IOP Publishing Modelling and Simulation in Materials Science and Engineering

Modelling Simul. Mater. Sci. Eng. **23** (2015) 074008 (18pp) doi:10.1088/0965-0393/23/7/074008

Interpolation effects in tabulated interatomic potentials

M Wen, S M Whalen, R S Elliott and E B Tadmor

Department of Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis, MN 55455, USA

E-mail: tadmor@aem.umn.edu

Received 4 December 2014, revised 10 April 2015
Accepted for publication 15 April 2015
Published 18 September 2015

Abstract

Empirical interatomic potentials are widely used in atomistic simulations due to their ability to compute the total energy and interatomic forces quickly relative to more accurate quantum calculations. The functional forms in these potentials are sometimes stored in a tabulated format, as a collection of data points (argument–value pairs), and a suitable interpolation (often spline-based) is used to obtain the function value at an arbitrary point. We explore the effect of these interpolations on the potential predictions by calculating the quasi-harmonic thermal expansion and finite-temperature elastic constant of a one-dimensional chain compared with molecular dynamics simulations. Our results show that some predictions are affected by the choice of interpolation regardless of the number of tabulated data points. Our results clearly indicate that the interpolation must be considered part of the potential definition, especially for lattice dynamics properties that depend on higher-order derivatives of the potential. This is facilitated by the Knowledgebase of Interatomic Models (KIM) project, in which both the tabulated data ('parameterized model') and the code that interpolates them to compute energy and forces ('model driver') are stored and given unique citeable identifiers. We have developed cubic and quintic spline model drivers for pair functional type models (EAM, FS, EMT) and uploaded them to the OpenKIM repository (https://openkim.org).

Keywords: interatomic potentials, spline interpolation, lattice dynamics, Knowledgebase of Interatomic Models

(Some figures may appear in colour only in the online journal)

1. Introduction

Empirical interatomic potentials (IPs) have long been used to model atomic bonding. Initially, IPs were used as part of theoretical studies of material behavior (such as the Lennard-Jones [1, 2] and Morse [3] potentials). Following the invention of the electronic computer in the 1940s, Monte Carlo (MC) and molecular dynamics (MD) simulations became possible, and this led to a renewed interest in the development of IPs for a variety of metallic and covalent systems. This trend accelerated in the 1980s and, since then, a large number of different and increasingly complex IP descriptions have been developed (see [4] for more on the history and scientific development of IPs and MD). The reduced computational cost of IPs relative to more accurate quantum methods makes it possible to simulate larger systems for longer times, and thus to tackle problems that are inaccessible to quantum simulations, such as plastic deformation, fracture, atomic diffusion, and phase transformations [5].

Empirical IPs typically include functional forms with parameters that are obtained by fitting to a relevant training set of experimental and quantum data. In practice, the fitted functional forms are often stored in a discretized tabulated format as a list of data points, and intermediate values are obtained by interpolation. For example, a pair potential function $\phi(r)$ would be stored as a set of values: (r_1, ϕ_1), (r_2, ϕ_2), …, (r_n, ϕ_n), where $\phi_i = \phi(r_i)$ and n is the number of data points. In some cases, an analytic form for the function does not exist and the fitting procedure directly generates the potential in its discretized form with a small number (typically 10–30) of ϕ_i values used as the fitting parameters. This is the case, for example, for the force-matching scheme pioneered by Ercolessi and Adams [6].

IPs have been tabulated since the early days of molecular simulations and continue to be tabulated in popular MD codes. Interpolation of tabulated data using polynomials or splines is computationally more efficient than calculating the total energy and forces directly from the analytic functional forms, especially when the analytic functional forms are complicated. Moreover, the use of tabulated IPs makes the subroutines for calculating the total energy and forces general to a family of potentials, e.g. a single routine can be used for all pair potentials. This facilitates the incorporation of user-defined IPs into existing codes because the task reduces to providing a table of numbers in a standard format. Initially, due to memory limitations, the number of data points n supported by these codes was limited (for example, in the original embedded atom method (EAM) potential, n was less than 10 [7]), but currently storing 10 000 data points per function is typical. For IPs based on an analytic form, the use of a large number of stored points provides a very smooth and accurate representation of the energy function. For IPs based directly on an interpolation through a small number of data points, the tabulation of additional points between the fitted values cannot improve the smoothness or accuracy of the energy function. However, in both cases, the large number of tabulated points minimizes differences between simulation codes that use different schemes to interpolate the data (e.g. cubic splines versus quintic splines).

