

Molecular Physics

An International Journal at the Interface Between Chemistry and Physics

ISSN: 0026-8976 (Print) 1362-3028 (Online) Journal homepage: <https://www.tandfonline.com/loi/tmph20>

Alchemical molecular dynamics for inverse design

Pengji Zhou, James C. Proctor, Greg van Anders & Sharon C. Glotzer

To cite this article: Pengji Zhou, James C. Proctor, Greg van Anders & Sharon C. Glotzer (2019) Alchemical molecular dynamics for inverse design, Molecular Physics, 117:23-24, 3968-3980, DOI: [10.1080/00268976.2019.1680886](https://doi.org/10.1080/00268976.2019.1680886)

To link to this article: <https://doi.org/10.1080/00268976.2019.1680886>



View supplementary material 



Published online: 23 Oct 2019.



Submit your article to this journal 



Article views: 89



View related articles 



View Crossmark data 

Alchemical molecular dynamics for inverse design

Pengji Zhou^{a,*}, James C. Proctor^{b,*}, Greg van Anders^{a,c,d} and Sharon C. Glotzer^{a,b,c,e}

^aDepartment of Chemical Engineering, University of Michigan, Ann Arbor, MI, USA; ^bDepartment of Materials Science and Engineering, University of Michigan, Ann Arbor, MI, USA; ^cDepartment of Physics, University of Michigan, Ann Arbor, MI, USA; ^dDepartment of Physics, Engineering Physics, and Astronomy, Queen's University, Kingston, Ontario, Canada; ^eBioInterfaces Institute, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

We present a molecular dynamics (MD) implementation of an extended statistical mechanical ensemble that includes ‘alchemical’ degrees of freedom describing particle attributes as thermodynamic variables. We demonstrate the use of this alchemical MD method in inverse design simulations of particles interacting via the Oscillating Pair Potential (OPP) and the Lennard–Jones–Gauss potential (LJG) – two general, previously studied models for which phase diagrams are known. We show that alchemical MD can quickly and efficiently optimise pair potentials for target structures within a specified design space in the low-temperature regime, where internal energy adequately represents the features of the alchemical free energy landscape. We show that alchemical MD can be also used to inversely design pair potentials to achieve target materials properties (here, bulk modulus) directly, without explicit knowledge of the structure–property relationship. Alchemical MD can easily be generalised and applied to any target materials properties or structures and used with any differentiable interaction potential.

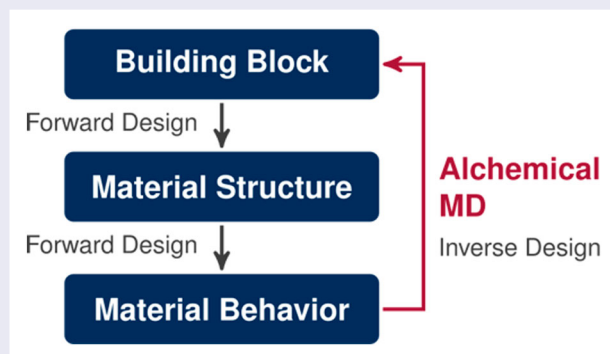
ARTICLE HISTORY

Received 1 June 2019

Accepted 30 September 2019

KEYWORDS

Digital alchemy; materials design; inverse design; self-assembly; molecular dynamics



1. Introduction

A holy grail of materials science is the ability to ‘inversely design’ a material with any arbitrary set of properties or behaviours [1–3]. This approach contrasts with traditional ‘forward’ approaches that start by considering atomic, molecular or nanoscale building blocks, then predict thermodynamic structures from building block attributes, and subsequently link structure to materials properties [4]. Several recent investigations have introduced techniques for inverse design that yield building blocks optimised for given target structures [5–14].

Here, we present an inverse design method for molecular dynamics simulation based on digital alchemy [11] that can be directly used to inversely design structures and properties (see Figure 1 for an illustration).

Digital alchemy was first introduced as an alchemical Monte Carlo (MC) algorithm that performed thermodynamic optimisation of hard nanoparticle shapes for a target colloidal crystal structure in an extended thermodynamic ensemble where particle shape is treated as a fluctuating thermodynamic variable [11,14,15]. Alchemical MC is straightforward to implement; in addition to

CONTACT Sharon C. Glotzer  sglotzer@umich.edu

*P. Z. and J. C. P. contributed equally to this work.

 Supplemental data for this article can be accessed here. <https://doi.org/10.1080/00268976.2019.1680886>

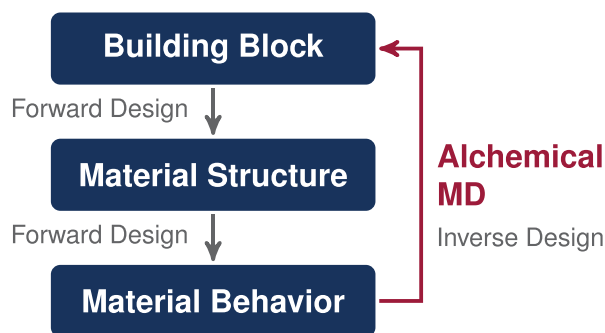


Figure 1. Traditional, forward, design starts from a material's basic atomic, molecular or nanoscale building blocks, then determines how those building blocks organise into a structure, and then how that structure determines material properties. We implement an 'Alchemical' Molecular Dynamics algorithm gives the 'inverse' design of building blocks that form a material with a target property.

particle translation and rotation moves, shape moves are attempted (by randomly proposing a change in some shape parameter α), and the newly generated microstate similarly accepted or rejected by comparison with the Boltzmann probability. When the conjugate 'alchemical potential' μ_α is set to zero, the Boltzmann probability in the $NVT\mu_\alpha$ ensemble is either one or zero, just as it is for an NVT ensemble of hard particles. Indeed, for hard particles, the only criterion for accepting any attempted change of microstate is that particles cannot overlap. Beyond hard shape, inclusion of interparticle interactions in alchemical MC is similarly straightforward.

Alchemical molecular dynamics is trickier, requiring the derivation and implementation of an 'alchemostat' allowing simulations in a fixed alchemical-potential ensemble where particle attributes are allowed to fluctuate, in analogy with a thermostat or barostat used for simulations in the NVT or NPT thermodynamic ensembles. Here we present an implementation of alchemical MD (Alch-MD) and derive finite difference approximations of the equations of motion for interacting particles in contact with an alchemical 'bath', or alchemostat, in analogy with the Nosé-Hoover thermostat [16].

We then perform MD simulations in the $NVT\mu_\alpha$ alchemical ensemble and demonstrate how this method can be used to optimise an interaction potential both for target structures and for target properties. As a simple but instructional example problem, we target bulk modulus, a measure of mechanical response and inversely design a pair-wise interaction potential that can effectively tune the bulk modulus of a crystal. We validate this prediction via direct, brute-force computations. We show that a few Alch-MD simulations can produce important design information previously obtained via many simulations in a grid sweep with fixed parameters. The Alch-MD

method is extensible and generalisable to any system of particles that can be modelled by interaction potentials.

And while this work focuses on a-priori investigations of particle interaction in simulation, thus the phrase 'digital alchemy', the methodology can accommodate parallel experimental research where such alchemical changes (or 'colloidal alchemy') are physically achievable, e.g. in colloidal systems where different manipulation of colloidal shape or size and effective interaction are possible [17,18]. In such systems, μ_α , the thermodynamic conjugate of the alchemical parameter corresponds physically to the energy cost of changing the effective interaction and the resulting driving force in alchemical space. For most physical systems, we don't expect this cost to be zero, and we present in this work different examples where μ_α is zero or nonzero and showcase the different applications with different μ_α .

2. Model and methods

We implement a Nosé-Hoover type thermostat which incorporates alchemical variables in the particle-based molecular simulation toolkit HOOMD-blue [19,20]. We describe the derivation of such Alch-MD method and the simulation protocols we used for inverse design with Alch-MD in the Results section. Below we describe the two families of interaction potentials used as example model systems in this work, and the methods for bulk modulus calculation and structure identification, which are used for inverse design examples studied in this paper.

