}) = \hat{\mathbf{F}}_{\rm int}(\mathbf{r}) - k_{\rm B}T\nabla\hat{\rho}(\mathbf{r}) - \hat{\rho}(\mathbf{r})\nabla V_{\rm ext}(\mathbf{r}), \tag{51}$$

which allows upon taking account of the covariance form (7) of $\chi_A(\mathbf{r})$ to rewrite equation (49) as equation (50) [82]. As Robitschko *et al* [82] argue, the covariance on the left hand side of equation (50) can also be written as merely the mean $\langle\beta\hat{\mathbf{F}}(\mathbf{r})\hat{A}(\mathbf{r}^N)\rangle$ because the factorized contribution vanishes, $\langle\hat{\mathbf{F}}(\mathbf{r})\rangle\langle\hat{A}(\mathbf{r}^N)\rangle = 0$, due to $\langle\hat{\mathbf{F}}(\mathbf{r})\rangle = \mathbf{F}_{\rm int}(\mathbf{r}) - k_{\rm B}T\nabla\rho(\mathbf{r}) - \rho(\mathbf{r})\nabla\beta V_{\rm ext}(\mathbf{r}) = 0$ according to the equilibrium force density balance (42).

That the hyperforce density balance (50) is obtained from the present conceptually quite simple route of parametrically differentiating the force density balance (42) in the extended ensemble is a further indication of the fundamental status of this relation (see also the arguments given in [82]). We take the apparent tight interconnection as a demonstration of the capacity of the generalized ensemble to yield non-trivial insight. As anticipated in our discussion of the hyperfluctuation profile $\chi_A(\mathbf{r})$ in its covariance form (7), this local measure of the coupling of fluctuations in density and in the observable $\hat{A}(\mathbf{r}^N)$ features prominently in the hyperforce density balance (49) both via the diffusion-like first term and the external force-like second term on the left hand side of equation (49).

3.2. Hyperforces and hyperdensity functionals

The conceptual bridge between the present hyperforce concepts and the density functional point of view of section 2 is provided by linking the force density balance (42) with the Euler–Lagrange equation (16). In the standard way [6] we hence build the gradient of the latter identity, which yields upon multiplication by $\beta\rho(\mathbf{r})$ the result

$$\rho(\mathbf{r})\nabla c_1(\mathbf{r};[\rho]) = \nabla\rho(\mathbf{r}) + \rho(\mathbf{r})\nabla\beta V_{\rm ext}(\mathbf{r}), \tag{52}$$

where we have simplified $\rho(\mathbf{r})\nabla\ln\rho(\mathbf{r}) = \nabla\rho(\mathbf{r})$. Then comparing equation (52) and the force density balance (42) allows one, on the basis of the two identical right hand sides, to identify:

$$\beta\mathbf{F}_{\rm int}(\mathbf{r}) = \rho(\mathbf{r})\nabla c_1(\mathbf{r};[\rho]). \tag{53}$$

One significant consequence of equation (53) is the implication that its left hand side is elevated from a mere average $\mathbf{F}_{\mathrm{int}}(\mathbf{r}) = \langle \hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r}) \rangle$ given by equation (43) to also being a density functional, $\mathbf{F}_{\mathrm{int}}(\mathbf{r};[\rho])$, as given via the right hand side of equation (53).

We recall that we still work in the extended ensemble and that hence both sides of equation (53) are understood in this way. In order to make progress we parametrically differentiate equation (53) with respect to λ. The result is the following identity:

$$\begin{aligned}&\mathrm{cov}\left(\beta\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r}), \hat{A}\left(\mathbf{r}^N\right)\right) + \mathbf{S}_A(\mathbf{r}) \\ &\quad = \chi_A(\mathbf{r})\nabla c_1(\mathbf{r};[\rho]) + \rho(\mathbf{r})\nabla c_A(\mathbf{r};[\rho]) \\ &\qquad + \rho(\mathbf{r})\int \mathrm{d}\mathbf{r}'\chi_A(\mathbf{r}')\nabla c_2(\mathbf{r},\mathbf{r}';[\rho]). \end{aligned} \tag{54}$$

The left hand side of equation (54) follows directly from equation (46). The three terms on the right hand side of equation (54) follow respectively from the right hand side of equation (53) via: (i) differentiating the prefactor $\rho(\mathbf{r})$ and using equation (9), (ii) differentiating $c_1(\mathbf{r};[\rho])$ parametrically upon fixing the density profile according to equation (18), and (iii) using the chain rule upon monitoring the changes in the functional argument $\rho(\mathbf{r})$ and using equation (9) to identify the hyperfluctuation profile $\chi_A(\mathbf{r})$.

We can now restore the original Hamiltonian in equation (54) by taking the limit $\lambda \to 0$, which in particular again sets the interparticle force density operator to that of the original interparticle interaction potential, $\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r}) = -\sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla_i u(\mathbf{r}^N)$. As a consistency check, we replace the left hand side of equation (54) on the basis of the hyperforce density balance (49) and regroup terms, which yields

$$\begin{aligned}&\nabla\chi_A(\mathbf{r}) + \chi_A(\mathbf{r})\nabla\beta V_{\mathrm{ext}}(\mathbf{r}) - \chi_A(\mathbf{r})\nabla c_1(\mathbf{r};[\rho]) \\ &\quad = \rho(\mathbf{r})\nabla c_A(\mathbf{r};[\rho]) + \rho(\mathbf{r})\int \mathrm{d}\mathbf{r}'\chi_A(\mathbf{r}')\nabla c_2(\mathbf{r},\mathbf{r}';[\rho]). \end{aligned} \tag{55}$$

Replacing the one-body direct correlation functional by using the Euler–Lagrange equation (16) allows one to simplify the second term on the left hand side. Upon dividing by the density profile the result is

$$\begin{aligned}&\frac{\nabla\chi_A(\mathbf{r})}{\rho(\mathbf{r})} + \chi_A(\mathbf{r})\frac{\nabla\ln\rho(\mathbf{r})}{\rho(\mathbf{r})} \\ &\quad = \nabla c_A(\mathbf{r};[\rho]) + \int \mathrm{d}\mathbf{r}'\chi_A(\mathbf{r}')\nabla c_2(\mathbf{r},\mathbf{r}';[\rho]), \end{aligned} \tag{56}$$

which is identical to the gradient of the hyper-Ornstein–Zernike relation (27).

In conclusion of this section, we find a tight integration of the recently investigated hyperforce correlations into the hyperdensity framework. This situation is structurally similar to the relationship between averages of bare forces in elementary statistical physics and in standard density functional theory, see [6] for an extended discussion of this force point of view.

4. Specific forms of observables

4.1. One-body observables

The considerations presented thus far have been general in the sense that no restrictions on the specific form of the observable $\hat{A}(\mathbf{r}^N)$ under consideration were applied other than that the resulting extended ensemble is well-defined or, analogously, that H_A defined via equation (2) can serve as a valid Hamiltonian and that indeed $H_A \to H$ in the limit $\lambda \to 0$.

Here we follow the SI of [40] and address specific types of $\hat{A}(\mathbf{r}^N)$ with the aim to classify the behaviour that arises according to different form of the dependence on the particle configuration $\mathbf{r}^N$. We first consider one-body forms

$$\hat{A}\left(\mathbf{r}^N\right) = \sum_i a_1(\mathbf{r}_i), \tag{57}$$

where $a_1(\mathbf{r})$ is a given function of position $\mathbf{r}$. Recalling the form of the density operator (4) as a sum of Dirac distributions, we can rewrite equation (57) as

$$\hat{A}\left(\mathbf{r}^N\right) = \int \mathrm{d}\mathbf{r}\hat{\rho}(\mathbf{r})a_1(\mathbf{r}). \tag{58}$$

Averaging equation (58) on both sides directly gives the thermal mean A as a density functional

$$A[\rho] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})a_1(\mathbf{r}), \tag{59}$$

where we have exchanged on the right hand side the order of the average and spatial integral and have identified $\rho(\mathbf{r})$ as the average (3). From equation (59) it is clear that a standard density functional treatment suffices to obtain all relevant information as within density functional theory the density profile is of course available.

For this simple one-body case we obtain the corresponding hyperdirect correlation functional via insertion of equation (59) into equation (36), which yields

$$c_A(\mathbf{r};[\rho]) = a_1(\mathbf{r}), \tag{60}$$

such that there is no density functional dependence. The simple result (60) can serve as a toy to illustrate more general hyperdirect correlations.

The corresponding hyperfluctuation profile $\chi_A(\mathbf{r})$ is according to the covariance form (7) obtained as

$$\chi_A(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' a_1(\mathbf{r}')\,\mathrm{cov}\left(\hat{\rho}(\mathbf{r}), \hat{\rho}(\mathbf{r}')\right) \tag{61}$$

$$= \int \mathrm{d}\mathbf{r}' a_1(\mathbf{r}')H_2(\mathbf{r},\mathbf{r}'), \tag{62}$$

where $H_2(\mathbf{r},\mathbf{r}') = \mathrm{cov}(\hat{\rho}(\mathbf{r}),\hat{\rho}(\mathbf{r}'))$ is the standard correlation function of density fluctuations [2, 5]; see our discussion in the context of the standard inhomogeneous Ornstein–Zernike equation (26). Inserting the results (60) and (62) into the general hyper-Ornstein–Zernike equation (27) then yields this relation upon dividing by β in the following one-body form

$$\int \mathrm{d}\mathbf{r}'' a_1(\mathbf{r}'') \Big[\rho(\mathbf{r}) \int \mathrm{d}\mathbf{r}' c_2(\mathbf{r},\mathbf{r}') H_2(\mathbf{r}',\mathbf{r}'') + \rho(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}'') - H_2(\mathbf{r},\mathbf{r}'')\Big] = 0. \tag{63}$$

The identity (63) is straightforward to prove directly as the term in brackets already vanishes due to the standard two-body Ornstein–Zernike relationship (26) [5, 6].