Many popular IP forms are routinely tabulated, including pair potentials, EAM [8–10], Finnis–Sinclair (FS) [11], effective medium theory (EMT) [12], modified EAM (MEAM) [13], and the angular-dependent potential (ADP) [14, 15], among others. There are two common approaches for tabulating IPs. The first ('energy only') tabulates only the energy values of the potential at selected points. Examples are the EAM potentials used in LAMMPS. If derivatives of the energy are needed, then they are computed as derivatives of the interpolated potential energy (e.g. a spline used to interpolate the energy is differentiated). The second ('energy and derivative(s)') approach tabulates the values of the energy as well as its derivative(s). Then, energy derivatives (if needed) can be obtained by interpolating the tabulated derivative data. Ercolessi and Adams store their glue potentials in this way [16]. The first

approach provides a self-consistent system (the derivative will match the numerical derivative of the energy), but may provide a very coarse description of the derivative function that has a low order of continuity. For example, the third-order derivative of a cubic spline is a piecewise constant function, and derivatives of order four and higher are identically zero. This can be problematic, especially when IPs of this type are fit using a small number of data points. The second approach can provide high-order continuous derivative functions, but they will be inconsistent in the sense that the derivative functions will not match the numerical derivatives of the energy function. This inconsistency can result in problems such as energy drift in microcanonical ensemble (NVE) MD simulations [4].

Traditionally, tabulated IPs were delivered with MD codes. More recently, emphasis has been placed on archival storage of tabulated IPs in a standardized data format within the NIST Interatomic Potential Repository Project [17]. A more general, related effort is the Knowledgebase of Interatomic Models (KIM) project that archives arbitrary IPs and includes the development of a standard application programming interface (API) [18] for atomistic simulations [19–21]. The KIM API enables any KIM-compliant IP to work seamlessly with any KIM-compliant simulation code, including LAMMPS [22, 23], IMD [24, 25], DL_POLY [26, 27], and GULP [28, 29]. All items within the open online KIM system (the OpenKIM Repository, hosted at https://openkim.org) have unique citeable identifiers that enable others to reproduce simulation results.

Users of tabulated IPs typically consider the data file containing the discretized functional forms as *the* potential without considering the nature of the interpolation. The reasoning behind this is that if enough data points are used, then the type of interpolation will not affect the results. As alluded to, this argument is largely correct for MC and MD simulations that, for finite temperatures, sample many of the interpolated data points and obtain a 'smeared' result that is effectively independent of the interpolation type. However, the nature of the interpolation can strongly affect methodologies that use higher-order derivatives of the IP without sampling, such as lattice dynamics calculations. For this reason, GULP (a leading lattice dynamics code) does not use tabulated potentials. This issue is also recognized within KIM, where both the tabulated data ('parameterized model') and the code that interpolates the tabulated data to compute energy and forces ('model driver') are stored in the OpenKIM Repository and together define the IP. A recent example where a calculation can 'go wrong' due to interpolation effects is the application of the vibrational self-consistent field (VSCF) method [30] to the second-generation Brenner model [31]. VSCF requires derivatives of the IP up to fourth order; however, the Brenner potential incorporates a cubic spline with knots (data points) at graphene and diamond geometries. This leads to discontinuities in the (numerically computed) third- and fourth-order derivatives at the spline knots and, hence, to a breakdown of the VSCF approach [32].

The objective of this article was to systematically study the effect of tabulation and interpolation on the predictions of an IP. We chose the simplest possible case of a one-dimensional (1D) chain of copper atoms interacting via a nearest-neighbor modified Morse potential [33]. We computed quasi-harmonic predictions for the thermal expansion and finite-temperature elastic constants as well as MD predictions for these properties. Although simple, this example includes all of the features that are expected to play out in three-dimensional lattice dynamics calculations. We studied five types of splines: natural cubic, cubic Hermite, clamped quartic, clamped quintic, and quintic Hermite. As a cautionary tale, we also include a sixth, 'naïve quartic' spline generated using an algorithm that appears reasonable but leads to significant errors. The predictions obtained from the spline computations were compared to the same computations performed with the analytic IP. The results show a strong effect of the interpolation on the computed properties. Strictly speaking, only the clamped quintic spline is able

to reproduce all results obtained from the analytic IP. If the number of data points is on the order of $n = 500$, then the quintic Hermite spline may also do a good job of reproducing the results. The observed strong interpolation effects motivated us to develop KIM model drivers for pair functional IPs (EAM, FS, EMT) for the quintic splines studied here, which have been uploaded to the OpenKIM Repository.