2.1. Models

While the Alch-MD framework is not limited to isotropic pair potentials, their computational efficiency and ubiquity make them an ideal candidate for the demonstration of the utility of Alch-MD. Colloids and nanoparticles with isotropic interactions have been demonstrated to manifest a broad range of interesting behaviours in experiments [21–24]. The ability to control the synthesis [25] and interactions [23,26] of those particles is increasing steadily. The development of methods for materials property design in systems with isotropic interactions is therefore crucial to guiding future synthesis. To illustrate the application of Alch-MD, we use two isotropic pair potential families that have been shown to have rich phase diagrams of diverse crystal structures in alchemical space.

2.1.1. Oscillating pair potential

One model suitable for testing our inverse design approach is provided by the oscillating pair potential

(OPP) of the form

$$V_{\text{OPP}}(r) = \frac{1}{r^{15}} + \frac{1}{r^3} \cos(k(r-1) + \phi), \quad (1)$$

where the parameters k and ϕ control the frequency of the oscillation and the phase shift, respectively. Although Equation (1) was originally proposed in the context of metals [27], these potentials have more recently been used in exhaustive structural investigations that discovered a multiplicity of complex crystals including icosahedral quasicrystals [28,29]. Prior thorough investigations of the OPP family facilitate the validation of our inverse design approach.

We have fixed units of energy by setting the overall multiplicative constant for this potential to one. To properly smooth the potential near the computational cut-off radius, the XPLOD cut-off scheme is applied in HOOMD to the OPP potential such that the actual potential becomes $V(r) \cdot S(r)$:

$$S(r) = \frac{(r_{\text{cut}}^2 - r^2)^2 \cdot (r_{\text{cut}}^2 + 2r^2 - 3r_{\text{on}}^2)}{(r_{\text{cut}}^2 - r_{\text{on}}^2)^3} (r_{\text{on}} \leq r \leq r_{\text{cut}}). \quad (2)$$

Here $r_{\text{on}} = 2.5$, $r_{\text{cut}} = 3.0$, r_{on} is the radius where the tail smoothing XPLOD function $S(r)$ starts to apply, and r_{cut} is the radius where the potential and its first derivative are set to zero.

2.1.2. Lennard–Jones–Gauss

Another suitable interaction potential is the Lennard–Jones–Gauss (LJG) potential, which has been shown to self-assemble a variety of 2D structures including quasicrystals [30]. The LJG potential, $V_{\text{LJG}}(r)$, is a double-well potential comprised additively of a simplified Lennard–Jones potential and an attractive Gaussian:

$$V_{\text{LJG}}(r) = \left(\frac{1}{r^{12}} - \frac{2}{r^6} - \epsilon e^{-(r-r_0)^2/2\sigma^2} \right) \epsilon_{\text{norm}}, \quad (3)$$

where r_0 and ϵ correspond to the location of the second well and its additive depth, respectively. σ^2 determines the width of that well and, following previous studies, is fixed at a value of 0.02. The variable ϵ_{norm} sets the energy units of the interaction potential. Setting $\epsilon_{\text{norm}} = \text{constant}$ would, due to the nature of the parameterisation, result in a trivial optimisation solution for all systems to minimise their free energy by maximising the depth of the second well. In Alch-MD, this would appear as an ever increasing ϵ value. Thus, to create a both reasonable and differentiable energy scaling, the energy normalisation is defined as the following functions of the other variables to ensure that the attractive portion of the

potential always has unit area:

$$\epsilon_{\text{norm}} = \frac{-1}{\int_{r_{\text{min}}}^{\infty} V_{\text{LJG}}|_{\epsilon_{\text{norm}}=1} dr}. \quad (4)$$

If we set the point at which the standard Lennard–Jones potential first becomes attractive, $r_{\text{min}} = 2^{(-1/6)}$, we can then define ϵ_{norm} simply as

$$\epsilon_{\text{norm}} = \left(\sqrt{\frac{\pi}{2}} \sigma \epsilon \left(\text{erf} \left(\frac{r_0 - 2^{-1/6}}{\sqrt{2}\sigma} \right) + 1 \right) + 2^{5/6} \frac{12}{55} \right)^{-1}. \quad (5)$$

2.2. Bulk modulus quantification

We used the inverse design of bulk modulus as an example for inverse property design using Alch-MD method. To calculate bulk modulus, we used the virial stress and pressure measurements implemented in HOOMD-blue [19,20]. The virial stress calculation is given by [31]

$$\tau_{ij} = \frac{1}{V} \sum_{k \in V}^N \left(-m^{(k)} (u_i^{(k)} - \bar{u}_i) (u_j^{(k)} - \bar{u}_j) + \frac{1}{2} \sum_{l=1}^N (x_i^{(l)} - x_i^{(k)}) f_j^{(kl)} \right), \quad (6)$$

where k and l denote different particles in the system, V is the volume of the system, $m^{(k)}$ is the mass of the particle k , $u_i^{(k)}$ is the i th component of the velocity of particle k , $\bar{u}_j$ is the j th component of the average velocity of particles in the volume, $x_i^{(k)}$ is the i th component of the position of particle k and $f_j^{(kl)}$ is the j th component of the force applied on particle k by particle l . The bulk modulus is obtained from the trace of the change in τ in response to an infinitesimal strain.

In an isotropic system at equilibrium the pressure is given by [32]

$$P = -\frac{1}{3} \text{Tr}(\tau), \quad (7)$$

and the bulk modulus is

$$B = -V \frac{dP}{dV}. \quad (8)$$

We computed the virial stress for 255 self-assembled nanocrystals produced via brute force non-AlchMD MD simulations following the protocol described in Simulation Protocol subsection, step 1. We found that the anisotropy of the bulk modulus, as indicated by the distribution of diagonal elements of τ , remained below 5% for the majority (>99%) of the self-assembled crystals studied in this work. Therefore – and for computational

efficiency – we approximated the nanocrystals as having an isotropic modulus and used Equation (8) to evaluate the property bias in our implementation of Equation (17). We used a second-order finite difference scheme with a 0.5% isotropic contraction and expansion:

$$\frac{dP}{dV} = \frac{P(V + dV) - P(V - dV)}{2dV}. \quad (9)$$

2.3. Structure identification methods

To identify the many self-assembled structures produced in this study, we used a machine learning method described in [33] to automate the task. Briefly, this method uses spherical harmonics $Y_l^m(\theta, \phi)$ as a numerical description of the neighbouring particle bonds to describe the environment of each particle. A maximum of 10 nearest neighbours and a maximum spherical harmonic degree, l , of 10 were used to generate high dimensional local environment descriptors. We used a set of 18 thermalised crystal structures from [33] (cP1-Po, cI2-W, cF4-Cu, cP4-Li, cP8-Cr3Si, cI16-Si, cP20-Mn, cP54-K4Si23, tI4-Sn, tP30-CrFe, hP1-Ca0.15Sn0.85, hP2-Mg, hP2-X, tI20-X, oP52-X, hP2-X, hP10-X, hP18-X), and used Pythia [33,34] to train a feed-forward artificial neural network with one hidden layer to predict the local structure around a particle based on its local environment descriptor using the Python library Keras [35]. The final reported structure of a nanocrystal was chosen as the majority vote of structures identified for each particle in the same nanocrystal. We also used Bond Orientational Order Diagrams (BOODs) [36] as a visual fingerprint to visualise the structural change of the system.

For the 2D LJG simulations, local environment descriptors were generated using Pythia with a maximum of 6–10 nearest neighbours. For both the distance and angle calculations, a singular value representation of the generated matrix was used for invariance to neighbour order [34]. A stochastic gradient descent algorithm was used to optimise a logistic regression [37] using labelled structures following naming conventions of Engel et al [30].

3. Results

Here we present a derivation of the alchemostat and Alch-MD equations of motions used in this work. We then demonstrate the use of Alch-MD in the inverse design of targeted 3D and 2D structures for the OPP and LJG systems, respectively, and targeted bulk modulus for the OPP system.

3.1. Derivation of Alch-MD equations of motion

To perform digital alchemy MD simulations in the alchemically extended canonical $NVT\mu_\alpha$ ensemble, we develop in this section an alchemostat – a type of thermostat that accounts for the fluctuating alchemical variables of this extended ensemble. We follow the derivation of the Nosé–Hoover thermostat as presented by Frenkel and Smit [16].