In alternative reasoning we note that equation (63) holds for any permissible form of $a_1(\mathbf{r})$. Hence we retain a valid identity upon functionally differentiating with respect to $a_1(\mathbf{r})$. Differentiating the right hand side of equation (63) trivially gives zero. Differentiating the left hand side gives the bracketed expression, which hence necessarily needs to vanish. Alternatively, one can argue that $a_1(\mathbf{r})$ is a mere test function of arbitrary form and hence equation (63) can only hold provided the bracketed expression vanishes. The result constitutes a derivation of the standard two-body Ornstein–Zernike relation starting from the hyper-Ornstein–Zernike relation (63), providing a demonstration of the self-consistency of the framework.

For the present one-body observables (57) the following two terms that appear in equation (54) become identical to each other:

$$\mathbf{S}_A(\mathbf{r}) = \rho(\mathbf{r})\nabla c_A(\mathbf{r}), \tag{64}$$

where we recall that $\mathbf{S}_A(\mathbf{r}) = \langle\sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla_i \hat{A}(\mathbf{r}^N)\rangle$; see the corresponding operator identity (45). To prove the relationship (64), we use equation (57) to re-write the left hand side as $\langle\sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla_i a_1(\mathbf{r}_i)\rangle = \rho(\mathbf{r})\nabla a_1(\mathbf{r}) = \rho(\mathbf{r})\nabla c_A(\mathbf{r})$, where in the first step we have started from $\hat{A}(\mathbf{r}^N) = \sum_i a_1(\mathbf{r}_i)$, which gives $\nabla_i \hat{A}(\mathbf{r}^N) = \nabla_i a_1(\mathbf{r}_i)$, and then identified the specific form $c_A(\mathbf{r}) = a_1(\mathbf{r})$ according to equation (60).

The resulting equation (60) allows one to simplify equation (54) in the following form

$$\int \mathrm{d}\mathbf{r}' a_1(\mathbf{r}')\,\mathrm{cov}\left(\beta\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r}),\hat{\rho}(\mathbf{r}')\right) = \chi_A(\mathbf{r})\nabla c_1(\mathbf{r};[\rho]) + \rho(\mathbf{r})\int \mathrm{d}\mathbf{r}' \chi_A(\mathbf{r}')\nabla c_2(\mathbf{r},\mathbf{r}';[\rho]). \tag{65}$$

On the above left hand side we have re-written $\mathrm{cov}(\beta\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r}),\hat{A}(\mathbf{r}^N)) = \int \mathrm{d}\mathbf{r}' a_1(\mathbf{r}')\mathrm{cov}(\beta\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r}),\hat{\rho}(\mathbf{r}'))$. We have furthermore used equation (64) to simplify. Functional differentiation of equation (65) with respect to $a_1(\mathbf{r}'')$ yields upon taking account of the form (62) of $\chi_A(\mathbf{r})$ the result:

$$\mathrm{cov}\left(\beta\hat{\mathbf{F}}_{\mathrm{int}}(\mathbf{r}),\hat{\rho}(\mathbf{r}'')\right) = H_2(\mathbf{r},\mathbf{r}'')\nabla c_1(\mathbf{r};[\rho]) + \rho(\mathbf{r})\int \mathrm{d}\mathbf{r}' H_2(\mathbf{r}',\mathbf{r}'')\nabla c_2(\mathbf{r},\mathbf{r}';[\rho]). \tag{66}$$

Here the right hand side can be obtained as a functional derivative of the one-body interparticle force density distribution, $-\delta\mathbf{F}_{\mathrm{int}}(\mathbf{r})/\delta V_{\mathrm{ext}}(\mathbf{r}'')$, using equation (53), the functional chain rule, the identity $H_2(\mathbf{r},\mathbf{r}'') = -\delta\rho(\mathbf{r})/\delta\beta V_{\mathrm{ext}}(\mathbf{r}'')$. The left hand side of of equation (66) is the identical expression $-\delta\mathbf{F}_{\mathrm{int}}(\mathbf{r})/\delta V_{\mathrm{ext}}(\mathbf{r}')$ according to the general identity $\delta A/\delta V_{\mathrm{ext}}(\mathbf{r}) = -\beta\mathrm{cov}(\hat{\rho}(\mathbf{r}),\hat{A})$ by Eckert *et al.* [95]; see also equation (10). We take the demonstration that the general hyperdensity functional theory reduces to known limits for the case of one-body observables (57) as an indication for the internal consistency of our approach.

A specific example could be a simple static model countoscope [107], where we choose $\hat{A}(\mathbf{r}^N) = \int \mathrm{d}\mathbf{r} a_1(\mathbf{r})\hat{\rho}(\mathbf{r})$ with $a_1(\mathbf{r}) = \Theta(l-|\mathbf{r}|)$ being an indicator function for the spherical countoscope of diameter $2l$; here $\Theta(\cdot)$ indicates the Heaviside unit step function. Alternatively, one could use a version with planar symmetry, $a_1(\mathbf{r}) = \Theta(l-|x|)$.

A formally even simpler case is recovered when considering the total number of particles N. Then, the framework reduces, up to a scaling factor β, to considering the local compressibility [89–96], $\chi_\mu(\mathbf{r}) = \beta\chi_A(\mathbf{r})$, as we demonstrate in the following. We choose $\hat{A} = N$, which corresponds to $a_1(\mathbf{r}) = 1$ in equation (57). Following equation (60) the hyperdirect correlation functional becomes constant unity, $c_A(\mathbf{r};[\rho]) = 1$. As a consequence the hyper-Ornstein–Zernike equation (27) reduces upon multiplication by $\rho(\mathbf{r})$ to

$$\rho(\mathbf{r})\int \mathrm{d}\mathbf{r}' c_2(\mathbf{r},\mathbf{r}';[\rho])\chi_\mu(\mathbf{r}') + \beta\rho(\mathbf{r}) = \chi_\mu(\mathbf{r}), \tag{67}$$

which constitutes the exact fluctuation Ornstein–Zernike relation [94, 95] for the local compressibility $\chi_\mu(\mathbf{r}) = \beta\chi_A(\mathbf{r})$. The general functional integral (38) reduces according to equation (59) to the explicit result $A[\rho] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})$. This clearly is the mean number of particles, expressed as a density functional, due to $\int \mathrm{d}\mathbf{r}\rho(\mathbf{r}) = \int \mathrm{d}\mathbf{r}\langle\hat{\rho}(\mathbf{r})\rangle = \langle\int \mathrm{d}\mathbf{r}\hat{\rho}(\mathbf{r})\rangle = \langle N\rangle = A$.

4.2. Two-body observables

The reduction of the hyperdensity functional framework to spatial integration of the standard Ornstein–Zernike relation extends beyond the above one-body forms of $\hat{A}(\mathbf{r}^N)$. In the two-body case we have

$$\hat{A}(\mathbf{r}^N) = \sum_{ij} a_2(\mathbf{r}_i,\mathbf{r}_j), \tag{68}$$

where the sums over i and j run over all particles and the function $a_2(\mathbf{r},\mathbf{r}')$ is a given two-body field. By inserting two density operators we can re-write equation (68) as

$$\hat{A}\left(\mathbf{r}^N\right) = \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\hat{\rho}\left(\mathbf{r}\right)\hat{\rho}\left(\mathbf{r}'\right)a_2\left(\mathbf{r},\mathbf{r}'\right). \tag{69}$$

In order to formulate the mean A, we build the thermal average of equation (69). Observing that $\langle\hat{\rho}(\mathbf{r})\hat{\rho}(\mathbf{r}')\rangle = H_2(\mathbf{r},\mathbf{r}') + \rho(\mathbf{r})\rho(\mathbf{r}')$ we obtain

$$\begin{aligned} A = &\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\rho\left(\mathbf{r}\right)\rho\left(\mathbf{r}'\right)a_2\left(\mathbf{r},\mathbf{r}'\right) \\ &+ \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'H_2\left(\mathbf{r},\mathbf{r}'\right)a_2\left(\mathbf{r},\mathbf{r}'\right), \end{aligned} \tag{70}$$

where the first term on the right hand side is an explicit density functional. Expressing the second term also as a density functional however requires to have access to $H_2(\mathbf{r},\mathbf{r}';[\rho])$ as a density functional. Conventionally one would base this on solving the inhomogeneous two-body Ornstein–Zernike equation (26) for the specific situation at hand, which can be numerically demanding; see e.g. [79, 102].