The article is organized as follows. In section 2, we briefly describe the modified Morse potential being used. In section 3, we discuss the basics of spline interpolation and the continuity and locality of splines, and apply the studied splines to the modified Morse potential. In section 4, we present the 1D chain results. We end in section 5 with a summary and conclusions.

2. The empirical interatomic potential: modified Morse potential

We consider 1D nearest-neighbor pair potential interactions where the total energy of a finite chain[1] of N atoms, with positions $x_1 < x_2 < \cdots < x_N$, is given by

$$V = \phi(x_2 - x_1) + \phi(x_3 - x_2) + \cdots + \phi(x_N - x_{N-1}). \tag{1}$$

The pair potential used in this article is a modified Morse potential [33], given by

$$\phi(r) = \begin{cases} \dfrac{D_0}{2B-1}\left[\mathrm{e}^{-2A\sqrt{B}(r-r_0)} - 2B\mathrm{e}^{-A(r-r_0)/\sqrt{B}}\right], & r < r_{\mathrm{cut}}, \\ 0, & r \geqslant r_{\mathrm{cut}} \end{cases} \tag{2}$$

with parameters for copper given by $r_0 = 2.5471$ Å, $A = 1.1857$ Å^{-1}, $D_0 = 0.5869$ eV, $B = 2.265$, and $r_{\mathrm{cut}} = 8.15$ Å. As part of this work, the aforementioned IP has been implemented and is now archived in OpenKIM [19, 34]. The standard Morse potential was modified by MacDonald and MacDonald [33] to improve the agreement with experimental values for the thermal expansion of copper. This was done by adjusting ϕ'', which is largely responsible for the slope of the thermal expansion curve. Specifically, the parameter B is determined by fitting the thermal expansion to the experimental result near the Debye temperature. The standard Morse potential is recovered for $B = 1$.

3. Spline interpolation of tabulated functions

3.1. Overview of spline interpolation

A variety of methods exist for interpolating between a discrete set of data points $(\hat{x}_1, \hat{y}_1), \ldots, (\hat{x}_n, \hat{y}_n)$. In this section, we focus on spline interpolation, which is one of the methods of choice in MD codes. A spline is a function consisting of distinct polynomial segments defined over adjacent intervals that has a specified level of continuity over its entire domain. The *locality* of a spline is related to how quickly a perturbation to a function value $\hat{y}_i$ at data point $\hat{x}_i$ decays away from it. In general, there is a tradeoff between locality and higher-order continuity [35].

Mathematically, a function $S(x)$ on the interval $[a, b]$ is a spline of degree k, if: (i) S is *at most* C^{k-1} continuous on $[a, b]$; (ii) there are sub-intervals $[x_{i-1}, x_i]$, $i = 1, \cdots, n$ with $a = x_0 < x_1 < \cdots < x_n = b$ on which S is polynomial of degree k; and (iii) there are points $x_{i-1} < \hat{x}_i \leqslant x_i$

[1] For a periodic 1D chain with length L, the total energy is given by $V = \phi(x_2 - x_1) + \phi(x_3 - x_2) + \cdots + \phi(x_N - x_{N-1}) + \phi(L + x_1 - x_N)$.

(the 'knots' of S) such that $S(\hat{x}_i) = \hat{y}_i$ for $i = 1, \cdots, n$ (see further, [36–39]). Thus, the spline is defined as

$$S(x) = \begin{cases} S_1(x) = \sum_{j=1}^{k+1} a_{1j}\, x^{j-1}, & x_0 \leqslant x \leqslant x_1, \\ \vdots & \\ S_n(x) = \sum_{j=1}^{k+1} a_{nj}\, x^{j-1}, & x_{n-1} \leqslant x \leqslant x_n, \end{cases} \tag{3}$$

where the $N_{\text{coef}} = n(k+1)$ coefficients a_{ij} are obtained from the requirements that the spline: (i) interpolates the data at the knots (i.e. $S(\hat{x}_i) = \hat{y}_i$); (ii) satisfies certain continuity conditions at the sub-interval boundaries; and (iii) satisfies zero or more boundary conditions (BCs) at the ends. Note that splines of the same degree can have different continuity across sub-interval boundaries depending on the selected conditions as shown here.