We begin by defining the alchemical Hamiltonian, H_A , of a system of point particles and a single alchemical variable as

$$H_A = \sum_i \frac{p_i^2}{2m_i s^2} + U(q, \alpha) + \frac{p_s^2}{2Q} + \frac{L}{\beta} \ln s + \frac{p_{s_\alpha}^2}{2Q_\alpha} + \frac{p_\alpha^2}{2M s_\alpha^2} + \mu_\alpha N \alpha + \frac{L_\alpha}{\beta_\alpha} \ln s_\alpha. \quad (10)$$

In this expression, i indexes particles; p_i , q_i and m_i are particle momenta, positions and masses, respectively; U is any defined energy function for the whole system as a function of q and α ; μ_α is the alchemical potential thermodynamically conjugate to α ; p_s and Q are the momentum and mass, respectively, of the Nosé–Hoover thermostat; s is the Nosé–Hoover coordinate [16]; L is the number of canonical degrees of freedom (dimensionality of the system times the number of particles plus one introduced by the Nosé–Hoover thermostat); p_α and M are the momenta and mass, respectively, associated with the alchemical variable; s_α is the Nosé–Hoover coordinate for the alchemostat; p_{s_α} and Q_α are the corresponding momentum and mass, respectively, of the alchemostat, and L_α is the number of alchemical degrees of freedom (equal to the number of the alchemical variables plus one).

The parameter X is introduced, defined by $T_\alpha \equiv XT$ (or equivalently $\beta_\alpha \equiv \beta/X$), so that the ratio of energy in the alchemical and canonical degrees of freedom can be set, where X is typically some relatively small value, i.e. $X \ll 1$, in this work $X = 10^{-4}$.

Because the current systems of interest are those with constant interactions, the resulting design does not include dynamics for the alchemical variables. Without these dynamics, many alchemical variables, such as Q_α and M , would not have a corresponding observable and therefore convenient values can be imposed. However, when modelling dynamic interactions in an experimental system, appropriate tuning of these variables would be critical to matching the dynamics observed.

The equations of motion for a system defined by the Hamiltonian in Equation (10) are

$$\begin{aligned}
 \dot{q} &= \frac{p}{ms^2} \\
 \dot{p} &= -\frac{\partial U(q, \alpha)}{\partial q} \\
 \dot{s} &= \frac{p_s}{Q} \\
 \dot{p}_s &= \frac{1}{s} \left(\frac{p^2}{ms^2} - \frac{L}{\beta} \right) \\
 \dot{\alpha} &= \frac{p_\alpha}{Ms_\alpha^2} \\
 \dot{p}_\alpha &= -\mu_\alpha N - \frac{\partial U(q, \alpha)}{\partial \alpha} \\
 \dot{s}_\alpha &= \frac{p_{s_\alpha}}{Q_\alpha} \\
 \dot{p}_{s_\alpha} &= \frac{1}{s_\alpha} \left(\frac{p_\alpha^2}{Ms_\alpha^2} - \frac{L_\alpha}{\beta_\alpha} \right).
 \end{aligned} \tag{11}$$

To allow for the simulation of canonical and alchemical degrees of freedom on separate timescales, following Martyna and Klein [38], a Trotter factorisation is used to derive the finite difference equations representing the desired equations of motion. These intermediate steps are detailed in the supporting information document. We also simplify our notation by defining $\Delta t \equiv n\delta t$, where δt is the canonical time step in the MD simulation and the alchemical timestep, Δt can be any constant factor n of δt . Following the Nosé–Hoover scheme, we define the relation $\xi \equiv s'p'_s/Q$ for both thermostat and ‘alchemostat’ variables.

Incorporating the virtual variables for both the canonical and alchemical degrees of freedom with the traditional Nosé–Hoover conventions (i.e. $t' \equiv t/s$, $p' \equiv p/s$, $q' \equiv q$ and $s' \equiv s$) we numerically integrate the equations of motion of the in Alch-MD system as follows. The first step in the integration algorithm is to update the alchemical variable and its corresponding momentum:

$$\begin{aligned}
 \alpha' \left(\frac{\Delta t}{2} \right) &= \alpha'(0) + \frac{\Delta t}{2} \frac{p'_\alpha(0)}{M} \\
 p'_\alpha \left(\frac{\Delta t}{2} \right) &= e^{-\xi_\alpha(0)(\Delta t/2)} \\
 &\times \left[p'_\alpha(0) - \frac{\Delta t}{2} \left(\frac{\partial U(q'(0), \alpha'(0))}{\partial \alpha'} + \mu_\alpha N \right) \right].
 \end{aligned} \tag{12}$$

We next update the particle positions, momenta and thermostat variable, ξ . This update proceeds in three steps, where for i from 1 to n we first update the momenta and

positions:

$$\begin{aligned}
 p' \left(\frac{(2i-1)\delta t}{2} \right) &= e^{-\delta t \xi (2(i-1)\delta t/2)} \left[p' \left(\frac{(2i-1)\delta t}{2} \right) \right. \\
 &\quad \left. - \frac{\delta t}{2} \left(\frac{\partial U \left(q' \left(\frac{(2i-1)\delta t}{2} \right), \alpha' \left(\frac{(2i-1)\delta t}{2} \right) \right)}{\partial q'} \right) \right] \\
 q' \left(\frac{2i\delta t}{2} \right) &= q' \left(\frac{(2i-1)\delta t}{2} \right) + \delta t \frac{p' \left(\frac{(2i-1)\delta t}{2} \right)}{m}.
 \end{aligned} \tag{13}$$

Then, we update the thermostat variable which controls the time scaling of the virtual system:

$$\begin{aligned}
 \xi \left(\frac{2i\delta t}{2} \right) &= \xi \left(\frac{(2i-1)\delta t}{2} \right) \\
 &\quad + \frac{\delta t}{Q} \left(\frac{p' \left(\frac{(2i-1)\delta t}{2} \right)^2}{m} - \frac{L}{\beta} \right).
 \end{aligned} \tag{14}$$

Next, we again update the momentum:

$$\begin{aligned}
 p' \left(\frac{2i\delta t}{2} \right) &= e^{-\delta t \xi (2i\delta t/2)} p' \left(\frac{(2i-1)\delta t}{2} \right) \\
 &\quad - \frac{\delta t}{2} \left(\frac{\partial U \left(q' \left(\frac{2i\delta t}{2} \right), \alpha' \left(\frac{2i\delta t}{2} \right) \right)}{\partial q'} \right).
 \end{aligned} \tag{15}$$

To complete one full alchemical time step, we update the alchemostat time scaling variable followed by the alchemical variable and its conjugate momentum:

$$\begin{aligned}
 \xi_\alpha(\Delta t) &= \xi_\alpha(0) + \frac{\Delta t}{Q_\alpha} \left(\frac{p'_\alpha \left(\frac{\Delta t}{2} \right)^2}{M} - \frac{L_\alpha X}{\beta} \right) \\
 p'_\alpha(\Delta t) &= e^{-(\Delta t/2)\xi_\alpha(\Delta t)} p'_\alpha \left(\frac{\Delta t}{2} \right) \\
 &\quad - \frac{\Delta t}{2} \left(\frac{\partial U \left(q'(\Delta t), \alpha'(\Delta t) \right)}{\partial \alpha'} \right) \\
 \alpha'(\Delta t) &= \alpha' \left(\frac{\Delta t}{2} \right) + \frac{\Delta t}{2} \frac{p'_\alpha(\Delta t)}{M}.
 \end{aligned} \tag{16}$$

The addition of angular momentum, multiple alchemical variables and/or alchemical forces arising from additional constraints is straightforward and implemented in code, but is not included here for brevity. Also not included are the numerous alchemical variants of thermodynamic relations which are a product of the thermodynamic rigor the Digital Alchemy framework provides, but are not used directly in this work [11,15].

A useful feature of this Alch-MD implementation is that we have the flexibility to incorporate different design biases or constraints into the system Hamiltonian directly

during our optimisation, i.e. adding harmonic springs attaching particles at specific crystal sites to bias a particular crystal structure or adding a measurement of the system as an energy bias term to the Hamiltonian to bias the system toward a certain property.

