

# Tesi di Laurea Magistrale

in Physics of Complex Systems

Academic Year 2024/2025

## Boltzmann inversion: measuring forces by watching movies

*Author*

Davide Paolino

*Supervisors*

Andrea Pagnani

Ludovic Berthier



Politecnico  
di Torino



Université  
Paris Cité



SORBONNE  
UNIVERSITÉ

université  
PARIS-SACLAY

UNIVERSITÉ  
FRANCO  
ITALIENNE

UNIVERSITÀ  
ITALO  
FRANCESE

UMR 7083  
Griver

# Abstract

Reconstructing effective interaction potentials from structural data is a central inverse problem in statistical physics, particularly relevant for coarse-grained modelling of polymer and colloidal interactions. In this work, we introduce an enhanced version of the Iterative Boltzmann Inversion method for recovering pairwise interaction potentials  $U(r)$  from a target radial distribution function  $g(r)$ . Standard iterative schemes require costly many-body simulations at each step to update structural observables. We overcome this bottleneck by incorporating the Borgis force-based estimator for  $g(r)$ , which allows the inversion procedure to proceed directly from a fixed ensemble of particle configurations. This approach preserves the minimal-assumption philosophy of Iterative Boltzmann Inversion while significantly improving computational efficiency. We validate the method across a diverse set of benchmark potentials, including Lennard-Jones, Weeks-Chandler-Andersen, power-law and shoulder-type interactions, demonstrating accurate reconstructions even under undersampled or experimental-like conditions. The framework is particularly suited for systems where running repeated simulations is too expensive or impractical, such as when working with large colloidal systems or active matter. These results highlight the potential of force-based inversion schemes as practical and robust tools for effective interaction reconstruction.

# Contents

|          |                                                                |           |
|----------|----------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                            | <b>1</b>  |
| 1.1      | Inverse Problems in Statistical Mechanics . . . . .            | 1         |
| 1.2      | Force Inference in Particle Systems . . . . .                  | 1         |
| 1.3      | Existing Inversion Algorithms . . . . .                        | 2         |
| 1.4      | Stop Making New Simulations . . . . .                          | 3         |
| <b>2</b> | <b>Theoretical Framework</b>                                   | <b>5</b>  |
| 2.1      | Pair statistics and the Radial Distribution Function . . . . . | 5         |
| 2.2      | Borgis Formula for the Radial Distribution Function . . . . .  | 6         |
| 2.3      | Iterative Boltzmann Inversion . . . . .                        | 8         |
| 2.4      | Description of the New Algorithm . . . . .                     | 8         |
| <b>3</b> | <b>Results and Validation of the New Algorithm</b>             | <b>10</b> |
| 3.1      | Test Systems and Simulations . . . . .                         | 10        |
| 3.2      | Smoothing the Reference $g(r)$ . . . . .                       | 11        |
| 3.3      | Practical Implementation Details . . . . .                     | 13        |
| 3.4      | Potentials Reconstruction . . . . .                            | 16        |
| <b>4</b> | <b>Conclusion</b>                                              | <b>18</b> |
|          | <b>References</b>                                              | <b>19</b> |
| <b>A</b> | <b>Laplacian Representation of the Dirac delta</b>             | <b>22</b> |
| <b>B</b> | <b>Derivation of Borgis formula</b>                            | <b>23</b> |

## List of Figures

|   |                                                         |    |
|---|---------------------------------------------------------|----|
| 1 | Different estimators for $g(r)$ . . . . .               | 7  |
| 2 | New algorithm schematic . . . . .                       | 9  |
| 3 | Smoothing the reference $g(r)$ . . . . .                | 12 |
| 4 | Iteration example on LJ system . . . . .                | 14 |
| 5 | Reconstruction of LJ and WCA potentials . . . . .       | 15 |
| 6 | Reconstruction of the inverse cubic potential . . . . . | 16 |
| 7 | Reconstruction of the shoulder potential . . . . .      | 17 |

# 1 Introduction

## 1.1 Inverse Problems in Statistical Mechanics

**Inverse Problems** – Learning Nature’s rules directly from observations should be inherently an interesting problem for a scientist, even though it reverses the usual theoretical approach of deriving phenomena from first principles. In statistical mechanics, for instance, the main goal is to predict emerging physical behaviours from a set of given microscopic rules that govern each constituent of the system. However, when facing directly experimental observations, the most natural questions that arise concern what kind of rules, parameters, or interactions are at the root of such phenomenology: this is known as an inverse problem.

**Inverse Problems in Literature** – The field of inverse statistical problems has evolved over five decades [1] with applications in different disciplines. In statistical physics, inferring spin couplings for Ising models has enabled quantitative analysis of complex networks, from gene regulatory interactions to financial market correlations, by mapping them to equivalent spin systems [2]. In biophysics, inverse methods have addressed the challenge of protein design by reverse engineering energy landscapes that reproduce folded structures [3]. Lastly, in living matter physics, these techniques have extracted interaction rules from collective behaviours, like flocking or pattern formation, providing validation for individuals-based models [4].

## 1.2 Force Inference in Particle Systems

**Inverting Interactions** – In physics, the inverse problem often involves finding the proper set of parameters for the system’s Hamiltonian based on empirical data. Choosing an appropriate initial set of observables is itself a non-trivial task for formulating the inverse procedure. While for many years X-ray scattering was a reliable method to extract information about the structure of many-body systems, modern advances in optical microscopy have provided direct access to particle positions, and thus to pair correlation statistics in the form of the radial distribution function [5, 6]. Proving hope that such reconstruction is indeed possible, Henderson’s theorem demonstrates that, for systems governed purely by pairwise interactions at fixed temperature and density, the radial distribution function uniquely determines the inter-particle potential [7]. Thus, knowledge of one should directly grant access to the other.

**Applications** – Extracting effective inter-particle potentials from structural data has broad applications across soft matter, polymer physics, and active mat-

ter systems. In the context of coarse-grained modelling, the goal is to replace a high-dimensional description involving all-atom interactions with a simplified, effective potential that faithfully reproduces the essential physics of the system. By reducing the complexity of molecular interactions while preserving key physical behaviours, coarse-grained potentials enable efficient simulations of large-scale or long-timescale phenomena [8]. In colloidal systems, which are widely used as benchmark platforms for testing predictions in statistical physics. However, engineering an experimental setup that precisely reproduces a desired interaction potential remains a complex and delicate task. An alternative and more direct approach is to measure the interactions present in existing experimental systems [6].

Beyond equilibrium systems, effective potentials also offer crucial insights into non-equilibrium behaviour. For instance, motility-induced phase separation, where active particles spontaneously cluster into a dense phase without explicit attractive interactions, could be better understood in terms of emergent effective attractions [9]. Although such systems are fundamentally out of equilibrium, their steady-state structures can often be effectively captured using equilibrium-like interaction models [10], providing a practical framework for both theoretical interpretation and computational analysis.

### 1.3 Existing Inversion Algorithms

**Different classes** – Over the years, at least three main classes of methods have been developed to address the so-called Henderson problem, namely recovering the pair potential  $U(r)$  from a known radial distribution function  $g(r)$ .

The first class of methods is based on solving the Ornstein–Zernike equation [11]:

$$g(r) - 1 = c(r) + \rho \int c(|\mathbf{r} - \mathbf{r}'|)(g(r') - 1) d\mathbf{r}', \quad (1)$$

where  $c(r)$  is the so-called direct correlation function. To establish a direct link between the radial distribution function and the pair potential, the Ornstein–Zernike equation must be complemented by a closure relation, which expresses  $c(r)$  in terms of  $g(r)$  and  $U(r)$ . Two widely used closures are the Percus–Yevick approximation,

$$c(r) \approx g(r) [1 - e^{U(r)/k_B T}] , \quad (2)$$

and the Hypernetted-Chain approximation,

$$c(r) \approx g(r) - 1 - \ln g(r) e^{U(r)/k_B T} . \quad (3)$$

The biggest advantage of this method is that inversion can be performed analytically, meaning one can write an explicit expression of the form  $U(r) = F[g(r)]$ .

However, these closures are inherently approximate and rely on assumptions that can limit both the accuracy and generality of the inversion procedure [12].

The second group aims to fit the target radial distribution function (RDF) by optimizing a parametrized family of potential functions. While this offers a straightforward solution to the problem, finding numerous applications [10], it depends strictly on the correct interaction parametrization and becomes increasingly less tractable and reliable as the number of fitting parameters grows, particularly when dealing with a variety of diverse interactions [13].

**Simulation-Based Algorithms** – Lastly, a third class of algorithms employs iterative schemes to overcome the limitations of previous approaches [14]. These methods begin with an initial potential estimate that is refined iteratively. At each step, a full many-body simulation must be performed to compute the RDF, after which an update rule minimizes its discrepancy with the reference RDF. This approach has the advantage of not placing any restrictions on the functional form of the potential, allowing any analytic shape to emerge naturally. However, it remains computationally demanding, as new simulations are required at each iteration.

The algorithms in this class differ mainly in the complexity of their update rules [15], which is often inversely related to the number of iterations needed for convergence. For example, Inverse Monte Carlo [16] achieves accurate results within relatively few steps, also matching higher-order statistics. Unfortunately, it requires long forward-simulations to obtain sufficiently converged statistics and a reasonably good initial guess to ensure stable convergence [17].