4.3. Interparticle energy as an observable

For the specific case of the considered observable being identical to the interparticle interaction potential, $\hat{A}(\mathbf{r}^N) = u(\mathbf{r}^N)$, the generic hyperfluctuation profile in covariance form (7) becomes

$$\chi_u\left(\mathbf{r}\right) = \mathrm{cov}\left(\hat{\rho}\left(\mathbf{r}\right), u\left(\mathbf{r}^N\right)\right). \tag{71}$$

Up to a scaling factor of $1/(k_\mathrm{B}T^2)$ this is identical to the interparticle contribution to the local thermal susceptibility $\chi_{T,\mathrm{int}}(\mathbf{r}) = \mathrm{cov}(\hat{\rho}(\mathbf{r}),u(\mathbf{r}^N))/(k_\mathrm{B}T^2)$, as identified by Eckert *et al* [95]. Hence $\chi_{T,\mathrm{int}}(\mathbf{r}) = \chi_u(\mathbf{r})/(k_\mathrm{B}T^2)$. These authors have proven a contact theorem at a hard wall, $\chi_{T,\mathrm{int}}(0^+) = \langle u(\mathbf{r}^N)\rangle/(Vk_\mathrm{B}T^2)$; see [95] for the derivation. From the general hypercontact theorem (40) we reproduce their result in the equivalent form $\chi_u(0^+) = \langle u(\mathbf{r}^N)\rangle/V$.

In the present case the general hyperforce density balance (49) reduces to the previously found identity [82]

$$\mathrm{cov}\left(\beta\hat{\mathbf{F}}\left(\mathbf{r}\right), u\left(\mathbf{r}^N\right)\right) = \mathbf{F}_\mathrm{int}\left(\mathbf{r}\right), \tag{72}$$

which we can re-write by using the definition of $\chi_u(\mathbf{r})$ and re-ordering as

$$\begin{aligned} &\mathrm{cov}\left(\beta\hat{\mathbf{F}}_\mathrm{int}\left(\mathbf{r}\right), u\left(\mathbf{r}^N\right)\right) \\ &\quad = \mathbf{F}_\mathrm{int}\left(\mathbf{r}\right) + \nabla\chi_u\left(\mathbf{r}\right) + \chi_u\left(\mathbf{r}\right)\nabla\beta V_\mathrm{ext}\left(\mathbf{r}\right). \end{aligned} \tag{73}$$

The general hyper-Ornstein–Zernike relation (27) retains its form as

$$c_u\left(\mathbf{r};[\rho]\right) = \frac{\chi_u\left(\mathbf{r};[\rho]\right)}{\rho\left(\mathbf{r}\right)} - \int \mathrm{d}\mathbf{r}'c_2\left(\mathbf{r},\mathbf{r}';[\rho]\right)\chi_u\left(\mathbf{r}'\right), \tag{74}$$

which can be viewed as the interparticle contribution to the fluctuation Ornstein–Zernike equation for the thermal susceptibility $\chi_\mathrm{T}(\mathbf{r})$ [94, 95]. For additional generalizations, we refer the Reader to the very recent study by Kampa *et al* [44], who derived and applied meta-Ornstein–Zernike relations on the basis of the explicit treatment of the pair interaction potential via neural metadensity functionals.

5. Machine learning neural hyperdensity functionals

5.1. Hyperdensity functional workflow

We first lay out the principal workflow for the application of the hyperdensity functional theory to a concrete physical problem. We assume that all necessary functional relationships are accessible in a concrete way. After establishing the formal procedure, we then describe how supervised machine learning allows one to train neural networks that provide the required functional closure for the theory. Figure 3 depicts a graphical representation of the hyperdensity functional workflow. The relationships between the different relevant functionals and correlation functions are illustrated on the basis of data for the clustering in the one-dimension hard rod model, which is taken as a prototypical case for a non-trivial observable in [40].

We turn to the general structure. We assume that a specific observable $\hat{A}$ has been chosen as the relevant target quantity of interest for a given many-body system, as specified by its Hamiltonian H. The extended ensemble, as described in section 2.1, requires no explicit treatment. The aim is to study the equilibrium statistical mechanical behaviour for, in general, spatially inhomogeneous systems at state points μ, T, and for given form of the external potential $V_\mathrm{ext}(\mathbf{r})$. Results for the equilibrium density profile $\rho(\mathbf{r})$ follow from standard density functional minimization (14) on the basis of the chosen excess free energy functional $F_\mathrm{exc}[\rho]$ and corresponding one-body direct correlation functional $c_1(\mathbf{r};[\rho])$, cf their functional derivative relationship (17). Typically one obtains the solution of the Euler–Lagrange equation (16) numerically. Then evaluating the grand potential functional $\Omega[\rho]$ gives access to the thermodynamics. Furthermore the two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$ follows from the functional derivative relationships (24) or (25), which gives access to the pair structure; see e.g. [39] for recent work.

This standard density functional structure provides a backdrop for the hyperfunctional application. Using the thus obtained equilibrium density profile $\rho(\mathbf{r})$ and either evaluating the hyperdensity functional $A[\rho]$ directly or via performing the functional integral (38) of the one-body hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ then gives the desired thermal mean A for the considered system.

The hyper-Ornstein–Zernike equation (27) allows one to determine the corresponding hyperfluctuation profile $\chi_A(\mathbf{r})$. We recall the role of $\chi_A(\mathbf{r})$ as a spatially resolved measure of correlations of the local particle number and the value of $\hat{A}$ via the covariance (7). Here, however, we do not have

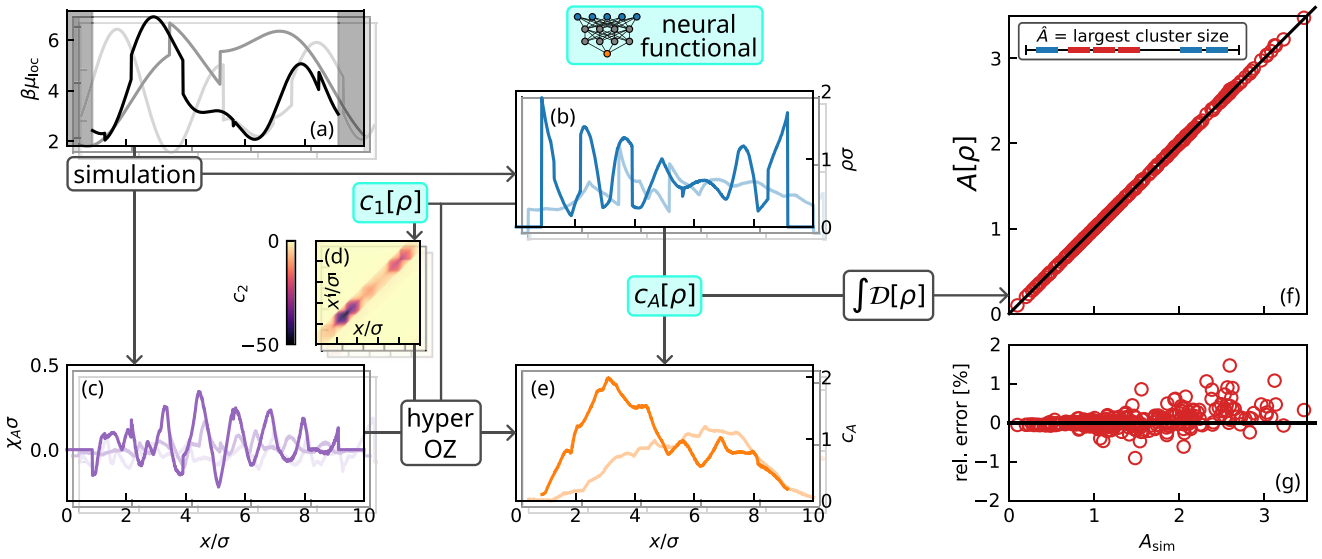

Figure 3. Overview of hyperdensity functional theory applied to the cluster statistics of one-dimensional hard rods of size σ. The observable $\hat{A}$ is the size of the largest cluster, where two particles are bonded provided their mutual distance is $< 1.2\sigma$. (a) The local chemical potential $\beta\mu_{\rm loc}(x) = \beta\mu - \beta V_{\rm ext}(x)$ creates spatially inhomogeneous systems. Shown are representative examples from 512 grand canonical Monte Carlo simulations with both randomized values of $\beta\mu$ and forms of $\beta V_{\rm ext}(x)$. (b) Corresponding scaled density profiles $\rho(x)\sigma$ sampled via equation (3). (c) Corresponding scaled hyperfluctuation profiles $\chi_A(x)\sigma$ obtained via equation (7). (d) Two-body direct correlation functions $c_2(x,x';[\rho])$, as obtained via automatic differentiation from $c_1(x;[\rho])$ [35, 36]. (e) Hyperdirect correlation functions $c_A(x)$ obtained by solving the hyper-Ornstein–Zernike equation (27). Using the density profile as input and the simulation results for $c_A(x)$ as target, supervised training yields a neural network that represents the hyperdirect correlation functional $c_A(x;[\rho])$. (f) Predicted values $A[\rho] = \int D[\rho]c_A(x;[\rho])$ from functional integration according to equation (38). For a test set of 256 systems not encountered during training the predictions of $A[\rho]$ are compared against reference simulation data $A_{\rm sim}$. (g) The relative numerical error of the predicted mean size A of the largest cluster is smaller than $\sim 1\%$. Reprinted (figure) with permission from [40], Copyright (2024) by the American Physical Society.

to resort to many-body sampling, as all information that is required for the application of equation (27) is attainable via two functional relationships: (i) evaluating the hyperdirect correlation functional gives the concrete hyperdirect correlation function $c_A(\mathbf{r}) = c_A(\mathbf{r};[\rho])$ for the specific system at hand, and (ii) evaluating the two-body direct correlation functional, $c_2(\mathbf{r},\mathbf{r}') = c_2(\mathbf{r},\mathbf{r}';[\rho])$, gives the standard two-body direct correlation function $c_2(\mathbf{r},\mathbf{r}')$, which is ready to act as a generalized convolution kernel in the hyper-Ornstein–Zernike equation (27). Due to $\chi_A(\mathbf{r})$ appearing both in bare form and inside of the generalized, i.e. fully position-dependent, spatial convolution one can solve equation (27) straightforwardly for $\chi_A(\mathbf{r})$, as this constitutes a system of linear equations when using spatial discretization.