In an example in Section 3.2.2, we do this by introducing nonzero μ_α values for the alchemical potential to bias the system toward different regions of design space for different structures. In another example, in Section 3.2.3, we introduce a bias term, $U_{\text{bias}}(q, \alpha)$, to Equation (10) as a part of the $U(q, \alpha)$ term. Here,

$$U_{\text{bias}}(q, \alpha) = \lambda(B(q, \alpha) - B_0)^2, \quad (17)$$

where $B(q, \alpha)$ represents a physical property in this work, bulk modulus of the system calculated during the simulation in real time, B_0 is the target value for the physical property, and λ is a free stiffness parameter.

In implementation, regarding the derivatives of $U_{\text{bias}}(q, \alpha)$, only $\partial U_{\text{bias}}(q, \alpha)/\partial \alpha$ is calculated. The other term, $\partial U_{\text{bias}}(q, \alpha)/\partial q$, is ignored because the fact the per-particle effect would be negligible since B is determined as a collective result of all the particle positions.

3.2. Application examples of Alch-MD

3.2.1. Inverse design of OPP for BCC structure

To validate the Alch-MD method and confirm that the expected behaviour from the derived $NVT\mu_\alpha$ ensemble is observed in simulation, we first perform Alch-MD at low temperature within a design region for a single structure, at which condition it can be expected that the potential energy will dominate any entropic terms of the free energy within a single structure. Thus the free energy minimising trajectories followed in alchemical space should be closely approximated by those that minimise the potential energy for a given structure. We used the BCC forming region in the OPP design space as a test case, the validation proceeded in two steps for 255 simulations with 17 different initial values of k (4–8, with $\Delta k = 0.25$) and 15 different initial values of ϕ (0.1π – 1.5π , with $\Delta \phi = 0.1\pi$).

Step 1 – Droplet assembly: We initialised a system of 4096 particles in a simple cubic lattice with number density 0.0569 (i.e. unit cell length $a = 2.6$) at temperature $T^* = 3.0$ with randomly initialised velocities based on a Boltzmann distribution (here we report the reduced temperature $T^* = kT/\varepsilon$, with units of energy ε). We randomised the particle positions by running for a few iterations at high temperature, and then cooled the system in two steps: first, a steep, linear cooling ramp from $T^* = 3.0$ to $T^* = 1.0$ over 5×10^6 time steps, and then a slow, linear cooling ramp from $T^* = 1.0$ to $T^* = 0.1$ over 9×10^7 time steps, all with $\delta t = 0.005$. During this

cooling, we observed the system of particles self-assemble into a larger nanocrystal with BCC crystal structure. We then ran for an additional 5×10^6 time steps at $T^* = 0.1$ to equilibrate the nanocrystal.

Step 2 – Alch-MD optimisation without property bias: Following the simulations that produced a BCC nanocrystal in step 1, Alch-MD simulations were run over 6×10^6 time steps with $\delta t = 0.001$, $T^* = 0.1$ to find the optimal interaction potential parameters for BCC within the OPP design space, yielding the BCC nanocrystal's ‘eigenpotential’ within the potential constraints and at the selected temperature.

We plot only the BCC-forming region, which is our chosen design space, and the corresponding potential energy per particle for each k and ϕ pair in Figure 2.

Starting with different self-assembled nanocrystals in the BCC region, we ran Alch-MD to find the optimum values of the alchemical variables corresponding to the OPP parameters k and ϕ for the BCC structure (*Step 2*). Eight randomly selected initial points from these simulations are shown by the lines plotted in Figure 2(a). These lines show the change of k and ϕ starting from different initial conditions and the path they follow in alchemical space to converge to the same optimum OPP potential, with minor deviation arising from nanocrystal defects. All data points converge around the optimum potential for BCC with $\bar{k} = 7.36 \pm 0.03$ and $\bar{\phi} = (0.722 \pm 0.003)\pi$ in this alchemical design space. Further comparison with the brute force obtained energy landscape confirms that the optimal potential energy is obtained with parameters that correspond to the region of k and ϕ that yield the BCC structure with the minimum energy, as shown in Figure 2(b).

Among all the points converging around the optimum (k, ϕ) pair, we found the one with the lowest potential energy at $k_* = 7.362 \pm 0.004$, $\phi_* = (0.724 \pm 0.001)\pi$, with a starting $k_0 = 6.0$ and $\phi_0 = 0.7\pi$. Nanocrystals at this state point (k_*, ϕ_*) have a per-particle potential energy $U = -3.461 \pm 0.002$.

The results discussed in Figure 2 indicate that in an unbiased digital alchemy simulation (where $\mu_{\{k, \phi\}} = 0$ and alchemical parameters are free to change without any energy associated with it), we can effectively optimise the potential to locate the regional optimum point, in this case the optimum potential for the BCC structure. While all these simulations are done with droplet systems with 4096 particles where the surface has non-negligible energy contribution to the whole system, we performed similar simulations across a large range of droplet sizes and showed that the effect of surface on optimal potential parameter diminishes much faster than its effect on interfacial energy (SI Figure 1), and our system size is large enough to obtain results close to the bulk system limit.

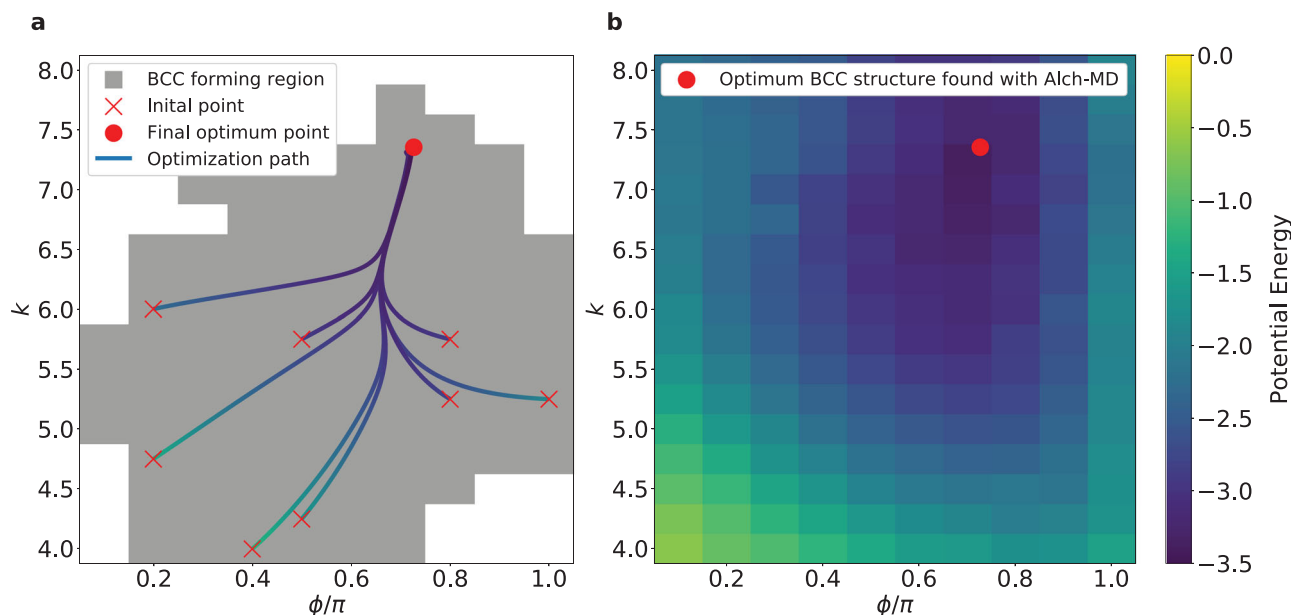


Figure 2. Unbiased optimisation of the BCC structure in the OPP design space. (a) The self-assembly region of the BCC structure in the OPP design space is shown in the background, the lines show the optimisation process of Alch-MD to find the optimum potential for BCC and are coloured according to the observed potential energy during optimisation. (b) A potential energy heat map, obtained from the results of the brute force self-assembly simulations depicted in the background of (a). The region with the lowest energy coincides with the optimised values of (k, ϕ) found via Alch-MD.