Iterative Boltzmann Inversion (IBI) is one of the most widely used and straightforward methods in this category, thanks to its simple update rule. However, convergence is found to be much slower than other algorithms, often requiring ten times more iterations [18]. Its robustness has led to widespread application in areas such as liquid interfaces [19], polymer physics [8], and biomolecular modelling [20]. Nonetheless, like all methods in this class, IBI still suffers from the core limitation of requiring costly many-body simulations at each step to obtain updated structural data.

## 1.4 Stop Making New Simulations

**New proposal** – Recently, an alternative approach based on the Test-Particle Insertion (TPI) formula has emerged as a promising solution [21, 22]. The TPI method estimates  $g(r)$  through the expression

$$g_{\text{TPI}}(r) = \frac{\langle \exp -\Psi/k_B T \rangle_r}{\langle \exp -\Psi/k_B T \rangle}, \quad (4)$$

where  $\Psi$  represents the interaction energy associated with inserting a test particle at a given location. The numerator  $\langle \cdot \rangle_r$  denotes a local average over insertions constrained to lie at a fixed distance  $r$  from existing particles, while the denominator accounts for the global average over all possible insertion positions. This approach provides a direct connection between  $g(r)$  and the potential energy landscape, which allows direct access to new structural information without the need for additional simulations. However, TPI faces significant limitations at high densities, where most insertion attempts become energetically unfavourable. Under these conditions, the method requires prohibitively many new particles to achieve accurate estimates, as the majority of insertions contribute negligibly to the averages.

To address these challenges, in this work we propose an alternative strategy that exploits the Borgis formula [23]. Similar to TPI, this approach reconstructs  $g(r)$  directly from a trial potential without full simulation. Nevertheless, it eliminates the need for test particles and remains valid across a broader range of thermodynamic conditions, including high densities where TPI fails. This makes our approach both computationally more efficient and more generally applicable than traditional TPI methods.

**Summary of the report** – The report begins with a concise yet rigorous theoretical foundation on the concept of pair distribution functions (subsection 2.1), followed by a detailed comparison between two approaches to evaluate them: the distance-histogram method and the Borgis formula (subsection 2.2). While the distance-histogram method is often preferred in cases where the interaction potential is completely unknown, it comes at a considerable computational cost for traditional Iterative Boltzmann Inversion schemes. This is because this formula requires expensive molecular dynamics simulations at each step of the iteration to generate updated structural information.

To overcome this limitation, we introduce and implement an enhanced version of the IBI algorithm (subsection 2.3), which directly exploits the knowledge of inter-particle forces using the Borgis formula to compute the radial distribution function. This modification allows the inversion procedure to proceed without running new simulations at each iteration, offering a substantial gain in computational efficiency without compromising accuracy.

To validate the reliability and generality of our algorithm, we applied it to a diverse set of benchmark pair potentials, each presenting unique challenges for inversion, which were simulated using molecular dynamics. The methodology and parameters used to set up these simulations are thoroughly discussed in subsection 3.1, ensuring the reproducibility of our results.

Special attention is paid to mimic realistic, experiment-like conditions in order to ease the application in such contexts: in particular, the target radial distribution function is constructed by smoothing and interpolating data collected from molecular dynamics simulations, improving numerical stability during the inversion process. These details are systematically discussed before we present the final results of our algorithm in [subsection 3.4](#), where we show successful examples of the inversion procedure.

## 2 Theoretical Framework

### 2.1 Pair statistics and the Radial Distribution Function

**The RDF** – A many-body system of  $N$  particles can be completely characterized by the knowledge of all the  $n$ -particle distribution functions  $g_N^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n)$  [11]. The computation of all these quantities for  $N > 2$  is often a nightmare, but fortunately the knowledge of low-order particle distribution functions  $n \leq 2$  is often sufficient to evaluate the equation of state and other thermodynamic properties. In particular, if the system is both isotropic and homogeneous, the local density is constant  $\rho(\mathbf{r}) = \rho$  and the pair distribution function  $g_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$  depends only on the separation  $r = \|\mathbf{r}_2 - \mathbf{r}_1\|$ . In this case, it is usually called radial distribution function (RDF) and referred as  $g(r)$ , taking the form

$$g(r) = \frac{1}{\rho} \left\langle \frac{1}{N} \sum_{i=1}^N \sum_{j \neq i} \delta(r - r_{ij}) \right\rangle, \quad (5)$$

where  $r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$ ,  $\delta(\cdot)$  is the Dirac delta function centered in zero, and  $\langle \dots \rangle$  represents an ensemble average.

**Distance Histogram** – Given the configurations of a homogeneous system, the target  $g_{\text{ref}}(r)$  can be computed using a histogram procedure, that is a simple discretization of [Equation 5](#). In practice, it simply amounts to counting the number of particles  $N(r)$  within a spherical shell of radius  $r$  and thickness  $\Delta r$ , normalized by the shell volume  $\Omega_d r^{d-1} \Delta r$  and the particle density  $\rho = N/V$ , where  $\Omega_d$  is the solid angle in  $d$  dimensions ( $\Omega_2 = 2\pi$  and  $\Omega_3 = 4\pi$ ). Averaging over an ensemble of configurations  $\langle \dots \rangle$  yields

$$g_{DH}(r) = \frac{\langle N(r) \rangle}{\rho \Omega_d r^{d-1} \Delta r}. \quad (6)$$

Note that [Equation 6](#) depends explicitly on the choice of  $\Delta r$ , and the variance of the RDF obtained via this method diverges as  $1/\Delta r^2$ . This approach scales

quadratically with the number of particles  $N^2$  and is ideal for generating  $g_{\text{ref}}(r)$  since it assumes no prior knowledge of the inter-particle potential. However, the RDF evaluated this way is uniquely tied to the potential used to generate the configurations.

## 2.2 Borgis Formula for the Radial Distribution Function

Inspired by previous work on quantum electronic densities, Borgis proposed a new formula to estimate the pair correlation function of a homogenous and isotropic system when the forces acting on each particle are known [23]

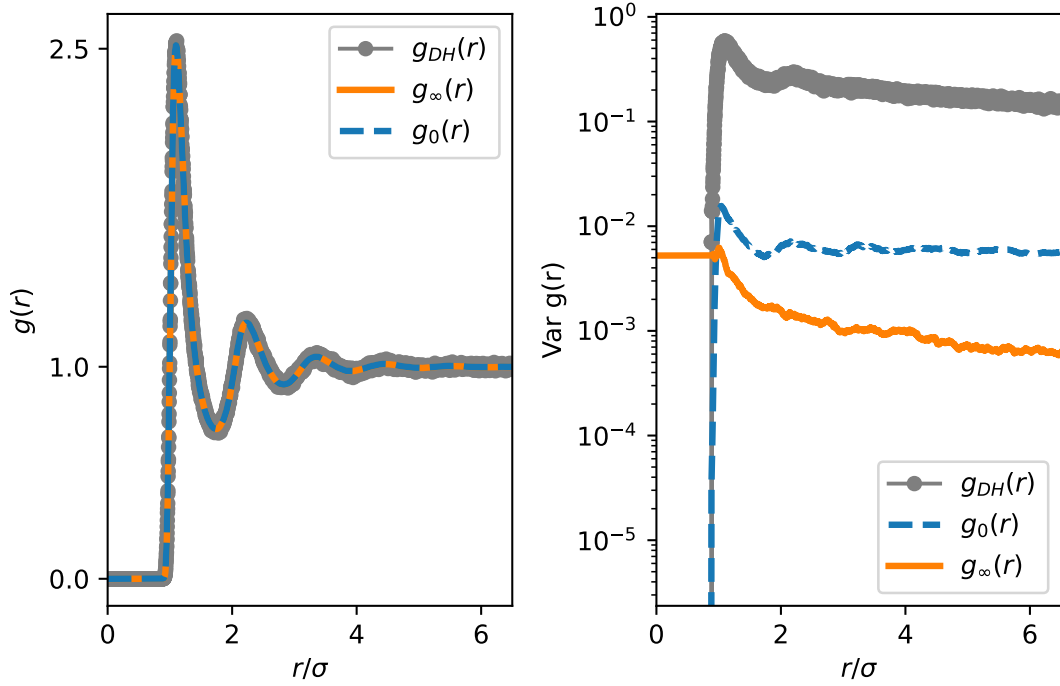
$$g_{\infty}(r) = 1 - \frac{1}{\rho N \Omega_d} \left\langle \sum_{i=1}^N \sum_{j<i} \frac{\mathbf{f}_i - \mathbf{f}_j}{k_B T} \cdot \frac{\mathbf{r}_{ij}}{r_{ij}^d} \Theta(r_{ij} - r) \right\rangle, \quad (7)$$

where  $\Theta(\cdot)$  is the Heaviside theta function. It is important to note that the presence of the total forces  $\mathbf{f}_i$ , acting on each particle  $i$ , is due to the gradient of the Boltzmann weight with respect to the particles positions, coming from an integration by parts, and that no assumption on the pairwise nature of the interaction is required. Moreover, this formula differs from the original work [23] by a factor 2 that was omitted, as already pointed out in [24]. The RDF estimate from Equation 7 matches perfectly the one obtained with the distance-histogram method, as shown in Figure 1.