Typically for odd-degree splines, the knots are taken to coincide with the sub-interval boundaries, $\hat{x}_i = x_i$, $i = 0, \cdots, n$. For even degree splines, the knots are usually taken as the sub-interval mid-points. As an example, a *natural cubic spline* is a third-degree spline ($k = 3$), which is of odd degree, so that the knots and sub-interval boundaries coincide, $\hat{x}_i = x_i$, and we can drop the hat notation. A natural cubic spline satisfies the interpolation condition,

$$S_i(x_{i-1}) = y_{i-1}, \quad S_i(x_i) = y_i, \quad i = 1, \ldots, n, \tag{4}$$

first- and second-derivative continuity at internal sub-interval boundaries,

$$S_i'(x_i) = S_{i+1}'(x_i), \quad S_i''(x_i) = S_{i+1}''(x_i), \quad i = 1, \ldots, n-1, \tag{5}$$

and zero second-derivative BCs at the ends,

$$S_1''(x_0) = 0, \quad S_n''(x_n) = 0. \tag{6}$$

These conditions yield a total of $N_{\text{eq}} = 2n + 2(n-1) + 2 = 4n$ equations, which is equal to the number of coefficients, $N_{\text{coef}} = n(k+1) = n\ (3+1) = 4n$. The spline coefficients are obtained by solving the system of linear equations in (4)–(6). A natural cubic spline is weakly local because, although the perturbation of a data point affects all the coefficients, in principle, this rapidly decays with distance from the data point.

Another type of cubic spline is the *cubic Hermite spline*, where second derivative continuity is not enforced at sub-interval boundaries. Instead, the coefficients of each segment are determined by the interpolation property in (4), and

$$S_i'(x_{i-1}) = m_{i-1}, \quad S_i'(x_i) = m_i, \quad i = 1, \ldots, n, \tag{7}$$

where m_i is the numerical derivative estimate of the slope at x_i. For example, using the 3-point central difference formula, the slopes at internal sub-interval boundaries ($i = 2, \cdots, n-1$) are $m_i = (y_{i+1} - y_{i-1})/(x_{i+1} - x_{i-1})$, and at the ends $m_0 = (y_1 - y_0)/(x_1 - x_0)$ and $m_n = (y_n - y_{n-1})/(x_n - x_{n-1})$. Because the coefficients of each segment are determined separately from all others, the cubic Hermite spline is strongly local.

A *clamped cubic spline* (not used in this work) is the same as the natural cubic spline except that the first derivatives are specified as BCs. In particular, (4) and (5) together with the boundary conditions $S_1'(x_0) = m_0$ and $S_n'(x_n) = m_n$ define the clamped cubic spline, where m_0 and m_n are the numerical derivative estimates of the slope at the left and right ends, respectively.

In a fashion similar to the cubic splines described, we define the odd-degree *clamped quintic* and *quintic Hermite splines*. Both are required to satisfy the interpolation property in (4).

The clamped quintic spline also satisfies C^4 continuity at sub-interval boundaries. Four additional BCs must be specified to close the system of equations. For the quintic Hermite spline, each segment's coefficients are determined from the values at the sub-interval boundaries and numerical derivative estimates for the first and second derivatives.

A direct extension of this discussion for odd-degree splines to the case of even-degree splines would impose the interpolation property in (4), C^{k-1} continuity at the sub-internal boundaries, and $k-1$ BCs to close the system of equations. We call even-degree splines constructed in this way 'naïve splines.' As their name suggests, it is inadvisable to use splines of this type due to the ill-posedness of the resulting equations for the spline coefficients[2]. Instead, even-degree splines should be constructed so that the knots are at the sub-interval mid-points [37], $\hat{x}_i = (x_{i-1} + x_i)/2$ for $i = 1, \cdots, n$. For example, the interpolation condition for a quartic spline becomes,

$$S_i(\hat{x}_i) = \hat{y}_i, \quad i = 1, \ldots, n, \tag{8}$$

and C^3 continuity at the internal sub-interval boundaries requires,

$$S_i^{(d)}(x_i) = S_{i+1}^{(d)}(x_i), \quad d = 0, \ldots, 3, \quad i = 1, \ldots, n-1. \tag{9}$$

Altogether there are $N_{\text{eq}} = n + 4(n-1) = 5n - 4$ equations, but $N_{\text{coef}} = n(k+1) = 5n$ coefficients need to be determined, which indicates four BCs should be specified. The *clamped quartic spline* is defined in this way using numerical estimates of the first and second derivatives at both ends ($x_0 = a$ and $x_n = b$) as BCs to close the system of equations.

Some of the splines described above are quite similar. As already mentioned, the natural cubic and clamped cubic splines only differ in BCs. Our computations show that they produce nearly identical results. Thus, we only present the natural cubic spline results here. We include results for a naïve quartic spline to illustrate the problems that can occur when naïve splines are used. A summary of the splines for which results are reported in this article is presented in table 1.