3.2.2. Inverse design of LJG potential for different target crystal structures

Many systems, however, may not contain a large energy basin in potential parameter design space of a structure like that observed in the OPP system for BCC structure. Here, we provide examples with a set of self-assembled 2D structures observed in previous work by Engel et al. in the LJG system [30]. Many of these structures do not have a local minimum in this extended energy landscape, as shown in Figure 3. This means that an unbiased Alch-MD simulation will move the system away from the corresponding self-assembly region and can allow the system to transition into a different, lower energy, structure as seen in Figure 4(a). For this case, we introduced nonzero alchemical potential μ_α in our simulation and studied how it could be used to add bias in potential design space for different target crystal structures.

The simulation for the 2D LJG system proceeded in 2 steps:

Step 1 – Droplet assembly: Using a system of 1024 particles, nanocrystals were self-assembled at randomised (r_0, ϵ) values chosen from each structure in the known self-assembly design landscape from Engel et al. [30]. Beginning with randomised particle positions and velocity, a linear cooling ramp from $T^* = 3.0$ to $T^* = 0.1$ was implemented over 5×10^7 time steps with $\delta t = 0.01$. The structures were then checked against the reported structures of Engel et al. to confirm the proper structure was assembled.

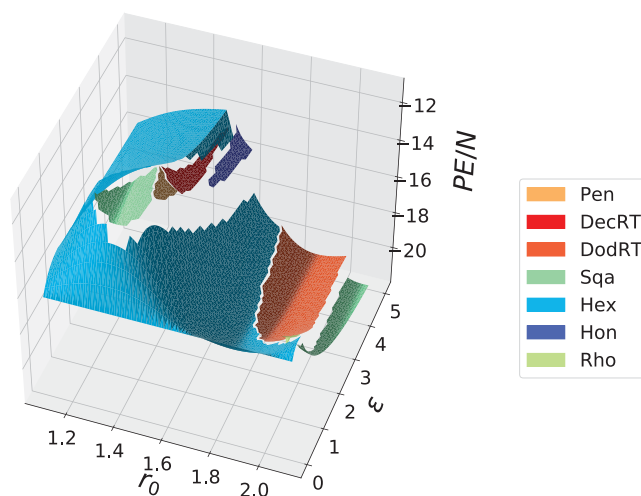


Figure 3. Alchemically extended potential energy landscape of low energy 2D LJG structures. Plotted regions are determined using the self-assembly phase diagram of Engel et al. The shading illustrates that no structure contains a minimum in both alchemical parameters within the studied region.

Step 2 – Alch-MD optimisation for different target structures: Following the simulations in step 1, we further ran Alch-MD simulations over 2×10^6 time steps with $\delta t = 0.001$ at $T^* = 0.1$ under three conditions: unbiased $\mu_\alpha = 0$, $\mu_\alpha = c$ where c is some constant, and under the constraint of an energy penalty function that effectively confines the system in a targeted region of the design space.

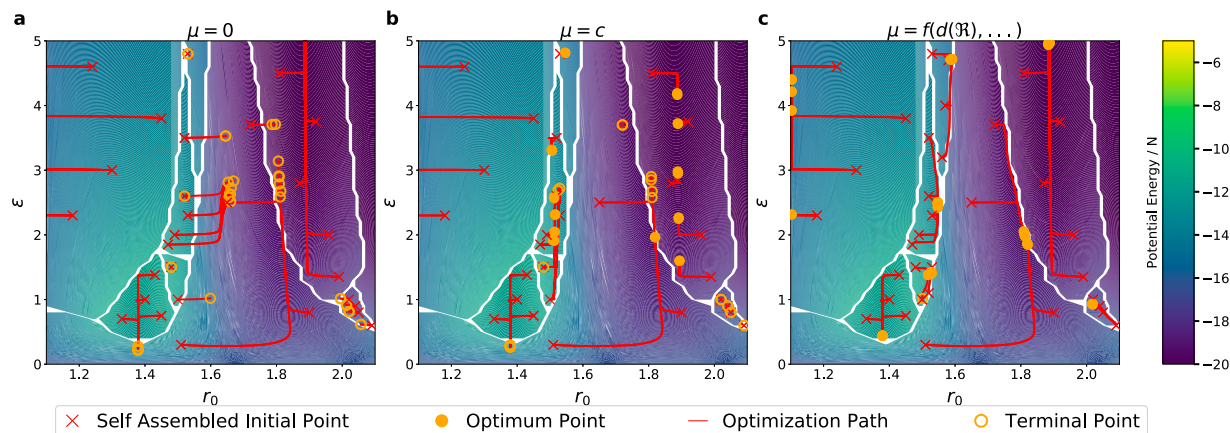


Figure 4. Alch-MD simulations of 2D LJG structures with different energy bias/alchemical potentials applied. (a) Alch-MD simulation with $\mu_\alpha = 0$ (no bias). (b) Alch-MD simulation with a constant μ_α value for each different structure in the format (structure:(μ_ϵ, μ_{r_0})) as follows: Pen: (0.1,3.0); DecRT: (0.42,3.3); DodRT: (0.39,0.0); Sq: (-0.02,0.0); Hex2: (0.2,0.0); Hex1: (0.0,0.0); Hon: (0.58,6.3); Rho: (0.9,-8.5). (c) Alch-MD simulation with a harmonic energy bias evaluated by the distance to the self assembly phase boundary.

As expected, simulations that start from a self-assembly region for a structure without a local minimum often lead to phase transitions as shown in Figure 4(a) with the open circles denoting these structural changes. It can be observed that a stable behaviour in the location of the second well is much more effective in preventing structural changes. Large movements in the relative energy depths of the wells are much better tolerated by systems.

In order to perform alchemical optimisations in these more complex energy landscapes, some form of bias must be introduced. The simplest form of bias is to add a non-zero alchemical potential μ_α , which assigns an energy penalty to changes in the alchemical variable. For the case of phase boundaries that are convex in the relevant design space for a structure, a constant value of the alchemical potential is sufficient to stabilise these structures, as shown in Figure 4(b). The exceptions are those that exhibit saddle points, particularly the square and rhombic structures. However, it can also be noted that the quasicrystals, while stable, do not converge. This is an effect of compressing all configuration information to a single structure label as the energy landscapes of these random-tiling/quasicrystal configurations are much more complex than those of periodic crystals.

When the underlying landscape is not convex, a more complex bias term must be used. Because the region of interest is known for each structure in this study, a simple harmonic energy penalty was applied to the systems based on the evaluated distance, d , from the region boundary. Thus our energy bias term looks like $U_{\text{bias}}(\alpha) = \lambda d(\alpha, \mathfrak{R})^2$ where $\mathfrak{R}$ defines the region of interest (in this case the phase boundary) and λ is the

spring constant for the harmonic spring. The results of this bounding bias can be seen in Figure 4(c).

When applied in parallel with an experimental system, the region of interest can similarly reflect what is physically realisable with current techniques, materials and desired application.

The relative success rates of the approaches in Figure 4(a)–(c) at retaining the structure of the tested self-assembled nanocrystals within the time of Alch-MD simulation is 31.6%, 56.1% and 94.2%, respectively. However, if an infinite amount of time was allowed, it would be found that only the hexagonal structure is thermodynamically stable with no bias applied. This is because the regions significantly outside the known self-assembly region correspond to interaction potentials with only single well, which, in 2D systems lacking pressure or anisotropy, is expected to stabilise in the hexagonal structure. These stable structures and conditions are thus a starting point for further complex optimisation, as discussed below for the case of property optimisation.

For many physical systems, it can be expected that changing the interparticle interactions would not be free, and thus $\mu_\alpha \neq 0$ is expected and it should be associated with the cost of changing interaction. For example for a system with tunable interaction, the alchemical potential should be considered as an additional parameter that can be linked physically to the system variable, e.g. for a system where interaction is tuned by changing magnetic field, μ_α would be related to intensity of the field.

Similar to the OPP nanocrystal system, we studied simulations across a large range of droplet sizes for 2D LJG system and confirmed that our system size used is large enough to obtain results close to the bulk system limit (details in SI Figure 2).