We next turn to three-dimensional hard sphere fluids and describe our machine learning strategy for the application of the hyperdensity functional approach to the cluster statistics, in generalization of the one-dimensional hard rod system shown in figure 3.

5.2. *Training neural hyperdensity functionals*

Carrying out the work programme described in section 5.1 requires one to have prescriptions for the excess free energy functional $F_{\rm exc}[\rho]$ and for the hyperdensity functional $A[\rho]$ to arrive at a closed theory. Alternatively and equivalently based on functional integration, which in practice is numerically straightforward [35, 36, 38–40], one can start with the one-body direct correlation functional $c_1(\mathbf{r};[\rho])$ and the hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$. We recall their respective relationships with $F_{\rm exc}[\rho]$ and $A[\rho]$ via functional differentiation according to equations (17) and (36), as well as via functional integration according to equations (19) and (38).

To facilitate concrete access to these functionals, we resort to simulation-based supervised machine learning and use local learning of one-body direct correlation functionals [35–40]. Constructing the density dependence of the direct correlation functional $c_1(\mathbf{r};[\rho])$ follows the methodology of [35–39] where the Euler–Lagrange equation (16) is used to generate training data that consists of target values $c_1(\mathbf{r})$ for given density profile $\rho(\mathbf{r}')$ within a limited range of positions $\mathbf{r}'$ around the target location $\mathbf{r}$. The application to the one-dimensional hard rod model [36, 37] indicates numerical performance that is in practice equivalent to that of Percus' exact solution [108, 109]. Functional differentiation, as is readily accessible via automatic differentiation [35–39, 57, 61] yields the two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$. As we are interested in hard sphere behaviour we re-use the trained neural functional of [35].

Training the neural hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ proceeds along very similar lines as training the standard one-body direct correlation functional $c_1(\mathbf{r};[\rho])$. The training data is acquired by grand canonical Monte Carlo simulations, where the chemical potential μ, the temperature T and the shape of the external potential $V_{\rm ext}(\mathbf{r})$ are prescribed (in practice in a randomized way). The density profile $\rho(\mathbf{r})$ is accessible via sampling according to equation (3) (or via more advanced methods [11, 18]). For the chosen observable $\hat{A}$, data for the hyperfluctuation profile $\chi_A(\mathbf{r})$ follows from sampling the covariance (7). As the two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$ is known from the above standard density functional treatment, the complete information is available to numerically evaluate the right hand side of the hyper-Ornstein–Zernike equation (27) and hence to construct the one-body hyperdirect correlation function $c_A(\mathbf{r})$ for the specific training system under consideration. This data is then used as the target for training the neural representation of the one-body hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ using solely the density profile $\rho(\mathbf{r}')$ as an input following the lines of [35–40]. We use a simple (fully connected) multi-layer perceptron to represent $c_A(\mathbf{r};[\rho])$ and refer the Reader to [110] for all technical details.

We emphasize that the above laid out machine learning strategy is heavily influenced by the successful neural functional methodology of local learning of one-body direct correlation functionals [35–40]. This method represents the classical density functional framework [2–6] in a very direct way. Moreover, it is computationally straightforward to implement in all its three key aspects of: data generation via simulation, training of the neural network, and numerical application; we refer to [39] for a demonstration of the breadth of applicability of the resulting neural theory in the context of gas–liquid phase separation.

This neural functional approach needs to be contrasted with brute force machine learning, where one could envisage bypassing the density functional relationships and working directly with the ensemble averages in order to represent the mean A via equation (5) and machine learn the relationship $V_{\rm ext}(\mathbf{r}) \to A$. The current strategy is very different, as it resorts to machine learning only the genuinely non-trivial hyperdirect correlation functionals and letting the encompassing structure be informed by the theoretical physics. This is analogous to putting the focus on $c_1(\mathbf{r};[\rho])$ in standard DFT treatments, instead of considering other possible mappings of interest [55, 56] or merely mimicking the simulation procedure, i.e. learning $V_{\rm ext}(\mathbf{r}) \to \rho(\mathbf{r})$.

5.3. Application to cluster statistics

We have applied the machine learning strategy of section 5.2 to three-dimensional hard spheres in planar inhomogeneous environments, choosing $\hat{A}$ as the size of the largest cluster in a given microstate. Specifically, we define two particles i and j as being bonded provided that their mutual distance is below a cut-off value, which we choose as $|\mathbf{r}_i - \mathbf{r}_j| < r_c = 1.2\sigma$, where σ is the hard sphere diameter. Clusters are then defined as groups of particles that are (transitively) bonded. Each cluster consists of a specific number of particles such that each configuration $\mathbf{r}^N$ is associated with a specific distribution of cluster sizes. We choose the number of particles in the largest occurring cluster as our target observable $\hat{A}$.

We use a fixed box size with lateral area $5\sigma \times 5\sigma$ and an elongated x-direction of length $L = 10\sigma$, along which the system is spatially inhomogeneous. We first consider confinement by two parallel hard walls represented by the scaled external potential $\beta V_{\rm ext}(x) = \infty$ for $x/\sigma < 1$ and $x/\sigma > 9$, and zero otherwise. We show in figure 4 (taken from the SI of [40]) results from the hyperdensity functional theory compared against stand-alone simulation results that provide the reference. The density profile exhibits the familiar oscillatory behaviour, as is induced by hard sphere packing, see Figure 4(a) for results over a range of increasing values of the scaled chemical potential $\beta\mu$. We have used the neural hard sphere functional of [35], which was shown to outperform in accuracy the White Bear Mk. II version of fundamental measure theory [111–113].

The corresponding one-body hyperdirect correlation functions $c_A(x)$ are obtained from evaluating the neural hyperdirect correlation functional at the respective equilibrium density profile $\rho(x)$. Here x denotes the coordinate of $\mathbf{r}$ along which the system is spatially inhomogeneous. The spatial variation of $c_A(x)$ is much smoother than the corresponding density profile and a pronounced broad peak is apparent at the centre of the system, see Figure 4(b). We attribute the latter feature to the strong effect on the size of the largest cluster that follows from having particles near the centre of the system. Crucially, via functional integration of $c_A(x;[\rho])$, the thermal average A, i.e. the mean size of the largest hard-sphere cluster, can be reproduced faithfully, as is evident from comparison to simulation data.

The scaled hyperfluctuation profiles $\chi_A(x)\sigma^3$ shown in Figure 4(c) possess very pronounced oscillations with an envelope that decays towards either wall. We recall that $\chi_A(x)$ measures the covariance (7) of having a particle at x with the size of largest cluster in the system.

In figure 5 we compare the clustering behaviour in the three following situations of i) bulk fluids, where the external potential vanishes, $V_{\rm ext}(x) = 0$, ii) planar confinement inside of a double well potential represented by the following form: $\beta V_{\rm ext}(x) = 3[(x/\sigma - 5)^2 - 2.5^2]^2/2.5^4 - 1.5$, and iii) the above hard wall case, which we use as a reference. The perhaps most striking difference between the double well and the hard wall confinement is the overcoming of the depression that both $c_A(x)$ and $\chi_A(x)$ display at the centre of the double-well system. This can be attributed intuitively to the hump of the potential well loosing its disrupting effect on the clustering upon increasing chemical potential.