Equation 7 represents a complete paradigm shift with respect to the histogram procedure described in Equation 6, because each pair  $(i, j)$  contributes now at all distances  $r \leq r_{ij}$ . This helps reducing the variance of the estimate, which now does not depend any more on the shell thickness  $\Delta r$ , meaning that the RDF can be evaluated with arbitrary resolution without fearing the divergent variance. However, the  $g(r)$  obtained using Equation 7 always shows a spurious non-zero value in the "hard-core" region as  $r \rightarrow 0$ , and its variance does not vanish either. This is unlike distance-histogram approaches, where both the mean and variance at small distances correctly go to zero, reflecting the physical exclusion of particles from the hard-core region. In contrast, at low distances Equation 7 includes contributions from all pairs and thus it approaches the correct zero limit only in the case of infinite statistics.

An alternative expression of this formula can be obtained by exchanging the integration boundaries in the integration by parts, leading to

$$g_0(r) = \frac{1}{\rho N \Omega_d} \left\langle \sum_{i=1}^N \sum_{j<i} \frac{\mathbf{f}_i - \mathbf{f}_j}{k_B T} \cdot \frac{\mathbf{r}_{ij}}{r_{ij}^d} \Theta(r - r_{ij}) \right\rangle. \quad (8)$$



**Figure 1:** Radial distribution functions (RDFs) computed using different methods for 500 equilibrium configurations of a Lennard-Jones system ( $\rho\sigma^2 = 0.56$ ,  $k_B T/\epsilon = 1$ ). Histograms were constructed with bin size  $\Delta r = 0.002\sigma$ . The right panel shows the variance of these estimates, demonstrating the non-vanishing asymptotic limit predicted by Equation 7.

This formula maintains the same structure as the previous one, but the Heaviside function now has an opposite sign, meaning that each pair affects all distances  $r \geq r_{ij}$ . Symmetrically to  $g_{\infty}(r)$ ,  $g_0(r)$  displays a vanishing variance at small distances, but its convergence to  $g(r \gg 1) = 1$  is not guaranteed in the case of limited statistics [25]. The three formulas ( $g_0(r)$ ,  $g_{\infty}(r)$ , and  $g_{DH}(r)$ ) for the radial distribution function provide unbiased estimates of the analytical quantity, as evidenced by their convergence to identical results in Figure 1.

The additional features of the Borgis formulas come with the requirement of knowing all the forces that act on the system: This is often an insurmountable request that makes it impossible to use Equation 7 and 8 for systems whose interactions are unknown. However, this aligns perfectly with the IBI framework, where forces are inherently defined through iterative potential updates. Thus, in the IBI context, Borgis formula become both viable and advantageous, bypassing

the need for new simulations at each step of the procedure.

### 2.3 Iterative Boltzmann Inversion

**IBI** – Iterative Boltzmann Inversion (IBI) is a commonly used inversion scheme aimed at reconstructing the potential  $U(r)$  from a target pair distribution function  $g_{\text{ref}}(r)$  obtained experimentally [26]. The procedure consists of updating the potential at each iteration  $t$  according to the rule proposed by Schommers [27]:

$$U_{t+1}(r) = U_t(r) + \alpha k_B T \ln \frac{g_t(r)}{g_{\text{ref}}(r)}, \quad (9)$$

where  $\alpha \in (0, 1]$  is a regularization factor used to control the stability of the procedure. The optimal potential is given by the fixed point  $U^*(r)$  of this iteration, provided that such a fixed point exists.

In the spirit of a minimal assumption algorithm and motivated by relations from statistical mechanics, the initial guess is typically chosen as

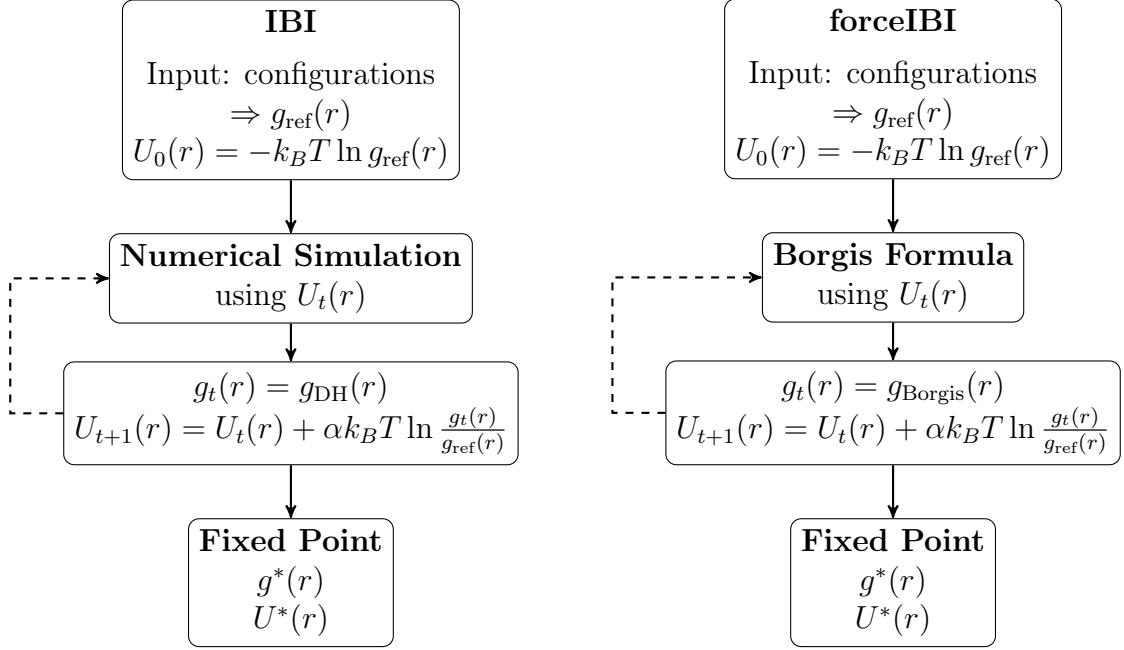
$$U_0(r) = -k_B T \ln g_{\text{ref}}(r), \quad (10)$$

which is often referred to as the potential of the mean force or the Boltzmann inverse of the RDF [28]. This estimate becomes exact in the limit of an infinitely dilute system, but it provides a reasonable starting point also for finite densities. Alternative initial guesses should not effect the final results [26]; in principle, they should not affect the final outcome of the inversion. However, the potential of mean force has the advantage of allowing the shape of the potential to emerge freely, without imposing strong biases toward a specific analytical form.

To apply Equation 9, one needs a protocol to evaluate all pair correlation functions. Although the distance-histogram method provides a direct route to compute  $g_{\text{ref}}(r)$ , unfortunately new ensemble configurations must be produced using the current potential  $U_t(r)$  to use it for each iteration  $t$ . Its iterative dependency on new simulations, which is illustrated in Figure 2, renders the IBI scheme computationally expensive. To circumvent this bottleneck, we propose an alternative approach to evaluate the running  $g_t(r)$ , namely the formula proposed by Borgis and discussed in the previous section.

### 2.4 Description of the New Algorithm

**forceIBI** – Finally, we present our improved iterative algorithm, *forceIBI*. As in the traditional Iterative Boltzmann Inversion (IBI) scheme, the procedure begins



**Figure 2:** Schematic of the standard IBI method (on the left) compared to the improved version that includes the Borgis formula feedback loop (on the right).

with an ensemble of system configurations, from which the target radial distribution function  $g_{\text{ref}}(r)$  is computed using the distance-histogram method presented in Equation 6.

The initial guess for the interaction potential is given by the direct Boltzmann inversion of the target distribution  $U_0(r) = -k_B T \ln g_{\text{ref}}(r)$ .

To avoid the need for costly molecular simulations at each iteration, we replace the traditional step of generating a new  $g_t(r)$  with the histogram method by using the Borgis formula (see Equation 7). We specifically adopt this "outer-integration" version rather than the inner one (see Equation 8) because it does not accumulate many contributions from close-contact pairs, which are prone to undersampling due to the divergence of hard-core potentials. These short-distance pairs must be treated with care, as discussed in detail in subsection 3.3.

The potential is then iteratively updated according to the Schommers scheme (Equation 9) until the algorithm meets the desired convergence criterion.

An overview of the full algorithmic pipeline is provided schematically in Figure 2.

## 3 Results and Validation of the New Algorithm

### 3.1 Test Systems and Simulations

The ultimate goal of our inversion algorithm is to reconstruct the pair potential  $u(r)$  starting from equilibrium configurations of the system. Experimentally, modern advances in confocal microscopy gives access to this information [6], but, before applying our procedure presented in figure [Figure 2](#) to experimental conditions, we want to test it against configurations obtained through numerical simulations, so that we can assess the quality of the inversion by comparing the inverted potential with the one used for the simulation.

**Test Potentials** – We’ll analyze different 2D systems of monodisperse particles of mass  $m$ , interacting with various widely used potentials, each one posing different challenges for the inversion.

The first is the ubiquitous 12 – 6 Lennard-Jones potential [29]

$$U_{LJ}(r) = \epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right], \quad (11)$$

which presents both a repulsive soft core and an attractive long-range behaviour, perfectly modelling van der Waals interactions for fluids with a characteristic energy minimum at  $r_{min} = 2^{1/6}\sigma$ . To limit interactions beyond a certain range, the potential is truncated at a distance  $r_{cut} = 2.5\sigma$  and then shifted so that  $U(r \geq r_{cut}) = 0$ .