3.2.3. Inverse design of OPP for property design

Following the structure optimisation described in the previous sections, our bulk modulus design protocol begins with the obtained optimised OPP potential for the BCC nanocrystal. An additional step is applied to optimise for a target bulk modulus:

Step 3 – Alch-MD optimisation with structure and property bias: Continuing from the final configuration of Step 2, an additional energy term is activated to bias the

Alch-MD simulation to satisfy the bulk modulus requirement. The system was kept at the same $T^* = 0.1$ and ran for an additional 3×10^5 time steps with $\delta t = 0.0001$ to allow the system to optimise for the target property bias. We used the last 1×10^5 time steps to sample k and ϕ for statistics. A stiffness constant, λ that tunes the energy penalty from the property bias, ranging from 0.1 to 0.5 and a set of targeted bulk moduli, B_0 , ranging from 1.0 to 7.0 were used. All simulations were replicated at

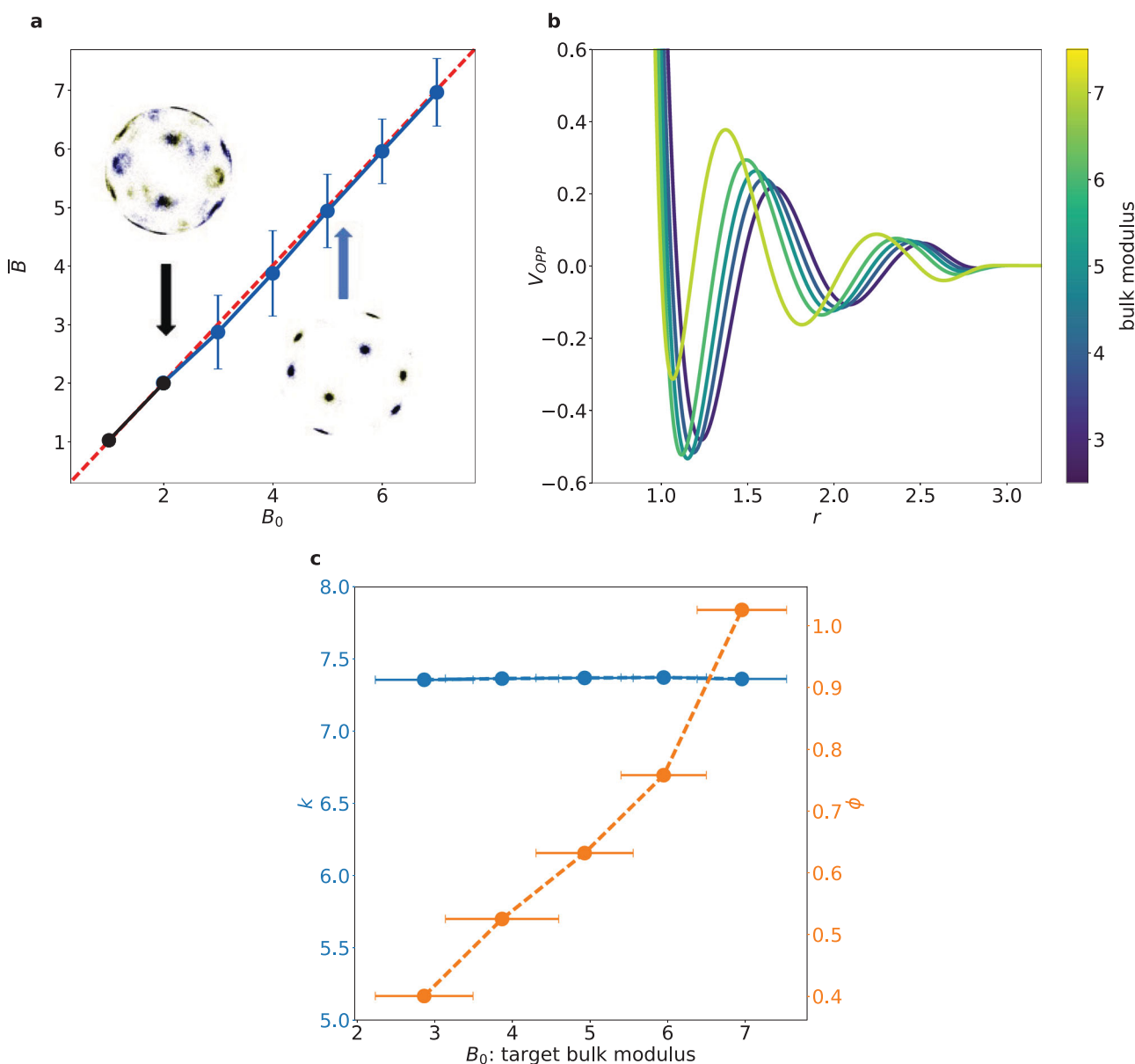


Figure 5. Designed systems with $\lambda = 0.1$, targeting bulk modulus values B_0 from 1 to 7. **(a)** Change of system bulk modulus with digital alchemy for different target B_0 . Blue data points indicate an effective tuning of the bulk modulus; the structure of the droplet remains perfectly crystalline, as illustrated by the clear BOOD (right inset). Black data points (in the low bulk modulus region) suggest that the target B_0 is beyond the design limit for this system and the simulation distorts the crystal into multiple grains to achieve the target, indicated by the BOOD (left inset). The red dotted line shows the $B_0 = \bar{B}$ line. Error bars in this plot show the statistical fluctuation of the $\bar{B}$ value at the end of simulation; error bars for the data points shown in black are very small and thus not visible in the plot. **(b)** Optimised OPP for different target values B_0 from low to high values, plotted in purple/blue/green/yellow colour, according to the depicted colour bar. **(c)** Correlation between targeted bulk modulus and design parameter k and ϕ , with $\lambda = 0.1$ (targeting B_0 between 3.0 and 7.0).

least once and no statistically significant difference was observed in any of the three steps.

We started with the best optimised potential for BCC without any property bias: Nanocrystals at this state point ($k_* = 7.362 \pm 0.004$, $\phi_*(0.724 \pm 0.001)\pi$) have bulk modulus $B = 5.92$ calculated from the last frame of the crystal. We carried out Alch-MD simulations with different target bulk modulus B_0 and coupling constant λ . Detailed results for $\lambda = 0.1$ are given in Figure 5. Other values of λ show no significant difference compared to the results for $\lambda = 0.1$; a summary of the results for all

values of λ is provided in the supplementary information (SI Figure 3) and suggests that λ has flexible working range with negligible variation for values ranging from 0.1 to 0.5.

Figure 5(a) shows a clear monotonic relationship between target bulk modulus, B_0 , and the average bulk modulus, $\bar{B}$, in thermalised crystalline nanocrystals of particles interacting via the designed isotropic pair potential. Note that the measured values of $\bar{B}$ fall directly on the design curve for $B_0 \geq 3.0$ (within statistical error); however, for $B_0 \leq 2.0$, we observe deviations in the bulk