6. Conclusions

In conclusion we have described in detail the recent generalization of classical density functional theory towards the description of the equilibrium statistical properties of virtually arbitrary observables $\hat{A}(\mathbf{r}^N)$. Thereby the dependence on

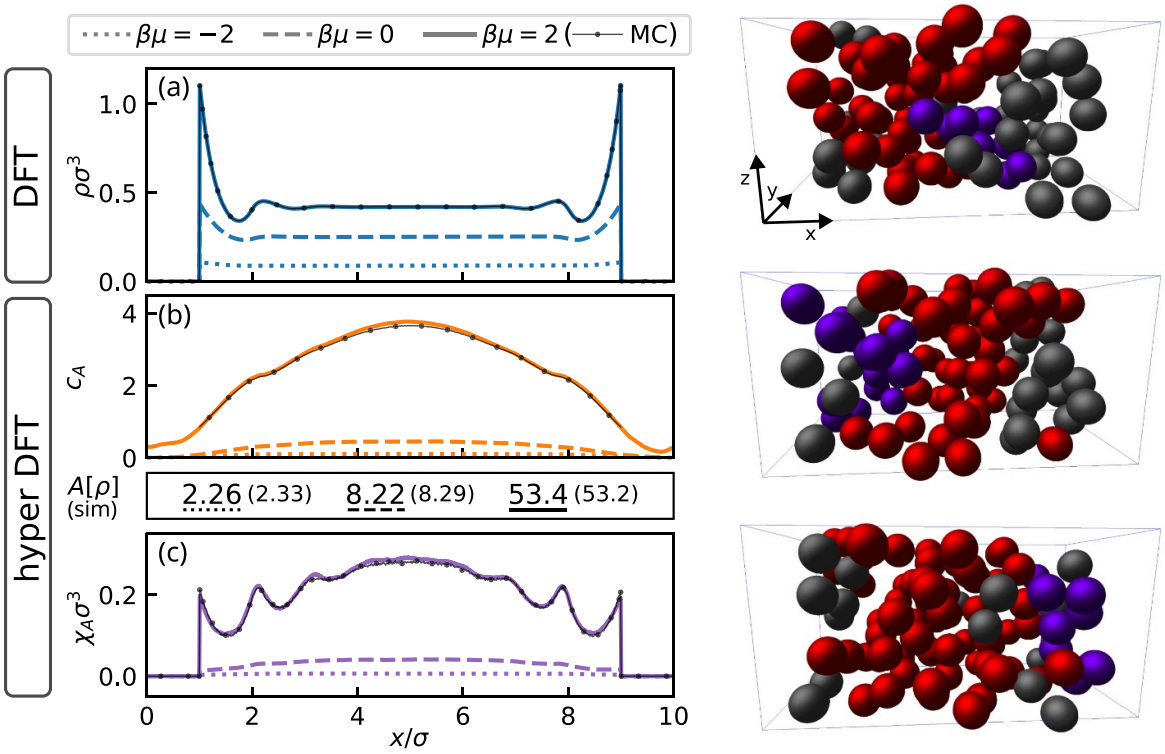

Figure 4. Cluster statistics of three-dimensional hard spheres confined between two parallel hard walls with distance 8σ. The observable $\hat{A}$ is the number of particles in the largest cluster. Theoretical results are shown for $\beta\mu = -2$ (dotted), 0 (dashed), and 2 (solid lines, symbols indicate reference simulation data). (a) The scaled density profile $\rho(x)\sigma^3$ as a function of the scaled distance x/σ across the slit is obtained from numerical solution of the Euler–Lagrange equation (16) with the neural hard sphere one-body direct correlation functional $c_1(x;[\rho])$ of [35]. (b) Corresponding hyperdirect correlation functions $c_A(x)$ from evaluating the neural hyperdirect correlation functional $c_A(x;[\rho])$ at the three respective density profiles. Functional integration according to equation (38) gives predictions for the mean A of the size of the largest cluster, as compared to the simulation reference $A = \langle\hat{A}\rangle$ (values in parenthesis). (c) Hyperfluctuation profiles $\chi_A(x)$ are obtained from solving the hyper-Ornstein–Zernike relation (27) for the three considered situations using as input $\rho(x)$ and the neural functionals $c_A(x;[\rho])$ and $c_2(x,x';[\rho]) = \delta c_1(x;[\rho])/\delta\rho(x')$. The simulation reference is obtained via sampling $\chi_A(x) = \mathrm{cov}(\hat{\rho}(x),\hat{A})$ according to equation (7). The three simulation snapshots (right column) show hard sphere configurations for $\beta\mu = 2$. The highlighted particles belong to the largest cluster (bright red) or to the second-largest cluster (dark violet). The number $\hat{A}$ of particles in the largest cluster fluctuates considerably over microstates. Reprinted (figure) with permission from [40], Copyright (2024) by the American Physical Society.

the position coordinates $\mathbf{r}^N$ of all particles in the system can be very general and hence can accommodate a multitude of physical quantities including complex, intricate order parameters of interest. Our terminology parallels Hirschfelder's hypervirial theorem [86] to refer to a generalization to arbitrary observables. The hyperdensity functional theory [40] is formally exact and we have provided detailed descriptions of the underlying generalized statistical ensemble, including the natural emergence of the hyperfluctuation profile $\chi_A(\mathbf{r})$ as a spatially resolved measure of fluctuations of the observable under consideration. The exact hyper-Ornstein–Zernike equation (27) relates $\chi_A(\mathbf{r})$ to the one-body hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ and the standard two-body direct correlation functional $c_2(\mathbf{r},\mathbf{r}';[\rho])$ plays its common role as a generalized convolution kernel that mediates spatial correlations over finite distances.

Our derivation of the hyper-Ornstein–Zernike relation (27) generalizes and connects to the standard two-body Ornstein–Zernike relation [6], the dynamical nonequilibrium Ornstein–Zernike framework [6, 105, 106], the local compressibility and the local thermal susceptibility [94, 95], and the very recent meta-Ornstein–Zernike framework for the explicit treatment of the form of the pair potential [44].

Besides constituting the spatial inhomogeneity in the hyper-Ornstein–Zernike equation (27) and thus driving the spatial structuring of the hyperfluctuation profile $\chi_A(\mathbf{r})$, the hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ also offers an avenue to express the thermal average of the observable as a genuine density functional, $A[\rho]$. The construction of the functional dependence involves formal functional integration, which can in practice be numerically carried out with great efficiency provided that the functional integrand $c_A(\mathbf{r};[\rho])$ is available. Centring the approach around the one-body hyperdirect correlation functional $c_A(\mathbf{r};[\rho])$ is motivated by its concrete accessibility via simulation-based supervised machine learning. Here a neural network is trained to act as a surrogate for the formally defined exact functional.

The training data is provided by grand canonical Monte Carlo simulations. Crucially, only standard techniques are thereby necessary, without any need to explicitly consider the extended ensemble on which the derivations are based formally. The ensemble extension merely serves as a device

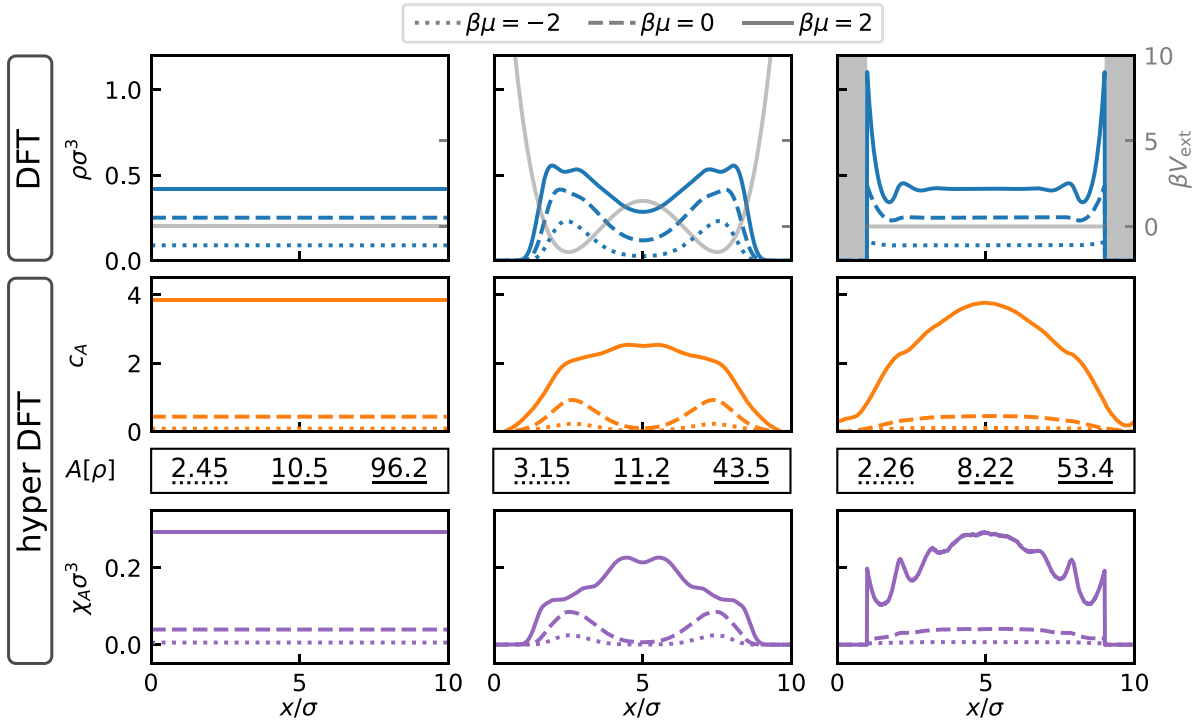

Figure 5. Hyperdensity functional results for the largest cluster in hard sphere fluids in bulk (left column), in a double well potential (middle column), and between two parallel hard walls (right column, reproduced from figure 4); see the corresponding scaled external potentials $\beta V_{\text{ext}}(x)$ (gray lines in the top panels). Shown are as a function of the scaled distance x/σ: the scaled density profile $\rho(x)\sigma^3$ (top row), the hyperdirect correlation function $c_A(x)$ (middle row), and the scaled hyperfluctuation profile $\chi_A(x)\sigma^3$ (bottom row). The scaled chemical potential has values $\beta\mu = -2, 0, 2$ (as indicated) and corresponding values of the mean size of the largest cluster $A[\rho]$ are given for each system. Note the striking structural difference between the double well potential (middle column) and the planar pore (right column).

to establish the required exact functional relationships that underlie the supervised machine learning. Pre-processing of data involves a (standard) generalized convolution operation with a kernel given by the two-body direct correlation functional, which in turn is represented by a trained neural functional [35–37]. In practice these computational demands are entirely manageable.