The second potential to be tested is the Weeks-Chandler-Anderson (WCA) potential [30], which is nothing but the LJ potential truncated at  $r_{cut} = 2^{1/6}\sigma$ , thus creating a purely repulsive reference potential which is often used in high-density regimes where the role of attractive forces diminishes.

Thirdly, we introduce a long-range power-law potential that scales as the inverse of the cubic distance

$$U_{R3}(r) = \epsilon \left( \frac{\sigma}{r} \right)^3. \quad (12)$$

The  $r^{-3}$  potential describes softer repulsive interactions compared to LJ and WCA, with slower decay relevant for dipole-like or screened electrostatic effects, common in colloidal systems [31]. This potential belongs to the class of inverse power-law interactions  $U^{(n)}(r) = \epsilon \left( \frac{\sigma}{r} \right)^n$ , which does not define separate energy and length scales, and thus its phase diagram depends only on the dimensionless inverse temperature  $\Gamma = (\sigma^d \rho)^{n/d} \cdot \epsilon / k_B T$  [32]. This contrasts with standard liquid-state systems, where temperature and density are treated as independent order parameters.

Lastly, we introduce the shoulder potential [33]

$$U_{sh}(r) = \epsilon \left( \frac{\sigma}{r} \right)^n + \frac{\epsilon}{2} \tanh \left( \frac{k_0}{\sigma} (r - r_0) \right), \quad (13)$$

that features two characteristic length scales, having an hard core  $r^{-n}$  and an outer softer shell represented by the hyperbolic tangent. The parameters are fixed as in the reference work [22], namely  $n = 14$ ,  $k_0 = 10$  and  $r_0 = 2.5\sigma$  with a cut-off of  $r_{cut} = 2.8\sigma$ .

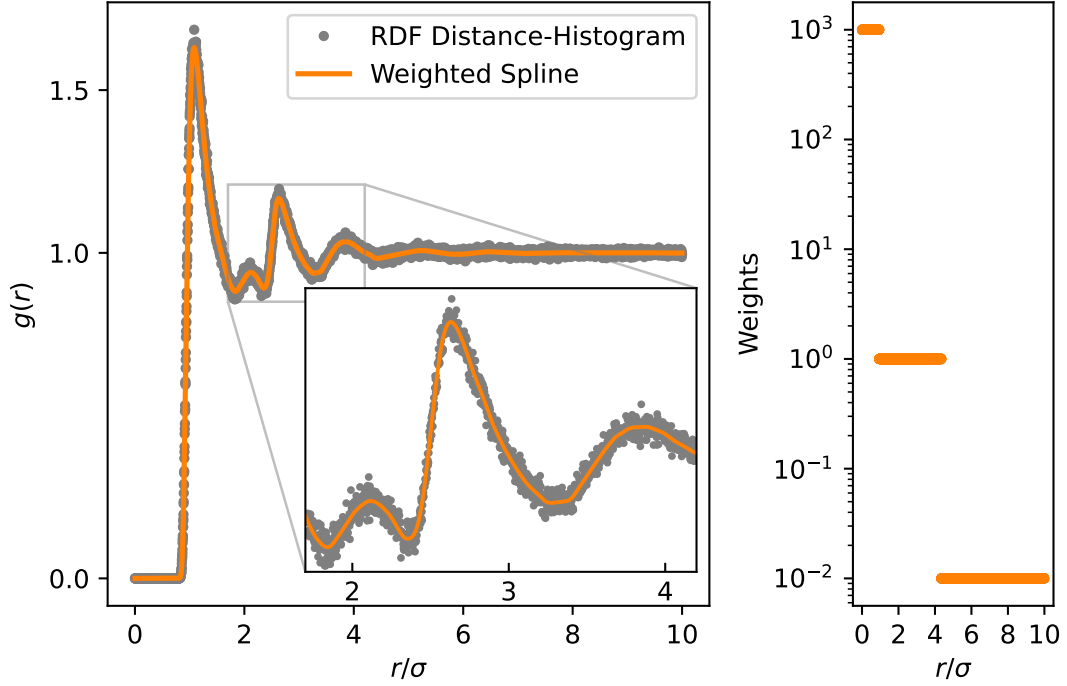
Note that for all the above potential choices, there is a common length scale  $\sigma$ .

**Molecular Dynamics** – The *forceIBI* scheme was tested using mono-atomic system configurations obtained through molecular dynamics simulations. These simulations were performed using the LAMMPS package, a powerful and open-source tool popular for its flexibility [34], which allowed precise control over the input potential for direct comparison with the reconstructed one.

All simulations were performed in a two-dimensional periodic box with lateral size  $L = 60\sigma$ , where  $\sigma$  is the characteristic length scale common to all the potentials discussed above. Particles were initially placed on a square lattice and evolved in the canonical  $NVT$  ensemble using a Nosé–Hoover thermostat. The number of particles  $N$  was selected to achieve densities ranging from approximately  $\rho\sigma^2 = 0.28$  to  $\rho\sigma^2 = 0.56$ , depending on the target potential. Exact values for  $\rho = N/L^2$ , temperature  $T$ , and other simulation parameters are reported in the captions accompanying the inversion results. For each input potential, the target radial distribution function  $g_{\text{ref}}(r)$  was computed from 500 independent snapshots using the distance-histogram method described previously.

### 3.2 Smoothing the Reference $g(r)$

**Smoothing** – The reference radial distribution function  $g_{\text{ref}}(r)$  must be carefully preprocessed to enable accurate potential inversion. A critical consideration is that the resolution of the inverted potential is directly inherited from the input RDF. While finer binning yields smoother potentials, essential for well-behaved force profiles via numerical differentiation, it introduces a trade-off: statistical noise in histogram-based RDFs scales as  $1/(\Delta r)^2$ , potentially creating artificial roughness in the reconstructed potential. Sufficient sampling can mitigate this effect, as larger numbers of MD configurations improve statistics for small bin sizes. However, experimental data often face stricter constraints, like limited measurement time, instrument resolution, or sample stability, restricting the number of usable configurations.



**Figure 3:** Comparison of radial distribution functions (RDFs) computed via direct distance-histogram and weighted spline interpolation methods for a system with shoulder potentials ( $\rho\sigma^2 = 0.56$ ,  $k_B T/\epsilon = 2.0$ ). Results are obtained for 500 configurations using a histogram bin size of  $\Delta r = 0.002\sigma$ . The right panel displays the spline-interpolation weights, highlighting discontinuities at  $r = 0.94\sigma$  and  $r = 4.35\sigma$ .

**Splines** – To balance resolution and noise while preserving key physical features, we apply spline interpolation to the RDF. This approach reduces noise by fitting the histogram data with a continuous function evaluable at any desired resolution. Given a dataset  $\mathcal{D} = \{(r_i, g_i)\}_{i=1}^M$ , the spline is obtained by minimizing the following loss functional

$$\mathcal{L}[f] = \sum_{i=1}^M w_i |g_i - f(r_i)|^2 + \lambda \int \left| \frac{\partial^2 f}{\partial r^2} \right|^2 dr \quad (14)$$

using piecewise cubic polynomials. The first term measures deviation from data points, while the second penalizes curvature, enforcing smoothness of the interpolated function  $f$ . Here  $\lambda$  tunes the trade-off between data fidelity and smoothness. Once the smoothing parameter  $\lambda$  and the weights  $w_i$  are chosen, the spline is constructed using the `scipy.signal` module in Python.

While conceptually easy, applying a single smoothing function across the entire radial distribution function poses various challenges, since the RDF behaves very differently across regions. At short distances, the hard-core region is heavily undersampled and exhibits a sharp rise from zero, which splines may fail to capture without introducing unphysical oscillations. Mid-range RDF peaks suffer from noise in distance-histogram methods, obscuring fine structural features. At long distances, where  $g(r) \approx 1$ , noise dominates, rendering interpolation unreliable due to overfitting. Among these, the short-distance regime is the most challenging and the most important for the reconstruction.

To address these issues systematically, we implement a region-specific weighting scheme  $w_i$  with three distinct zones: the hard-core region (from  $r = 0$  to the midpoint of the first ascent) receives a weight 1000 times stronger than the intermediate region with unitary weighting. Instead, the noisy long-range region where the RDF approaches  $g(r) \approx 1$  is deliberately suppressed with weights reduced to  $1/100$ . Lastly, the smoothing parameter  $\lambda$  is determined automatically through generalized cross-validation. This approach maintains simplicity and transferability across diverse RDF shapes while delivering robust performance for inversion procedures, as demonstrated in [Figure 3s](#) by the comparison between distance-histogram and smoothed spline results.