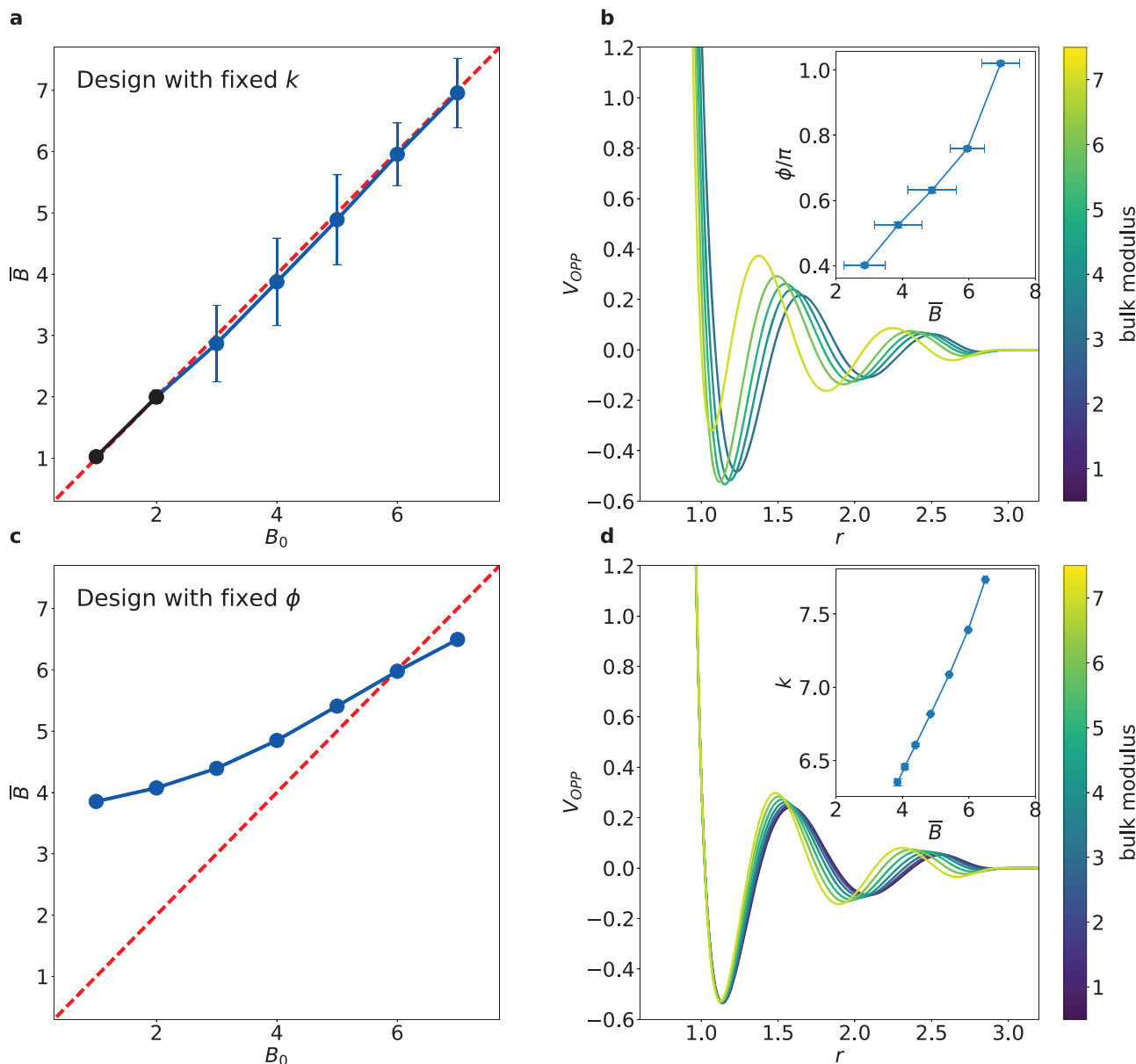


Figure 6. Bulk modulus biased design with only one degree of freedom (k or ϕ). (a) System bulk modulus after optimisation of ϕ across different target values B_0 . (b) ϕ -optimised OPPs for different bulk moduli. The correlation between measured bulk modulus and the design parameter ϕ (with fixed $k = 7.3629$) is shown in the inset. (c) System bulk modulus after optimisation of k across different target values B_0 . (d) k -optimised OPPs for different bulk moduli. The correlation between measured bulk modulus and the design parameter k (with fixed $\phi = 0.7237\pi$) is shown in the inset.

modulus from the target value. We believe this deviation occurs because those target values are approaching the physical limit of the bulk modulus achievable in this design space. Plots of the BOODs are inset in Figure 5(a); all nanocrystals with $3.0 \leq B_0 \leq 7.0$ are BCC, but lower structural quality and indications for polycrystallinity are observed for $B_0 \leq 2.0$. Figure 5(b) shows that the pair potential changes systematically to accommodate the change in bulk modulus.

Figure 5(c) indicates that the designed potentials are localised in a very narrow band in k , $7.32 \leq k \leq 7.39$, for $\lambda = 0.1$ over all measured values of $\bar{B}$. In contrast, changes in ϕ are predominantly responsible for determining the bulk modulus via a clear, monotonic relation.

The design rules indicated by Figure 5(c) suggest primary control through ϕ and an insensitivity to k within the narrow range previously noted. These findings raise two questions:

- (1) Can ϕ act as an effective control parameter for the bulk modulus entirely on its own (or is the variation in k small but important)?
- (2) Is k useless as a control parameter for the bulk modulus, or is its impact simply small compared to effects in ϕ ?

To answer these questions, we performed Alch-MD simulations with $\lambda = 0.1$ with (i) fixed $k = k_*$ and variable ϕ and (ii) with fixed $\phi = \phi_*$ and variable k . The results are plotted in Figure 6. Figure 6(a) indicates that Alch-MD with fixed $k = k_*$ is as successful as designing with variable k . Figure 6(b) shows that the design rule for ϕ with fixed $k = k_*$ yields similar ϕ values to the variable k results in Figure 5(b) and Figure 5(c). Note that, if plotted in Figure 6(b), the optimised potentials for variable k design are indistinguishable from fixed k . In contrast, Figure 6(c) indicates that while fixed $\phi = \phi_*$ design produces a consistent monotonic relationship between target and designed bulk modulus, it achieves much lower fidelity. Figure 6(d) shows that for Alch-MD with fixed $\phi = \phi_*$ design, the k design rule is drastically altered. This suggests that k can be used to tune the bulk modulus over a limited range, but that ϕ is a more effective control parameter.

4. Discussion and conclusion

We developed and implemented digital alchemy in molecular dynamics simulation (Alch-MD). We demonstrated that Alch-MD can serve as an inverse design approach that produces particle interactions leading to

materials with a targeted structure and/or bulk modulus as an example physical property.

First, a new thermostat was derived and implemented to allow Alch-MD simulations to be performed in the $NVT\mu_\alpha$ ensemble. Alch-MD was utilised to efficiently determine the optimal set of interaction potential parameters within an extended free energy landscape basin. Using self-assembled structures, the optimal parameters that Alch-MD revealed were shown to be only dependent on the structure and the alchemical potential if the simulation stayed within the same free energy basin. We validated our simulation method with two model pair potentials, OPP and LJG, using both brute force simulations and data from previous publications, and showed that our method can find the optimal potential parameters for a desired structure. To validate the optimal potential, we evaluated the potential energy as a proxy for free energy. We expect that the entropy contribution within a simulation should be relatively small and nearly constant due the restriction of simulations at low temperature and single structure, respectively.

While we only focus on low temperature cases in this work, where known potential energy landscape can be a good proxy for free energy, applying Alch-MD to systems at higher temperature where entropy is relevant, is straightforward. The validation of Alch-MD simulations at these temperatures becomes difficult, however, due to the difficult task of computing the extended free energy landscape which covers all degrees of freedom. For this reason in particular, we consider digital alchemy methods an exciting tool given that it samples appropriately in both the configuration and alchemical spaces. This thermodynamically rigorous approach binds Alch-MD to the relevant free energy landscape and we therefor think this could be a rich field for followup work.

We then applied our inverse design approach to property design in the OPP model system. In this system, we showed that the BCC structure can be optimised with zero alchemical potential and is both self-assembled and stable over a large region of the design space. The design constraint of a target bulk modulus was then introduced as a simple energy penalty via an external energy parameter in Alch-MD to bias the system away from its unconstrained optimal point while obtaining the targeted property value.

We showed that – within statistical error – the assembled nanocrystals exhibit bulk moduli that fall precisely on the design curve over a range of target values, and that deviations only become appreciable as target values approach physical system limits. Carrying out this design process for several target bulk moduli yields a set of clear design rules indicating which design parameters are most relevant to the target property, and what values

the design parameters should take for the chosen target physical response.

This work focused on systems of particles which assemble into nanocrystals not large enough to span the simulation volume in a $NVT\mu_\alpha$ simulations, so changing interaction potential allows for changing lattice spacings and the system is free to explore different configuration and density. If a crystal was to span the box in any dimension, it would impose artificial constraints. While this means the system will have non-negligible surface effect in many ways, our simulations were performed at sizes sufficiently large to approximate the constant surface effect of larger and larger nanocrystals, based on our testing detailed in the SI Figures 1 and 2. Comparing these results to those obtained for truly bulk crystals with the $NPT\mu_\alpha$ ensemble is an interesting direction for future work.