The hyperfluctuation profile $\chi_A(\mathbf{r})$ has previously emerged as a key quantity in the context of the hyperforce correlation theory [82] that is based on considering thermal gauge invariance [83, 84] of the mean value A of the given observable. Earlier versions are the local compressibility $\chi_\mu(\mathbf{r})$ [89–91] and the thermal susceptibility $\chi_T(\mathbf{r})$ [92, 94, 95]. While both the statistical mechanical Noetherian methodology [40, 75–82] and density functional theory itself [2–5] put functional relationships at the centre of theory construction, both approaches are complementary and hence the respective prominent occurrence of $\chi_A(\mathbf{r})$ is very noteworthy, see equation (49). Together with a corresponding interparticle hyperforce contribution, the total one-body hyperforce density correlation function is restored and ultimately related via equation (50) to the mean localized phase space gradient of the observable $\hat{A}(\mathbf{r}^N)$ under consideration [82], see the defining equation (45) of the corresponding hyperforce phase space function.

We have shown that the hyperforce density balance relationship, see equation (49), follows naturally from the extended ensemble. The derivation is based on starting from the standard force density balance in the extended ensemble and parametrically differentiating with respect to the coupling parameter λ that tunes the strength of the generalized contribution. We take the existence of this additional route to the hyperforce balance, besides the derivations from Noether invariance [82], from ad hoc phase space integration by parts [82], and from operator methods within gauge invariance [83, 84] both as an indicator for the fundamental status of this sum rule as well as for a sign of consistent treatment of the statistical mechanics within the hyperforce, hyperdensity, and gauge invariance frameworks.

These theoretical developments are interesting due to the fundamental structure that they reveal to be present in the many-body statistical physics. Furthermore, as we demonstrated, they provide a concrete blueprint for devising machine learning schemes and in particular for working with neural direct and hyperdirect correlation functionals. These neural networks enable: (i) immediate access to the relevant functional dependencies, (ii) efficient functional calculus via fast numerical functional integration and via automatic functional differentiation, as is relevant for accessing generating functionals and higher-order correlation functions, (iii)

carrying out consistency checks for the theoretical concepts and for the numerics, and (iv) gaining fresh and significantly deep insight into the correlated many-body physics under investigation.

The relevance of the hyperdensity functional theory lies in its combination of the rigorous aspects of classical density functional theory, and hence the modern view of statistical mechanics of working with functional relationships, with the wide spectrum of different types of observables, order parameters, and general quantities of interest that are in present-day use and often directly accessible via many-body simulations. In such work one commonly faces the task of finding physical meaning and structure in the simulation output. The hyperdensity functional theory offers a formal framework for performing this task without having to resort to costly many-body simulations.

To provide a specific example for this strategy, we have provided a hard wall hypercontact theorem that relates the contact value $\chi_A(0^+)$ of the hyperfluctuation profile at a hard wall with the mean bulk value A_b of the corresponding observable per system volume. Given this quite counter-intuitive relationship, one would surely be hard-pressed to discover the relation by mere inspection of simulation data. A similar contact theorem holds for Evans and coworkers' local compressibility $\chi_\mu(\mathbf{r})$ [89–91] as used in a multitude of interfacial studies. As we demonstrated, the local compressibility is recovered in the general framework when making the specific choice $\hat{A} = \beta N$.

For several different classes of specific forms of the general observable $\hat{A}(\mathbf{r}^N)$ we have investigated the thus arising simplifications. In particular, for one-body forms of $\hat{A}(\mathbf{r}^N)$ all relevant information is already available within a standard density functional treatment. Treating two-body observables within the standard approach requires (numerical) solution of the inhomogeneous two-body Ornstein–Zernike equation, which typically comes at non-negligible computational expense. We have shown that the hyperdensity approach is consistent with the standard method, while it does not suffer from higher-order restrictions, as it entirely operates on the one-body level. As a challenging test, we have considered cluster statistics of hard spheres, which are inaccessible in standard density functional treatments. As we have demonstrated, the hyperdensity functional framework gives ready access to this complex order parameter and its associated correlation and fluctuation measures.

Future work could address the analytical construction of hyperdensity functionals for specific observables. It would be interesting to see whether concepts from fundamental measure theory [113–115] could be used in the construction of hard sphere hyperdensity functionals. Also generalizations of fundamental measure concepts beyond hard sphere systems acquire new relevance, in particular the generalized weight functions for soft interactions [116–120]. As was the case for the Rosenfeld functional [113–115], inspiration could come from liquid integral equation theory [5], see [121] for recent work. Also investigating relationships with the internal-energy functional formulation [122] and with functional thermodynamics [99] could be interesting. We have here worked with a fixed finite system size. Investigating the scaling behaviour of the cluster statistics with changing system size is a relevant topic that could possibly be addressed using the (inverse) system size as an input to the neural functional, see [39]. This would allow to address multi-scale questions [123–126], as previously demonstrated successfully for neural density functionals [35–38].

Apart from these important conceptual points, it would be highly interesting to use machine learning and consider the application of neural hyperdensity functionals to a wider variety of order parameters and systems. This would potentially facilitate the investigation of physical phenomena at the much increased numerical efficiency that the neural hyperdensity functional theory delivers. We here have provided the full theoretical and methodological background required for engaging in such endeavours. The generalization of the hyperdensity functional framework to the multivariate case of several simultaneous observables of interest will be presented elsewhere [127].

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5281/zenodo.13318634.

Acknowledgments

We thank Silas Robitschko, Sophie Hermann, and Johanna Müller for useful discussions. This work is supported by the DFG (Deutsche Forschungsgemeinschaft) under Project No. 551294732.

ORCID iDs

Florian Sammüller https://orcid.org/0000-0002-3605-329X
Matthias Schmidt https://orcid.org/0000-0002-5015-2972

References

[1] Mermin N D 1965 Thermal properties of the inhomogeneous electron gas *Phys. Rev.* **137** A1441
[2] Evans R 1979 The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids *Adv. Phys.* **28** 143
[3] Evans R 1992 Density functionals in the theory of nonuniform fluids *Fundamentals of Inhomogeneous Fluids*, ed D Henderson (Marcel Dekker)
[4] Evans R, Oettel M, Roth R and Kahl G 2016 New developments in classical density functional theory *J. Phys.: Condens. Matter* **28** 240401
[5] Hansen J P and McDonald I R 2013 *Theory of Simple Liquids* 4th edn (Academic)
[6] Schmidt M 2022 Power functional theory for many-body dynamics *Rev. Mod. Phys.* **94** 015007