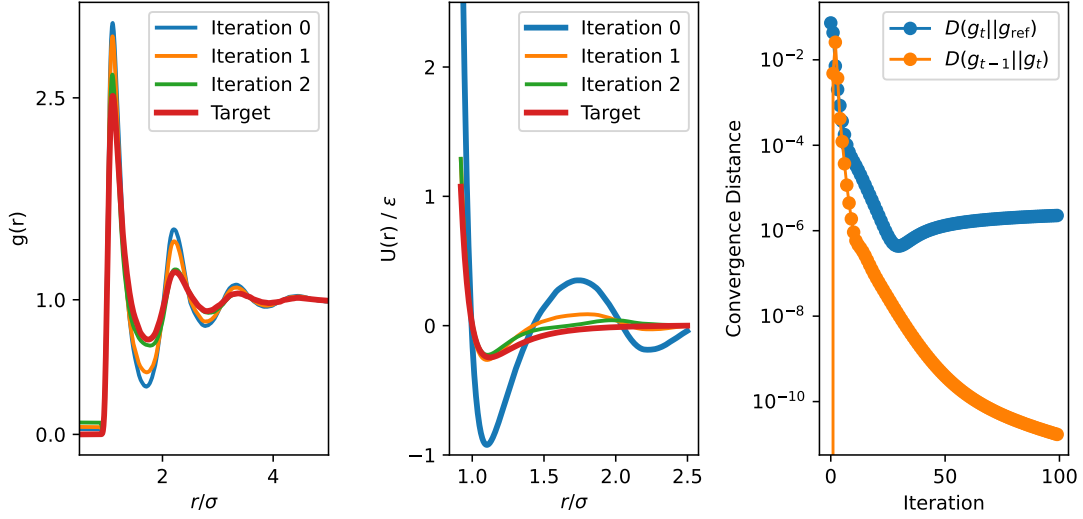
### 3.3 Practical Implementation Details

**Inversion Parameters** – Although one might hope to reconstruct the interaction potential across the full range of distances  $r$ , meaningful results are only achievable within a finite window  $r \in [r_{\text{low}}, r_{\text{inv-cut}}]$ . This limitation arises from fundamental physical and computational constraints tied to these parameters.

The upper cut-off  $r_{\text{inv-cut}}$  truncates the potential where interactions become negligible, reducing unnecessary computational cost. In experimental applications, this cutoff is typically refined iteratively: starting from a high value and progressively lowering it while monitoring potential stability. For consistency with forward simulations, we fix  $r_{\text{inv-cut}} = r_{\text{cut}}$ .

The lower bound  $r_{\text{low}}$  marks the hard-core repulsion scale, defining the distance below which the potential  $U(r) \gg k_B T$  becomes so strongly repulsive that the Boltzmann factor  $\exp[-U(r)/k_B T]$  effectively suppresses configurations with particles at such close separations. This exclusion leads to severe undersampling, making accurate potential reconstruction in this regime practically impossible without infinite statistics.

Although the choice of  $r_{\text{low}}$  is arbitrary due to the unknown true potential, it is the most delicate parameter for the convergence of the inversion procedure. If set too small, undersampling causes numerical instabilities in the radial distribu-



**Figure 4:** The first few iteration steps are shown for a Lennard-Jones potential with density  $\rho\sigma^2 = 0.56$  and  $k_B T/\epsilon = 1.0$ . The starting potential labelled as '0' in the central panel is the potential of the mean force described in Equation 10. On right-most panel, two different convergence metrics are shown.

tion function. If set too large, one must explicitly handle particle pairs  $(i, j)$  with  $r_{ij} < r_{\text{low}}$ . In our implementation, we extend the potential quadratically in this region, which yields stable results unless  $r_{\text{low}}$  is excessively large. To balance these competing requirements, we adopt a pragmatic strategy, as  $r_{\text{low}}$  is incrementally increased until the inversion stabilizes. After extensive testing, we identified the most stable initialization: for each configuration, one first computes the minimum distance between any pair of particles, then sets  $r_{\text{low}}$  to the most frequent value in this distribution of minimum distances.

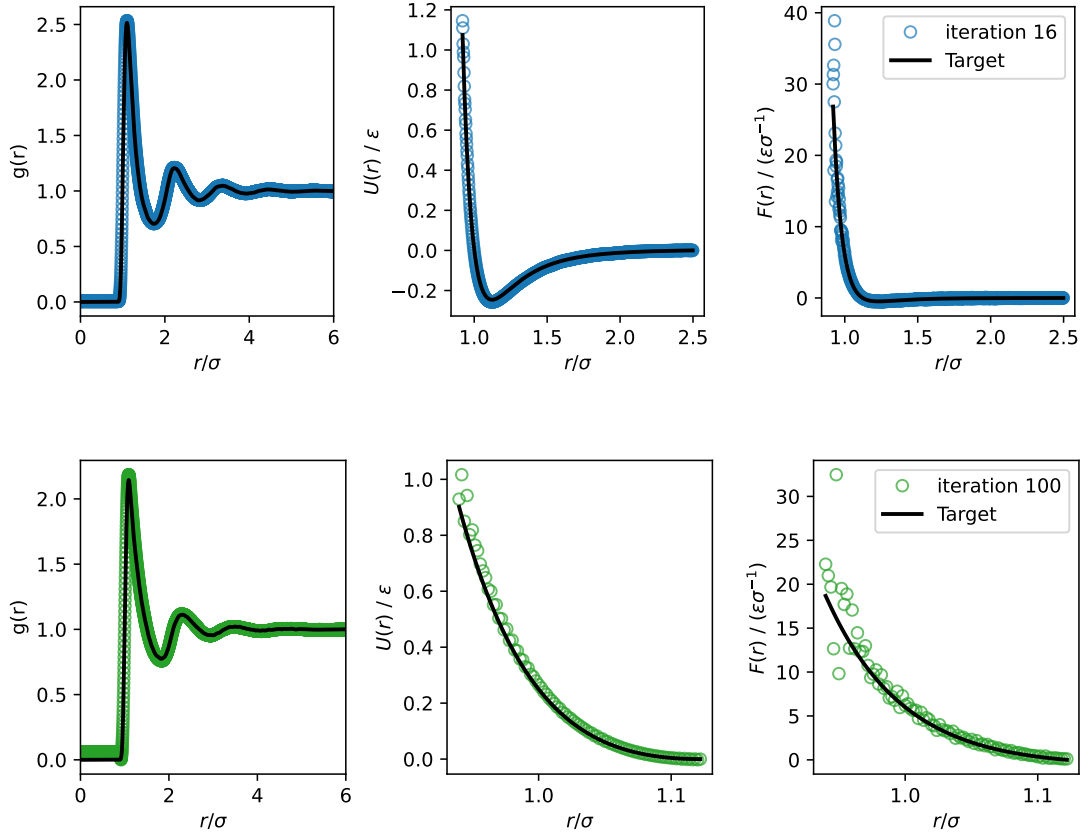
**Convergence Criteria** – Any iterative algorithm requires a criterion to determine when convergence has been reached. Given some distance metric  $D(\cdot||\cdot)$  defined on a (discretized) functions space, the most intuitive approach is to stop the procedure at step  $t^*$  when the current estimate  $g_{t^*}(r)$  is closer to the target than a desired precision  $\Delta$ , namely

$$t^* : D(g_{t^*}||g_{\text{ref}}) \leq \Delta. \quad (15)$$

Instead, the approach we adopt defines convergence as

$$t^* : D(g_{t^*}||g_{t^*-1}) \leq \Delta. \quad (16)$$

This has several advantages: it is more transferable across different pair distribution shapes and does not assume that the distance to the reference decreases



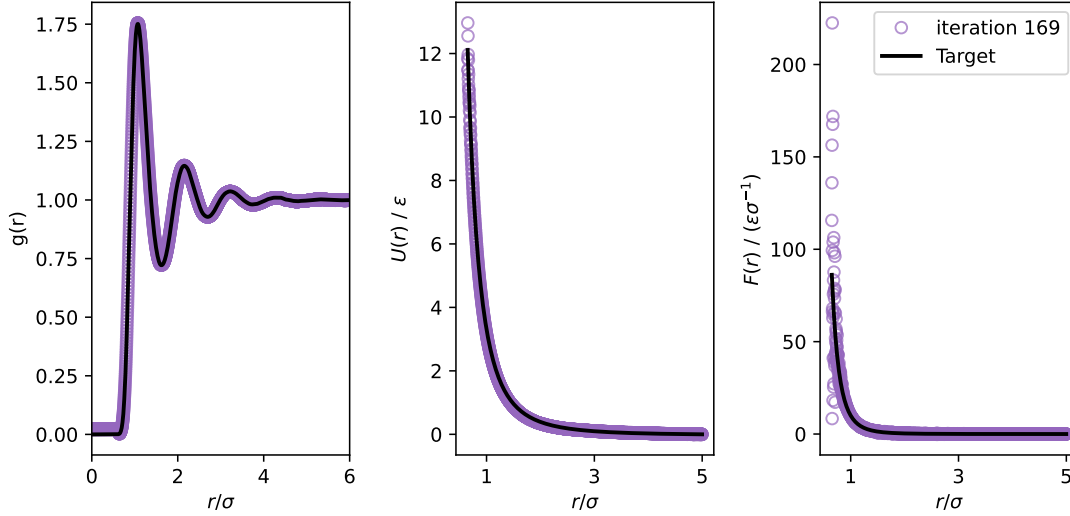
**Figure 5:** Reconstruction results for LJ and WCA potential, in the upper and lower panel respectively. Both systems were simulated using  $\rho\sigma^2 = 0.56$  and  $k_B T/\epsilon = 1.0$ . The reconstruction was performed using  $\alpha = 0.2$  and  $r_{low} = 0.92\sigma$  for both systems.

monotonically. In practice, both criteria yield similar results, as shown in [Figure 4](#). Throughout this work, the chosen metric  $D(\cdot||\cdot)$  is the mean squared error (MSE)

$$D(f||g) = \text{MSE}(f, g) = \frac{1}{r_{low} - r_{inv-cut}} \int_{r_{low}}^{r_{inv-cut}} |f(r) - g(r)|^2 dr \quad (17)$$

computed over the potential reconstruction window  $[r_{low}, r_{inv-cut}]$ .