This investigation was intended as a proof-of-principle; the Alch-MD approach can be extended in several ways. We restricted our investigation to monoatomic interacting via two families of isotropic pair potentials with two parameters. We chose the two pair potential families LJG and OPP because both have been exhaustively investigated elsewhere [27–30], and therefore we had access to reference data for validation. Extensions to potentials with more than two parameters, potentials for anisotropic, patchy interactions or multiple species of particles, are straightforward. Alch-MD can also be utilised to study unknown design spaces, as trajectories will locally track the gradient of the free energy surface in alchemical space. By running many trajectories, we can map out attractors and saddles in alchemical space.

Here we designed for the system's bulk modulus, however, extensions to other physical properties are limited only by the cost of computing the relevant property. Alch-MD is particular apt for designing dynamic properties because it is backed by the defined timescales of molecular dynamics. Note that, especially for high-dimensional design spaces, this method can still be very computationally effective even if the cost of calculating a property is high, because a single Alch-MD simulation can eliminate the need to blanket the design space with Edisonian simulations.

While using Alch-MD to guide materials design in experiments, in cases where the optimal alchemical parameters cannot be reached, non-zero alchemical potential can be effectively confine our alchemical 'search' to a physically relevant design space. While we focused on the case in which Alch-MD can be used for the inverse design of particles with constant interactions and the resulting properties in this work, experimentally realisable systems with dynamic interactions could also

fit well into the Alch-MD framework. The application of our alchemostat to these systems is trivial if the normalisation and alchemical potential accurately reflect all experienced states at the relevant temperatures and the energy cost of the associated changes, respectively. In this case, the alchemical potential would have a physical interpretation as the force with which the process driving the dynamics exerts in the alchemical space.

We believe that future work extending this approach along the above-stated directions, as well as working closely with experimentally achievable 'colloidal alchemy' examples, will give concrete steps toward achieving synthesisable colloidal and other materials with tailored properties.

Acknowledgements

P. Z. thanks Julia Dshemuchadse for helpful discussion on an earlier version of the manuscript and Matthew Spellings for providing the structure identification code and training data set. J. C. P. thanks Joshua A. Anderson for guidance with the alchemostat's implementation in HOOMD-blue. GvA thanks Jens Glaser for discussions on Trotter factorisations.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported in part by the National Science Foundation, Division of Materials Research under Computational and Data-Enabled Science and Engineering (CDS&E) Award # DMR 1409620 (2014–2018) and DMR 1808342 (2019–2023), and also by a grant from the Simons Foundation (256297, to SCG). Support for computational resources and services were provided by Advanced Research Computing at the University of Michigan, Ann Arbor.

References

- [1] A. Franceschetti and A. Zunger, *Nature* **402** (6757), 60–63 (1999).
- [2] M.Z. Miskin and H.M. Jaeger, *Nat. Mater.* **12**, 326–331 (2013).
- [3] M.Z. Miskin and H.M. Jaeger, *Soft Matter* **10**, 3708–3715 (2014).
- [4] A. Jain, J.A. Bollinger and T.M. Truskett, *AIChE J.* **60** (8), 2732–2740 (2014).
- [5] É Marcotte, F.H. Stillinger and S. Torquato, *J. Chem. Phys.* **138** (6), 061101 (2013).
- [6] A. Jain, J.R. Errington and T.M. Truskett, *Phys. Rev. X* **4**, 031049 (2014).
- [7] A. Jain, J.R. Errington and T.M. Truskett, *Soft Matter* **9**, 3866–3870 (2013).
- [8] M.Z. Miskin, G. Khaira, J.J. de Pablo and H.M. Jaeger, *Proc. Natl. Acad. Sci. USA* **113** (1), 34–39 (2016).
- [9] E. Edlund, O. Lindgren and M.N. Jacobi, *Phys. Rev. Lett.* **107**, 085503 (2011).

- [10] R.B. Jadrich, J.A. Bollinger, B.A. Lindquist and T.M. Truskett, *Soft Matter* **11** (48), 9342–9354 (2015).
- [11] G. van Anders, D. Klotz, A.S. Karas, P.M. Dodd and S.C. Glotzer, *ACS Nano* **9**, 9542–9553 (2015).
- [12] B.A. Lindquist, R.B. Jadrich and T.M. Truskett, *Soft Matter* **12** (12), 2663–2667 (2016).
- [13] C.S. Adorf, J. Antonaglia, J. Dshemuchadse and S.C. Glotzer, *J. Chem. Phys.* **149** (20), 204102 (2018).
- [14] Y. Geng, G. van Anders, P.M. Dodd, J. Dshemuchadse, and S.C. Glotzer, *Sci. Adv.* **5** (7), eaaw0514 (2019).
- [15] R. Cersonsky, G. van Anders, P.M. Dodd and S.C. Glotzer, *Proc. Natl. Acad. Sci. USA* **115**, 1439–1444 (2018).
- [16] D. Frenkel and B. Smit, *Understanding Molecular Simulations: From Algorithms to Applications* (Elsevier, 2002).
- [17] Y. Kim, R.J. Macfarlane, M.R. Jones and C.A. Mirkin, *Science* **351** (6273), 579–582 (2016).
- [18] J.J. Juárez and M.A. Bevan, *Adv. Funct. Mater.* **22** (18), 3833–3839 (2012).
- [19] J.A. Anderson, C.D. Lorenz, and A. Travesset, *J. Comput. Phys.* **227** (10), 5342–5359 (2008).
- [20] J. Glaser, T.D. Nguyen, J.A. Anderson, P. Lui, F. Spiga, J.A. Millan, D.C. Morse, and S.C. Glotzer, *Comput. Phys. Commun.* **192**, 97–107 (2015).
- [21] C. Zhang, R.J. Macfarlane, K.L. Young, C.H.J. Choi, L. Hao, E. Auyeung, G. Liu, X. Zhou and C.A. Mirkin, *Nat. Mater.* **12**, 741–746 (2013).
- [22] Y. Wang, Y. Wang, X. Zheng, É Ducrot, J.S. Yodh, M. Weck and D.J. Pine, *Nat. Commun.* **6**, 7253 (2015).
- [23] W.B. Rogers and V.N. Manoharan, *Science* **347** (6222), 639–642 (2015).
- [24] Y. Zhang, S. Pal, B. Srinivasan, T. Vo, S. Kumar and O. Gang, *Nat. Mater.* **14**, 840–847 (2015).
- [25] Y.J. Lee, N.B. Schade, L. Sun, J.A. Fan, D.R. Bae, M.M. Mariscal, G. Lee, F. Capasso, S. Sacanna, V.N. Manoharan and G.R. Yi, *ACS Nano* **7** (12), 11064–11070 (2013).
- [26] M. Rey, A.D. Law, D.M.A. Buzza and N. Vogel, *J. Am. Chem. Soc.* **139** (48), 17464–17473 (2017).
- [27] M. Mihalkovič and C. Henley, *Phys. Rev. B.* **85** (9), 092102 (2012).
- [28] M. Engel, P.F. Damasceno, C.L. Phillips and S.C. Glotzer, *Nat. Mater.* **14**, 109–116 (2015).
- [29] J. Dshemuchadse, P.F. Damasceno, C.L. Phillips, M. Engel, and S.C. Glotzer, In preparation (2019).
- [30] M. Engel and H.R. Trebin, *Phys. Rev. Lett.* **98**, 225505 (2007).
- [31] J.A. Zimmerman, E.B. WebbIII, J. Hoyt, R.E. Jones, P. Klein and D.J. Bammann, *Model. Simul. Mater. Sci. Eng.* **12** (4), S319–S332 (2004).
- [32] L.D. Landau and E.M. Lifshitz, *Theory of Elasticity*, 3rd ed. (Butterworth-Heinemann, Oxford, 1986).
- [33] M. Spellings and S.C. Glotzer, *AIChE J.* **64** (6), 2198–2206 (2018).
- [34] M. Spellings and S.C. Glotzer, Pythia (2019). < <https://pythia-learn.readthedocs.io> > .
- [35] F. Chollet et al., Keras (2015). < <https://keras.io> > .
- [36] V. Ramasubramani, B.D. Dice, E.S. Harper, M.P. Spellings, J.A. Anderson and S.C. Glotzer, arXiv e-prints arXiv:1906.06317 (2019).
- [37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
- [38] G.J. Martyna, M.E. Tuckerman, D.J. Tobias and M.L. Klein, *Mol. Phys.* **87** (5), 1117–1157 (1996).