[7] Hohenberg P and Kohn W 1964 Inhomogeneous electron gas *Phys. Rev.* **136** B864
[8] Kohn W 1999 Nobel lecture: electronic structure of matter–wave functions and density functionals *Rev. Mod. Phys.* **71** 1253
[9] Ornstein L S and Zernike F 1914 Accidental deviations of density and opalescence at the critical point of a single substance *Proc. Acad. Sci.* **17** 793 (available at: https://dwc.knaw.nl/DL/publications/PU00012727.pdf)
[10] Percus J K 1962 Approximation methods in classical statistical mechanics *Phys. Rev. Lett.* **8** 462
[11] Frenkel D and Smit B 2023 *Understanding Molecular Simulation: From Algorithms to Applications* 3rd edn (Academic)
[12] Wilding N B 2001 Computer simulation of fluid phase transitions *Am. J. Phys.* **69** 1147
[13] Brukhno A V, Grant J, Underwood T L, Stratford K, Parker S C, Purton J A and Wilding N B 2021 DL_MONTE: a multipurpose code for Monte Carlo simulation *Mol. Simul.* **47** 131
[14] Borgis D, Assaraf R, Rotenberg B and Vuilleumier R 2013 Computation of pair distribution functions and three-dimensional densities with a reduced variance principle *Mol. Phys.* **111** 3486
[15] de las Heras D and Schmidt M 2018 Better than counting: density profiles from force sampling *Phys. Rev. Lett.* **120** 218001
[16] Coles S W, Borgis D, Vuilleumier R and Rotenberg B 2019 Computing three-dimensional densities from force densities improves statistical efficiency *J. Chem. Phys.* **151** 064124
[17] Coles S W, Mangaud E, Frenkel D and Rotenberg B 2021 Reduced variance analysis of molecular dynamics simulations by linear combination of estimators *J. Chem. Phys.* **154** 191101
[18] Rotenberg B 2020 Use the force! Reduced variance estimators for densities, radial distribution functions and local mobilities in molecular simulations *J. Chem. Phys.* **153** 150902
[19] Mangaud E and Rotenberg B 2020 Sampling mobility profiles of confined fluids with equilibrium molecular dynamics simulations *J. Chem. Phys.* **153** 044125
[20] Coles S W, Morgan B J and Rotenberg B, RevelsMD: reduced variance estimators of the local structure in molecular dynamics (arXiv:2310.06149)
[21] Renner J, Schmidt M and de las Heras D 2023 Reduced-variance orientational distribution functions from torque sampling *J. Phys.: Condens. Matter* **35** 235901
[22] Moustafa S G, Schultz A J and Kofke D A 2015 Very fast averaging of thermal properties of crystals by molecular simulation *Phys. Rev.* E **92** 043303
[23] Schultz A J, Moustafa S G, Lin W, Weinstein S J and Kofke D A 2016 Reformulation of ensemble averages via coordinate mapping *J. Chem. Theory Comput.* **12** 1491
[24] Moustafa S G, Schultz A J and Kofke D A 2017 Harmonically assisted methods for computing the free energy of classical crystals by molecular simulation: a comparative study *J. Chem. Theory Comput.* **13** 825
[25] Moustafa S G, Schultz A J, Zurek E and Kofke D A 2017 Accurate and precise *abinitio* anharmonic free-energy calculations for metallic crystals: application to hcp Fe at high temperature and pressure *Phys. Rev.* B **96** 014117
[26] Schultz A J and Kofke D A 2018 Comprehensive high-precision high-accuracy equation of state and coexistence properties for classical Lennard-Jones crystals and low-temperature fluid phases *J. Chem. Phys.* **149** 204508
[27] Purohit A, Schultz A J, Moustafa S G, Errington J R and Kofke D A 2018 Free energy and concentration of crystalline vacancies by molecular simulation *Mol. Phys.* **116** 3027
[28] Moustafa S G, Schultz A J and Kofke D A 2018 Effects of thermostatting in molecular dynamics on anharmonic properties of crystals: application to fcc Al at high pressure and temperature *J. Chem. Phys.* **149** 124109
[29] Purohit A, Schultz A J and Kofke D A 2020 Implementation of harmonically mapped averaging in LAMMPS and effect of potential truncation on anharmonic properties *J. Chem. Phys.* **152** 014107
[30] Moustafa S G, Schultz A J and Kofke D A 2022 Reformulation of expressions for thermoelastic properties of crystals using harmonic mapping *Phys. Rev.* B **106** 104105
[31] Lin W S, Schultz A J and Kofke D A 2018 Electric-field mapped averaging for the dielectric constant *Fluid Phase Equilib.* **470** 174
[32] Trokhymchuk A, Schultz A J and Kofke D A 2019 Alternative ensemble averages in molecular dynamics simulation of hard spheres *Mol. Phys.* **117** 3734
[33] Schultz A J and Kofke D A 2019 Alternatives to conventional ensemble averages for thermodynamic properties *Curr. Opin. Chem. Eng.* **23** 70
[34] Purohit A, Schultz A J and Kofke D A 2019 Force-sampling methods for density distributions as instances of mapped averaging *Mol. Phys.* **117** 2822
[35] Sammüller F, Hermann S, de las Heras D and Schmidt M 2023 Neural functional theory for inhomogeneous fluids: fundamentals and applications *Proc. Natl. Acad. Sci.* **120** e2312484120
[36] Sammüller F, Hermann S and Schmidt M 2024 Why neural functionals suit statistical mechanics *J. Phys.: Condens. Matter* **36** 243002
[37] Sammüller F Neural functional theory for inhomogeneous fluids – Tutorial (available at: https://github.com/sfalmo/NeuralDFT-Tutorial)
[38] Sammüller F and Schmidt M 2024 Neural density functionals: local learning and pair-correlation matching *Phys. Rev.* E **110** L032601
[39] Sammüller F, Schmidt M and Evans R 2024 Neural density functional theory of liquid-gas phase coexistence (resubmitted to Phys. Rev. X) (arXiv:2408.15835)
[40] Sammüller F, Robitschko S, Hermann S and Schmidt M 2024 Hyperdensity functional theory of soft matter *Phys. Rev. Lett.* **133** 098201
[41] de las Heras D, Zimmermann T, Sammüller F, Hermann S and Schmidt M 2023 Perspective: how to overcome dynamical density functional theory *J. Phys.: Condens. Matter* **35** 271501
[42] Zimmermann T, Sammüller F, Hermann S, Schmidt M and de las Heras D 2024 Neural force functional for non-equilibrium many-body colloidal systems *Mach. Learn.: Sci. Technol.* **5** 035062
[43] Bui A T and Cox S J Learning classical density functionals for ionic fluids (arXiv:2410.02556)
[44] Kampa S M, Sammüller F, Schmidt M and Evans R 2024 Metadensity functional theory for classical fluids: extracting the pair potential (arXiv:2411.06972)
[45] Carrasquilla J and Melko R G 2017 Machine learning phases of matter *Nat. Phys.* **13** 431
[46] Bedolla E, Padierna L C and Castañeda-Priego R 2021 Topical review: machine learning for condensed matter physics *J. Phys.: Condens. Matter* **33** 053001
[47] Chertenkov V, Burovski E and Shchur L 2023 Finite-size analysis in neural network classification of critical phenomena *Phys. Rev.* E **108** L032102
[48] Arnold J, Schäfer F, Edelman A and Bruder C 2024 Mapping out phase diagrams with generative classifiers *Phys. Rev. Lett.* **132** 207301

[49] Carvalho F S and Braga J P 2022 Physics-informed neural networks applied to liquid state theory *J. Mol. Liq.* **367** 120504

[50] Chen W, Gao P and Stinis P 2024 Physics-informed machine learning of the correlation functions in bulk fluids *Phys. Fluids* **36** 017133

[51] Wu J and Gu M 2023 Perfecting liquid-state theories with machine intelligence *J. Phys. Chem. Lett.* **14** 10545

[52] Lin S-C and Oettel M 2019 A classical density functional from machine learning and a convolutional neural network *SciPost Phys.* **6** 025

[53] Lin S-C, Martius G and Oettel M 2020 Analytical classical density functionals from an equation learning network *J. Chem. Phys.* **152** 021102

[54] Cats P, Kuipers S, de Wind S, van Damme R, Coli G M, Dijkstra M and van Roij R 2021 Machine-learning free-energy functionals using density profiles from simulations *APL Mater.* **9** 031109

[55] Yatsyshin P, Kalliadasis S and Duncan A B 2022 Physics-constrained Bayesian inference of state functions in classical density-functional theory *J. Chem. Phys.* **156** 074105

[56] Malpica-Morales A, Yatsyshin P, Duran-Olivencia M A and Kalliadasis S 2023 Physics-informed Bayesian inference of external potentials in classical density functional theory *J. Chem. Phys.* **159** 104109

[57] Dijkman J, Dijkstra M, van Roij R, Welling M, van de Meent J W and Ensing B 2024 Learning neural free-energy functionals with pair-correlation matching (arXiv:2403.15007)

[58] Kelley M M, Quinton J, Fazel K, Karimitari N, Sutton C and Sundararaman R 2024 Bridging electronic and classical density-functional theory using universal machine-learned functional approximations *J. Chem. Phys.* **161** 144101

[59] Simon A, Weimar J, Martius G and Oettel M 2024 Machine learning of a density functional for anisotropic patchy particles *J. Chem. Theory Comput.* **20** 1062

[60] Simon A, Belloni L, Borgis D and Oettel M 2024 The orientational structure of a model patchy particle fluid: simulations, integral equations, density functional theory and machine learning (arXiv:2411.06973)

[61] Stierle R, Bauer G, Thiele N, Bursik B, Rehner P and Gross J 2024 Classical density functional theory in three dimensions with GPU-accelerated automatic differentiation: computational performance analysis using the example of adsorption in covalent-organic frameworks *Chem. Eng. Sci.* **298** 120380

[62] Yang J, Pan R, Sun J and Wu J 2024 High-dimensional operator learning for molecular density functional theory (arXiv:2411.03698)

[63] Nagai R, Akashi R, Sasaki S and Tsuneyuki S 2018 Neural-network Kohn-Sham exchange-correlation potential and its out-of-training transferability *J. Chem. Phys.* **148** 241737

[64] Schmidt J, Benavides-Riveros C L and Marques M A L 2019 Machine learning the physical nonlocal exchange-correlation functional of density-functional theory *J. Phys. Chem. Lett.* **10** 6425

[65] Zhou Y, Wu J, Chen S and Chen G 2019 Toward the exact exchange-correlation potential: a three-dimensional convolutional neural network construct *J. Phys. Chem. Lett.* **10** 7264

[66] Nagai R, Akashi R and Sugino O 2020 Completing density functional theory by machine learning hidden messages from molecules *npj Comput. Mater.* **6** 43

[67] Li L, Hoyer S, Pederson R, Sun R, Cubuk E D, Riley P and Burke K 2021 Kohn-Sham equations as regularizer: building prior knowledge into machine-learned physics *Phys. Rev. Lett.* **126** 036401

[68] Li H, Wang Z, Zou N, Ye M, Xu R, Gong X and Duan W 2022 Deep-learning density functional theory Hamiltonian for efficient ab initio electronic-structure calculation *Nat. Comput. Sci.* **2** 367

[69] Pederson R, Kalita B and Burke K 2022 Machine learning and density functional theory *Nat. Rev. Phys.* **4** 357