These same prescriptions could be extended to potential differences, although we emphasize that in experimental scenarios where the true interaction is unknown, the target-based convergence criterion in [Equation 15](#) cannot be applied.



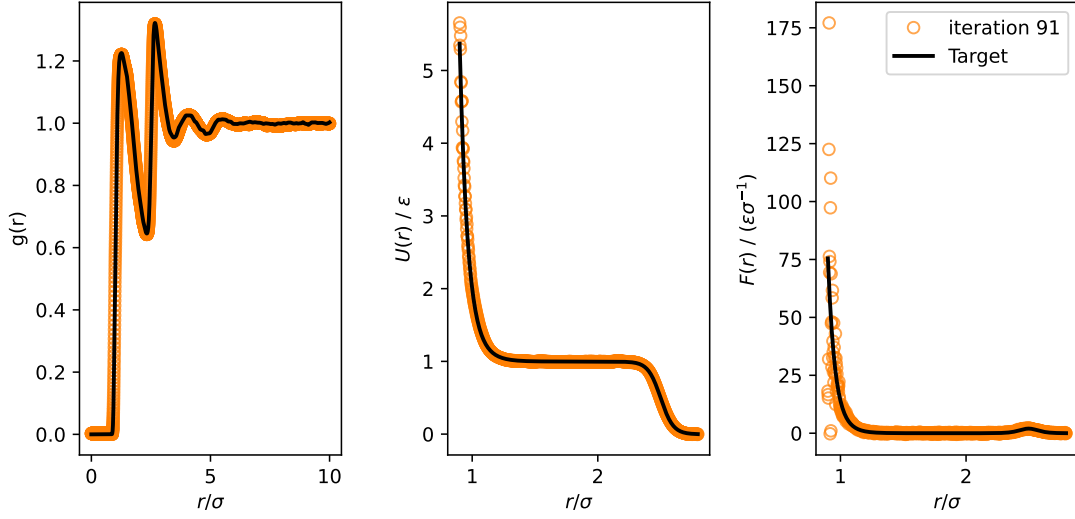
**Figure 6:** Reconstruction results for the  $r^{-3}$ . Simulations of the  $r^{-3}$  were performed using  $\rho\sigma^2 = 0.80$  and  $k_B T/\epsilon = 0.3$ , leading to an reduced inverse temperature  $\Gamma = 0.8^{3/2} \times 0.3^{-1}$ . The reconstruction is performed from  $r_{low} = 0.65\sigma$  using  $\alpha = 0.5$ .

### 3.4 Potentials Reconstruction

To assess the reliability and versatility of our inversion procedure, we tested it on several benchmark potentials introduced in [subsection 3.4](#), using only 125 independent snapshots for the reconstruction out of the 500 employed to build the target. Each potential poses distinct challenges to the inversion process, making them valuable test cases for assessing convergence and reconstruction accuracy.

[Figure 4](#) illustrates the early iterations of the inversion scheme for the Lennard-Jones (LJ) potential, showing the evolution of both the radial distribution function  $g_t(r)$  and the reconstructed potential  $U_t(r)$ . Notably,  $g_t(r)$  rapidly converges toward the reference distribution, despite the initial guess for the interaction potential differs significantly from the target. While the global shape of the potential is largely recovered in the first iteration, residual artifacts, such as spurious double attractive wells, require additional steps to resolve. The rightmost panel of the figure shows the evolution of different convergence metrics, demonstrating that after roughly 50 iterations, the distance to the reference  $g_{ref}(r)$  saturates, while the change between successive iterations falls below one part per million.

The overall agreement between the reconstructed and target potentials is excellent, even for particularly challenging systems. The Lennard-Jones potential



**Figure 7:** Reconstruction results for the shoulder potential, simulated using  $\rho\sigma^2 = 0.28$  and  $k_B T/\epsilon = 1.0$ . The inversion parameters were set to  $\alpha = 0.2$  and  $r_{low} = 0.90\sigma$ .

(Figure 5), in particular, characterized by steep short-range repulsion and a shallow attractive tail, is notoriously difficult to invert, especially because its attractive component plays a minor role in the resulting structure and is easily missed by standard algorithms. In contrast, the WCA potential, which is non-zero over a very narrow region, is accurately resolved, highlighting the robustness of the method in handling short-range interactions. The long-range  $r^{-3}$  potential, often problematic for inverse methods due to its slow decay, is also faithfully reconstructed, as shown in Figure 6. Finally, the shoulder potential (Figure 7), featuring both a steep repulsive core and an intermediate plateau, is successfully recovered, demonstrating the ability of the method to capture potentials with multiple intrinsic length scales.

Across all test cases, potential reconstructions exceed expectations, with discrepancies appearing primarily in the derived force profiles. However, this limitation comes from the numerical differentiation used to compute forces. In applications where the goal is force reconstruction, additional post-processing, such as smoothing or filtering of the underlying potential, can significantly improve the quality of the force field.

## 4 Conclusion

We have presented a robust and flexible inversion framework capable of reconstructing effective pair potentials from radial distribution functions, with strong performance across a range of benchmark interactions. By removing the expensive numerical simulation required at each iterative step, our method drastically reduces the computational cost, exploiting the Borgis formula for the radial distribution function. Nevertheless, it matches the performance of traditional algorithms in recovering potentials, even in challenging regimes with steep repulsions, short interaction ranges, or multiple characteristic length scales.

**Applications** – Beyond its technical success, the method holds significant promise for practical applications. In experimental soft matter systems, such as colloidal suspensions [6], structural data from scattering or microscopy can be combined with this inversion scheme to infer underlying inter-particle interactions, providing otherwise inaccessible insights. Moreover, in active matter systems, where microscopic dynamics violate detailed balance, this approach enables the extraction of effective pairwise interactions that reproduce steady-state structures. Such coarse-grained representations are invaluable for interpreting phenomena like motility-induced phase separation, where emergent clustering can be attributed to effective attractions.

In the future, we intend to apply the developed algorithm directly to these scenarios, leveraging the interdisciplinary environment of our laboratory where such experimental contexts are readily available.

## References

- <sup>1</sup>J. T. Chayes, L. Chayes, and E. H. Lieb, “The inverse problem in classical statistical mechanics”, [Communications in Mathematical Physics](#) **93**, 57–121 (1984).
- <sup>2</sup>H. C. Nguyen, R. Zecchina, and J. Berg, “Inverse statistical problems: from the inverse Ising problem to data science”, [Advances in Physics](#) **66**, 197–261 (2017).
- <sup>3</sup>B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, and D. Baker, “Design of a Novel Globular Protein Fold with Atomic-Level Accuracy”, [Science](#) **302**, 1364–1368 (2003).
- <sup>4</sup>R. Lukeman, Y.-X. Li, and L. Edelstein-Keshet, “Inferring individual rules from collective behavior”, [Proceedings of the National Academy of Sciences](#) **107**, 12576–12580 (2010).
- <sup>5</sup>A. D. Dinsmore, E. R. Weeks, V. Prasad, A. C. Levitt, and D. A. Weitz, “Three-dimensional confocal microscopy of colloids”, [Applied Optics](#) **40**, 4152–4159 (2001).
- <sup>6</sup>C. P. Royall, A. A. Louis, and H. Tanaka, “Measuring colloidal interactions with confocal microscopy”, [The Journal of Chemical Physics](#) **127**, 044507 (2007).
- <sup>7</sup>R. L. Henderson, “A uniqueness theorem for fluid pair correlation functions”, [Physics Letters A](#) **49**, 197–198 (1974).
- <sup>8</sup>G. Maurel, F. Goujon, B. Schnell, and P. Malfreyt, “Multiscale Modeling of the Polymer–Silica Surface Interaction: From Atomistic to Mesoscopic Simulations”, [The Journal of Physical Chemistry C](#) **119**, 4817–4826 (2015).
- <sup>9</sup>A. Brossollet, E. Lempereur, S. Mallat, and G. Biroli, *Effective Energy, Interactions And Out Of Equilibrium Nature Of Scalar Active Matter*, (Dec. 19, 2024) <http://arxiv.org/abs/2412.15175>, pre-published.
- <sup>10</sup>H. Wang and S. Torquato, “Equilibrium states corresponding to targeted hyperuniform nonequilibrium pair statistics”, [Soft Matter](#) **19**, 550–564 (2023).
- <sup>11</sup>J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, Oxford, Jan. 1, 2013).
- <sup>12</sup>J. K. Percus, “Approximation Methods in Classical Statistical Mechanics”, [Physical Review Letters](#) **8**, 462–463 (1962).
- <sup>13</sup>S. Izvekov, M. Parrinello, C. J. Burnham, and G. A. Voth, “Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: a new method for force-matching”, [The Journal of Chemical Physics](#) **120**, 10896–10913 (2004).

- <sup>14</sup>F. Delbary, M. Hanke, and D. Ivanizki, “A generalized Newton iteration for computing the solution of the inverse Henderson problem”, [Inverse Problems in Science and Engineering](#) **28**, 1166–1190 (2020).
- <sup>15</sup>M. P. Bernhardt, M. Hanke, and N. F. A. Van Der Vegt, “Iterative integral equation methods for structural coarse-graining”, [The Journal of Chemical Physics](#) **154**, 084118 (2021).
- <sup>16</sup>A. P. Lyubartsev and A. Laaksonen, “Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach”, [Physical Review E](#) **52**, 3730–3737 (1995).
- <sup>17</sup>T. Murtola, E. Falck, M. Karttunen, and I. Vattulainen, “Coarse-grained model for phospholipid/cholesterol bilayer employing inverse Monte Carlo with thermodynamic constraints”, [The Journal of Chemical Physics](#) **126**, 075101 (2007).
- <sup>18</sup>S. Jain, S. Garde, and S. K. Kumar, “Do Inverse Monte Carlo Algorithms Yield Thermodynamically Consistent Interaction Potentials?”, [Industrial & Engineering Chemistry Research](#) **45**, 5614–5618 (2006).
- <sup>19</sup>M. Jochum, D. Andrienko, K. Kremer, and C. Peter, “Structure-based coarse-graining in liquid slabs”, [The Journal of Chemical Physics](#) **137**, 064102 (2012).
- <sup>20</sup>H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. de Jong, S. M. Gopal, X. Periole, and S. J. Marrink, “The power of coarse graining in biomolecular simulations”, [WIREs Computational Molecular Science](#) **4**, 225–248 (2014).
- <sup>21</sup>A. E. Stones, R. P. A. Dullens, and D. G. A. L. Aarts, “Model-Free Measurement of the Pair Potential in Colloidal Fluids Using Optical Microscopy”, [Physical Review Letters](#) **123**, 098002 (2019).
- <sup>22</sup>C. R. Rees-Zimmerman, J. Martín-Roca, D. Evans, M. A. Miller, D. G. A. L. Aarts, and C. Valeriani, “Numerical methods for unraveling inter-particle potentials in colloidal suspensions: A comparative study for two-dimensional suspensions”, [The Journal of Chemical Physics](#) **162**, 074103 (2025).
- <sup>23</sup>D. Borgis, R. Assaraf, B. Rotenberg, and R. Vuilleumier, “Computation of pair distribution functions and three-dimensional densities with a reduced variance principle”, [Molecular Physics](#) **111**, 3486–3492 (2013).
- <sup>24</sup>A. Purohit, A. J. Schultz, and D. A. Kofke, “Force-sampling methods for density distributions as instances of mapped averaging”, [Molecular Physics](#) **117**, 2822–2829 (2019).
- <sup>25</sup>S. W. Coles, E. Mangaud, D. Frenkel, and B. Rotenberg, “Reduced variance analysis of molecular dynamics simulations by linear combination of estimators”, [The Journal of Chemical Physics](#) **154**, 191101 (2021).

- <sup>26</sup>D. Reith, M. Pütz, and F. Müller-Plathe, “Deriving effective mesoscale potentials from atomistic simulations”, [Journal of Computational Chemistry](#) **24**, 1624–1636 (2003).
- <sup>27</sup>W. Schommers, “A pair potential for liquid rubidium from the pair correlation function”, [Physics Letters A](#) **43**, 157–158 (1973).
- <sup>28</sup>A. K. Soper, “Empirical potential Monte Carlo simulation of fluid structure”, [Chemical Physics](#) **202**, 295–306 (1996).
- <sup>29</sup>J. E. Jones and S. Chapman, “On the determination of molecular fields.—I. From the variation of the viscosity of a gas with temperature”, [Proceedings of the Royal Society of London. Series A](#) **106**, 441–462 (1924).
- <sup>30</sup>J. D. Weeks, D. Chandler, and H. C. Andersen, “Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids”, [The Journal of Chemical Physics](#) **54**, 5237–5247 (1971).
- <sup>31</sup>K. Zahn, R. Lenke, and G. Maret, “Two-Stage Melting of Paramagnetic Colloidal Crystals in Two Dimensions”, [Physical Review Letters](#) **82**, 2721–2724 (1999).
- <sup>32</sup>S. C. Kapfer and W. Krauth, “Two-Dimensional Melting: From Liquid-Hexatic Coexistence to Continuous Transitions”, [Physical Review Letters](#) **114**, 035702 (2015).
- <sup>33</sup>N. V. Gribova, Yu. D. Fomin, D. Frenkel, and V. N. Ryzhov, “Waterlike thermodynamic anomalies in a repulsive-shoulder potential system”, [Physical Review E](#) **79**, 051202 (2009).
- <sup>34</sup>A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in ’t Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, “LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales”, [Computer Physics Communications](#) **271**, 108171 (2022).

## A Laplacian Representation of the Dirac delta

In the study of physics, one often observes that not all equations hold the same significance; some appear more frequently due to their fundamental nature. Consider, for instance, the equation:

$$\mathcal{L}G(\mathbf{r}) = \delta(\mathbf{r}), \quad (18)$$

where  $\delta$  denotes the Dirac delta function and  $G$  is the Green's function associated with the linear operator  $\mathcal{L}$ .

Solving such equations is generally non-trivial, and Green's functions find applications in many areas of physics and mathematics. In the particular case where  $\mathcal{L} = \nabla^2$ , the equation becomes:

$$\nabla \cdot (\nabla G(\mathbf{r})) = \delta(\mathbf{r}). \quad (19)$$

This equation closely resembles Gauss's law in electrostatics, one of Maxwell's equations. If we define the electric field as

$$\mathbf{E}(\mathbf{r}) = \nabla G(\mathbf{r}),$$

then the equation takes the form:

$$\begin{cases} \nabla \cdot \mathbf{E}(\mathbf{r}) = \delta(\mathbf{r}), \\ \mathbf{E}(\mathbf{r}) = \nabla G(\mathbf{r}). \end{cases} \quad (20)$$

This is precisely Maxwell's first equation in the electrostatic case, where the delta function represents a unit point charge located at the origin. Therefore, the Green's function  $G$  can be interpreted as the electric potential generated by such a point charge.

To derive a general expression for the electric field, we can apply the divergence theorem by integrating over a sphere of radius  $r$  centered around the origin.

$$\int_V \nabla \cdot \mathbf{E}(\mathbf{r}) \, dV = \int_{\partial V} \mathbf{E}(\mathbf{r}) \cdot d\mathbf{S}, \quad (21)$$

along with the assumption of rotational symmetry  $\mathbf{E}(\mathbf{r}) = E(r) \hat{\mathbf{r}}$ , where  $r = |\mathbf{r}|$ . The flux through the boundary then becomes:

$$\int_{\partial V} \mathbf{E}(\mathbf{r}) \cdot d\mathbf{S} = E(r) \Omega_d r^{d-1}, \quad (22)$$

where  $\Omega_d$  is the surface area of the unit sphere in  $\mathbb{R}^d$ . Since the divergence of  $\mathbf{E}$  integrates to 1 (due to the delta function), we obtain:

$$E(r) = \frac{1}{\Omega_d r^{d-1}}, \quad \mathbf{E}(\mathbf{r}) = \nabla G(\mathbf{r}). \quad (23)$$

We can now integrate to find the Green's function:

$$G(r) = \begin{cases} -\frac{1}{2}r & \text{for } d = 1, \\ -\frac{1}{2\pi} \log r & \text{for } d = 2, \\ -\frac{1}{4\pi r} & \text{for } d = 3. \end{cases} \quad (24)$$

Substituting this expression back into the original differential equation provides a useful representation of the Dirac delta function

$$\delta(\mathbf{r}) = -\nabla^2 G(r). \quad (25)$$

## B Derivation of Borgis formula

The pair correlation function for a binary mixture of species  $\mu, \nu$  is defined as

$$g_{\mu\nu}(\mathbf{r}) = \epsilon_{\mu\nu} \left\langle \sum_{i=1}^{N_\mu} \sum_{j=1}^{N_\nu'} \delta(\mathbf{r} - \mathbf{r}_{ij}) \right\rangle, \quad \mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j, \quad (26)$$

where the configurational average is represented by

$$\langle \dots \rangle = \frac{1}{Z_N} \int d\{\mathbf{r}_k\} (\dots) e^{-\beta U(\{\mathbf{r}_k\})}, \quad Z_N = \int d\{\mathbf{r}_k\} e^{-\beta U(\{\mathbf{r}_k\})}, \quad (27)$$

while the prefactor

$$\epsilon_{\mu\nu} = \left(1 - \frac{1}{2}\delta_{\mu\nu}\right) \frac{V}{N_\mu N_\nu}, \quad (28)$$

with  $V$  being the system volume and  $N_\mu, N_\nu$  the particle numbers of each species.