[70] Gedeon J, Schmidt J, Hodgson M J P, Wetherell J, Benavides-Riveros C L and Marques M A L 2022 Machine learning the derivative discontinuity of density-functional theory *Mach. Learn.: Sci. Technol.* **3** 015011

[71] Baydin A G, Pearlmutter B A, Radul A A and Siskind J M 2018 Automatic differentiation in machine learning: a survey *J. Mach. Learn. Res.* **18** 1 (available at: https://jmlr.org/papers/v18/17-468.html)

[72] Baus M 1984 Broken symmetry and invariance properties of classical fluids *Mol. Phys.* **51** 211

[73] Evans R and Parry A O 1990 Liquids at interfaces: what can a theorist contribute? *J. Phys.: Condens. Matter* **2** SA15

[74] Henderson J R 1992 Statistical Mechanical sum rules, *Fundamentals of Inhomogeneous Fluids*, ed D Henderson (Marcel Dekker)

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

[76] Hermann S and Schmidt M 2022 Why Noether's theorem applies to statistical mechanics *J. Phys.: Condens. Matter* **34** 213001

[77] Hermann S and Schmidt M 2022 Variance of fluctuations from Noether invariance *Commun. Phys.* **5** 276

[78] Hermann S and Schmidt M 2022 Force balance in thermal quantum many-body systems from Noether's theorem *J. Phys. A: Math. Theor.* **55** 464003

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

[80] Sammüller F, Hermann S, de las Heras D and Schmidt M 2023 Noether-constrained correlations in equilibrium liquids *Phys. Rev. Lett.* **130** 268203

[81] Hermann S, Sammüller F and Schmidt M 2024 Noether invariance theory for the equilibrium force structure of soft matter *J. Phys. A: Math. Theor.* **57** 175001

[82] Robitschko S, Sammüller F, Schmidt M and Hermann S 2024 Hyperforce balance from thermal Noether invariance of any observable *Commun. Phys.* **7** 103

[83] Müller J, Hermann S, Sammüller F and Schmidt M 2024 Gauge invariance of equilibrium statistical mechanics *Phys. Rev. Lett.* **133** 217101

[84] Müller J, Sammüller F and Schmidt M, Why gauge invariance applies to statistical mechanics (arXiv:2409.14166)

[85] Rotenberg B 2024 Viewpoint: symmetry spotted in statistical mechanics *Physics* **17** 163

[86] Hirschfelder J O 1960 Classical and quantum mechanical hypervirial theorems *J. Chem. Phys.* **33** 1462

[87] Yvon J 1935 La théorie statistique des fluides et l'équation d'état (in French) *Actualités Scientifiques et Industrielles* (Hermann & Cie.)

[88] Born M and Green H S 1946 A general kinetic theory of liquids I. The molecular distribution functions *Proc. R. Soc.* A **188** 10

[89] Evans R and Stewart M C 2015 The local compressibility of liquids near non-adsorbing substrates: a useful measure of solvophobicity and hydrophobicity? *J. Phys.: Condens. Matter* **27** 194111

[90] Evans R, Stewart M C and Wilding N B 2019 A unified description of hydrophilic and superhydrophobic surfaces in terms of the wetting and drying transitions of liquids *Proc. Natl Acad. Sci.* **116** 23901

[91] Coe M K, Evans R and Wilding N B 2022 Density depletion and enhanced fluctuations in water near hydrophobic

solutes: identifying the underlying physics *Phys. Rev. Lett.* **128** 045501
[92] Coe M K, Evans R and Wilding N B 2022 Measures of fluctuations for a liquid near critical drying *Phys. Rev.* E **105** 044801
[93] Coe M K, Evans R and Wilding N B 2023 Understanding the physics of hydrophobic solvation *J. Chem. Phys.* **158** 034508
[94] Eckert T, Stuhlmüller N C X, Sammüller F and Schmidt M 2020 Fluctuation profiles in inhomogeneous fluids *Phys. Rev. Lett.* **125** 268004
[95] Eckert T, Stuhlmüller N C X, Sammüller F and Schmidt M 2023 Local measures of fluctuations in inhomogeneous liquids: statistical mechanics and illustrative applications *J. Phys.: Condens. Matter* **35** 425102
[96] Wilding N B, Evans R and Turci F 2024 What is the best simulation approach for measuring local density fluctuations near solvo/hydrophobes? *J. Chem. Phys.* **160** 164103
[97] Sammüller F, de las Heras D and Schmidt M 2023 Inhomogeneous steady shear dynamics of a three-body colloidal gel former *J. Chem. Phys.* **158** 054908
[98] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (Oxford University Press)
[99] Anero J G, Español P and Tarazona P 2013 Functional thermo-dynamics: a generalization of dynamic density functional theory to non-isothermal situations *J. Chem. Phys.* **139** 034106
[100] Marconi U M B and Tarazona P 1999 Dynamic density functional theory of fluids *J. Chem. Phys.* **110** 8032
[101] Español P and Löwen H 2009 Derivation of dynamical density functional theory using the projection operator technique *J. Chem. Phys.* **131** 244101
[102] Sammüller F, Hermann S and Schmidt M 2023 Comparative study of force-based classical density functional theory *Phys. Rev.* E **107** 034109
[103] Levy M 1979 Universal variational functionals of electron densities, first-order density matrices and natural spin-orbitals and solution of the v-representability problem *Proc. Natl. Acad. Sci.* **76** 6062
[104] Dwandaru W S B and Schmidt M 2011 Variational principle of classical density functional theory via Levy's constrained search method *Phys. Rev.* E **83** 061133
[105] Brader J M and Schmidt M 2014 Dynamic correlations in Brownian many-body systems *J. Chem. Phys.* **140** 034104
[106] Brader J M and Schmidt M 2013 Nonequilibrium Ornstein-Zernike relation for Brownian many-body dynamics *J. Chem. Phys.* **139** 104108
[107] Mackay E K R, Marbach S, Sprinkle B and Thorneywork A L 2024 The countoscope: measuring self and collective dynamics without trajectories *Phys. Rev.* X **14** 041016
[108] Percus J K 1976 Equilibrium state of a classical fluid of hard rods in an external field *J. Stat. Phys.* **15** 505
[109] Robledo A and Varea C 1981 On the relationship between the density functional formalism and the potential distribution theory for nonuniform fluids *J. Stat. Phys.* **26** 513
[110] Sammüller F, Robitschko S, Hermann S and Schmidt M 2024 Hyperdensity functional theory of soft matter (*Zenodo*) Source code, simulation data, and neural functionalsW (https://doi.org/10.5281/zenodo.13318634)
[111] Roth R, Evans R, Lang A and Kahl G 2002 Fundamental measure theory for hard-sphere mixtures revisited: the White Bear version *J. Phys.: Condens. Matter* **14** 12063
[112] Hansen-Goos H and Roth R 2006 Density functional theory for hard-sphere mixtures: the White Bear version mark II *J. Phys.: Condens. Matter* **18** 8413
[113] Roth R 2010 Fundamental measure theory for hard-sphere mixtures: a review *J. Phys.: Condens. Matter* **22** 063102
[114] Rosenfeld Y 1989 Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing *Phys. Rev. Lett.* **63** 980
[115] Rosenfeld Y 1988 Scaled field particle theory of the structure and the thermodynamics of isotropic hard particle fluids *J. Chem. Phys.* **89** 4272
[116] Schmidt M 1999 Density-functional theory for soft potentials by dimensional crossover *Phys. Rev.* E **60** R6291
[117] Schmidt M 2000 A density functional for additive mixtures *Phys. Rev.* E **62** 3799
[118] Schmidt M 2000 Fluid structure from density functional theory *Phys. Rev.* E **62** 4976
[119] Schmidt M and Jeffrey M R 2007 Peel or coat spheres by convolution, repeatedly *J. Math. Phys.* **48** 123507
[120] Schmidt M 2011 Isometric and metamorphic operations on the space of local fundamental measures *Mol. Phys.* **109** 1253
[121] Pihlajamaa I and Janssen L M C 2024 Comparison of integral equation theories of the liquid state *Phys. Rev.* E **110** 044608
[122] Schmidt M 2011 Statics and dynamics of inhomogeneous liquids via the internal-energy functional *Phys. Rev.* E **84** 051203
[123] Schmid F 2022 Editorial: multiscale simulation methods for soft matter systems *J. Phys.: Condens. Matter* **34** 160401
[124] Brini E, Algaer E A, Ganguly P, Li C, Rodríguez-Ropero F and van der Vegt N F A 2013 Systematic coarse-graining methods for soft matter simulations – a review *Soft Matter* **9** 2108
[125] Delle Site L, Krekeler C, Whittaker J, Agarwal A, Klein R and Höfling F 2019 Molecular dynamics of open systems: construction of a mean-field particle reservoir *Adv. Theory Simul.* **2** 1900014
[126] Baptista L A, Dutta R C, Sevilla M, Heidari M, Potestio R, Kremer K and Cortes-Huerto R 2021 Density-functional-theory approach to the Hamiltonian adaptive resolution simulation method *J. Phys.: Condens. Matter* **33** 184003
[127] Sammüller F and Schmidt M Multivariate hyperdensity functional theory (to be published)