Introduce a Green function  $G(\mathbf{x})$  satisfying

$$\Delta_{\mathbf{x}} G(\mathbf{x}) = -\delta(\mathbf{x}). \quad (29)$$

Then

$$\delta(\mathbf{r} - \mathbf{r}_{ij}) = -\Delta_{\mathbf{r}} G(\mathbf{r} - \mathbf{r}_{ij}), \quad (30)$$

and

$$g_{\mu\nu}(\mathbf{r}) = -\epsilon_{\mu\nu} \sum'_{i,j} \frac{1}{Z_N} \int d\{\mathbf{r}_k\} \Delta_{\mathbf{r}} G(\mathbf{r} - \mathbf{r}_{ij}) e^{-\beta U}. \quad (31)$$

Because  $G$  depends on  $\mathbf{r}$  only via  $\mathbf{r} - \mathbf{r}_{ij}$ , the Laplacian may be viewed as acting on  $\mathbf{r}_{ij}$ :

$$\Delta_{\mathbf{r}} G(\mathbf{r} - \mathbf{r}_{ij}) = \Delta_{\mathbf{r}_{ij}} G(\mathbf{r} - \mathbf{r}_{ij}). \quad (32)$$

Fixing all coordinates except  $\mathbf{r}_{ij}$  and applying Green's first identity in the  $\mathbf{r}_{ij}$ -variable with  $\phi = e^{-\beta U}$  and  $\psi = G(\mathbf{r} - \mathbf{r}_{ij})$  yields

$$\int_{\mathbb{R}^d} e^{-\beta U} \Delta_{\mathbf{r}_{ij}} G(\mathbf{r} - \mathbf{r}_{ij}) d\mathbf{r}_{ij} = \int_{\partial B_R} e^{-\beta U} \nabla_{\mathbf{r}_{ij}} G \cdot d\mathbf{S} - \int_{\mathbb{R}^d} \nabla_{\mathbf{r}_{ij}} e^{-\beta U} \cdot \nabla_{\mathbf{r}_{ij}} G d\mathbf{r}_{ij}, \quad (33)$$

where  $B_R$  is a ball of radius  $R \rightarrow \infty$  and the integrals are taken for fixed remaining coordinates; subsequently integrate over the remaining coordinates and divide by  $Z_N$ . The pair correlation function therefore splits into a surface contribution plus a bulk contribution:

$$g_{\mu\nu}(\mathbf{r}) = \epsilon_{\mu\nu} \sum'_{i,j} \left[ S_{ij}(\mathbf{r}) - B_{ij}(\mathbf{r}) \right], \quad (34)$$

with

$$S_{ij}(\mathbf{r}) = \frac{1}{Z_N} \int d\{\mathbf{r}_k\} \left( \lim_{R \rightarrow \infty} \int_{\partial B_R} e^{-\beta U} \nabla_{\mathbf{r}_{ij}} G(\mathbf{r} - \mathbf{r}_{ij}) \cdot d\mathbf{S} \right), \quad (35)$$

$$B_{ij}(\mathbf{r}) = \frac{1}{Z_N} \int d\{\mathbf{r}_k\} \nabla_{\mathbf{r}_{ij}} e^{-\beta U} \cdot \nabla_{\mathbf{r}_{ij}} G(\mathbf{r} - \mathbf{r}_{ij}). \quad (36)$$

Evaluate the surface term  $S_{ij}$ . For physically reasonable, short-ranged potentials one has  $e^{-\beta U} \rightarrow 1$  on  $\partial B_R$  as  $R \rightarrow \infty$ . Use the asymptotic flux of  $\nabla G$ :

$$\lim_{R \rightarrow \infty} \int_{\partial B_R} \nabla_{\mathbf{x}} G(\mathbf{x}) \cdot d\mathbf{S}_{\mathbf{x}} = \frac{1}{\Omega_d} \int_{\mathbb{S}^{d-1}} d\Omega = 1, \quad (37)$$

where  $\Omega_d$  is the surface area of the unit sphere in  $d$  dimensions and we used the standard normalization of  $G$  so that this flux equals 1. Thus for each fixed pair  $(i, j)$  the surface integral contributes 1 after averaging and taking  $R \rightarrow \infty$ , hence

$$\epsilon_{\mu\nu} \sum'_{i,j} S_{ij}(\mathbf{r}) = \epsilon_{\mu\nu} \sum'_{i,j} 1. \quad (38)$$

The combinatorics of the sum together with  $\epsilon_{\mu\nu}$  give

$$\epsilon_{\mu\nu} \sum'_{i,j} 1 = 1, \quad (39)$$

both for  $\mu \neq \nu$  and for  $\mu = \nu$  (the factor  $1 - \frac{1}{2}\delta_{\mu\nu}$  and the restriction  $j \neq i$  ensure the correct counting). Therefore the surface term yields the additive baseline 1 in  $g_{\mu\nu}$ .

We now treat the bulk term. Compute the gradient of the Boltzmann factor:

$$\begin{aligned}\nabla_{\mathbf{r}_{ij}} e^{-\beta U} &= -\beta e^{-\beta U} \nabla_{\mathbf{r}_{ij}} U = -\beta e^{-\beta U} (\nabla_{\mathbf{r}_i} U - \nabla_{\mathbf{r}_j} U) \\ &= -\beta e^{-\beta U} (-\mathbf{f}_i + \mathbf{f}_j) = \beta e^{-\beta U} (\mathbf{f}_i - \mathbf{f}_j),\end{aligned}\quad (40)$$

where  $\mathbf{f}_\ell = -\nabla_{\mathbf{r}_\ell} U$  is the total force on particle  $\ell$ .

Consequently

$$g_{\mu\nu}(\mathbf{r}) = 1 + \epsilon_{\mu\nu} \sum'_{i,j} \frac{1}{Z_N} \int d\{\mathbf{r}_k\} \beta e^{-\beta U} (\mathbf{f}_i - \mathbf{f}_j) \cdot \nabla_{\mathbf{r}_{ij}} G(\mathbf{r} - \mathbf{r}_{ij}). \quad (41)$$

Use the explicit radial form of  $\nabla G$ . In  $d$  dimensions one may write

$$\nabla_{\mathbf{x}} G(\mathbf{x}) = \frac{\mathbf{x}}{\Omega_d |\mathbf{x}|^d}, \quad \mathbf{x} = \mathbf{r} - \mathbf{r}_{ij}, \quad (42)$$

so the bulk term becomes

$$g_{\mu\nu}(\mathbf{r}) = 1 + \frac{\epsilon_{\mu\nu}}{\Omega_d} \left\langle \sum'_{i,j} \beta \frac{\mathbf{r} - \mathbf{r}_{ij}}{|\mathbf{r} - \mathbf{r}_{ij}|^d} \cdot (\mathbf{f}_i - \mathbf{f}_j) \right\rangle. \quad (43)$$

We now perform the angular average to obtain the scalar radial correlation  $g_{\mu\nu}(r)$  which depends only on  $r = |\mathbf{r}|$ . Define the angular average over the direction  $\hat{\mathbf{r}}$  as

$$g_{\mu\nu}(r) = \frac{1}{\Omega_d} \int_{\mathbb{S}^{d-1}} g_{\mu\nu}(r\hat{\mathbf{r}}) d\Omega_{\hat{\mathbf{r}}}. \quad (44)$$

Insert the bulk term and exchange the order of the  $\hat{\mathbf{r}}$ -integral and the configurational average. Assuming that the forces  $\mathbf{f}_i$  are independent of  $\hat{\mathbf{r}}$ , namely that the system is isotropic, we need to evaluate the angular integral

$$\frac{1}{\Omega_d} \int_{\mathbb{S}^{d-1}} \frac{\mathbf{r} - \mathbf{r}_{ij}}{|\mathbf{r} - \mathbf{r}_{ij}|^d} d\Omega_{\hat{\mathbf{r}}} = \frac{\mathbf{r}_{ij}}{r_{ij}^d} \Theta(r - r_{ij}), \quad (45)$$

where  $\Theta$  is the Heaviside step function. This identity follows from the fact that the integral vanishes for  $r < r_{ij}$  by symmetry, while for  $r > r_{ij}$  it equals the flux of  $\nabla G$  through a sphere of radius  $r$  centered at the origin, which is independent of  $r$  and equals 1 in the direction of  $\mathbf{r}_{ij}$ . Thus, putting all together, under the isotropy assumption we arrive at Borgis' formula:

$$g_{\mu\nu}(r) = 1 + \frac{\epsilon_{\mu\nu}}{\Omega_d} \left\langle \sum'_{i,j} \frac{\mathbf{r}_{ij}}{r_{ij}^d} \cdot \beta (\mathbf{f}_i - \mathbf{f}_j) \Theta(r - r_{ij}) \right\rangle. \quad (46)$$

It's worth noting that the derivation did not require the forces to be pairwise additive, only that they are derivable from a potential energy function  $U$ . Moreover, since  $g(r \rightarrow \infty) = 1$ , we obtain that

$$\left\langle \sum'_{i,j} \frac{\mathbf{r}_{ij}}{r_{ij}^d} \cdot \beta(\mathbf{f}_i - \mathbf{f}_j) \right\rangle = 0. \quad (47)$$

Then by using the identity  $\Theta(r - r_{ij}) = 1 - \Theta(r_{ij} - r)$ , we can rewrite Borgis' formula in an alternative form:

$$g_{\mu\nu}(r) = -\frac{\epsilon_{\mu\nu}}{\Omega_d} \left\langle \sum'_{i,j} \frac{\mathbf{r}_{ij}}{r_{ij}^d} \cdot \beta(\mathbf{f}_i - \mathbf{f}_j) \Theta(r_{ij} - r) \right\rangle. \quad (48)$$