3.2. Computational cost of splines

IP implementations, such as the KIM-compliant EAM Dynamo driver [41], typically precompute and store the coefficients of each spline segment. The energy value and energy derivatives at a point are then computed easily using these coefficients. For example, using Horner's rule [42],

$$P(x) = (a_0 + x(a_1 + x(a_2 + \cdots + x(a_{m-1} + xa_m)\cdots),$$

for the evaluation of an mth-order polynomial at x, requires $2m$ flops. Thus, in principle, quartic and quintic splines require 33% and 66% more execution time than a cubic spline, respectively. However, in practice, today's general-purpose simulation codes (even the highly optimized ones) have enough additional overhead to make this performance difference largely inconsequential. To illustrate this point, we have timed MD simulations, using natural cubic and clamped quintic splines, that determine the equilibrium lattice constant at $T = 10$ K. The times are 2403.87 s and 2505.88 s, respectively. Thus, the run time for a quintic spline calculation is only 4.24% longer than the cubic spline computation[3].

[2] Here we refer to the mathematical definition, due to J Hadamard [40], of well-posed problems that require solutions to: (i) exist; (ii) be unique; and (iii) depend continuously on the problem data. In the case of the naïve even-degree splines, it is condition (iii) that fails, resulting in an extreme sensitivity of the solution to small perturbations in the data $\hat{x}_i$ and $\hat{y}_i$. See also the discussion of parabolic splines elsewhere [36].

[3] The tests were conducted on a machine with 32 kB L1, 256 kB L2, and 8 MB L3 caches and 3 GB RAM. The code used was the MiniMol MD and MS package [4]. If a more general-purpose MD code (e.g. LAMMPS) were used, we expect an even smaller difference in the timing values, due to the additional complexity of such codes.

Table 1. Summary of splines used in this article. 'Label' refers to the labels used in figures.

Spline	Label	Continuity	Locality	BCs		#Knots used to estimate derivatives
				Left end	Right end	
Natural Cubic	3N	C^2	weakly local	$S'' = 0$	$S'' = 0$	—
Cubic Hermite	3H	C^1	strongly local	—	—	5
Naïve Quartic	4*	C^3	nonlocal	S'	S', S'''	—
Clamped Quartic	4C	C^3	weakly local	S', S''	S', S''	—
Clamped Quintic	5C	C^4	weakly local	S', S''	S', S''	—
Quintic Hermite	5H	C^2	strongly local	—	—	7

Note: The BCs indicate the additional conditions specified at the spline ends. When not zero, the imposed values are either estimated numerically or computed from the interpolated function. The number of knots used for derivative estimates is applicable to the internal knots (fewer knots are used near the ends).

3.3. Spline interpolation of the modified Morse potential

To explore the behavior of the splines listed in table 1, we applied them to the modified Morse potential in (2). As discussed in section 1, IPs based on analytic functions are typically tabulated using one of two approaches ('energy only' or 'energy and derivative(s)'). In this article, we adopt the first approach ('energy only'). The energy function was interpolated over the range $r \in [0, 8.15]$ Å with equidistant knots at regular intervals. The derivatives of the IP were computed from the energy spline. The interpolation results are presented in figure 1, where the spline labels are given in table 1 and 'A' (short for Analytic) indicates values that are computed directly from (2). The graphs only show the region close to the minimum of the function ($r_0 = 2.5471$ Å) to highlight the differences.

Examining figure 1, we see that all splines (except naïve quartic and quintic Hermite) yield very good interpolation results of the modified Morse potential function and its derivatives up to an order that is one degree lower than the degree of the spline. (The first derivative curves are not shown because, like the energy curves, all the splines do a good job of matching the analytic function.) Although the second derivative of the cubic Hermite spline is discontinuous, this discontinuity is found to be negligibly small in comparison to its error in this case (see the inset of figure 1(b)). Cubic splines have piecewise constant third derivatives, as seen in figure 1(c), because, for third-degree splines, $S_i'''(x) = 6a_{i4}$, which is distinct for each segment (see (3)). A similar analysis applies to the fourth derivative of the quartic splines (because they are only C^3 continuous) and the fifth derivative of the two quintic splines (not shown in the figure). Note, in figure 1(e) the blue naïve quartic (4*) curve is discontinuous and alternates between large and small values surrounding the analytic fourth derivative. The amplitude of the alternations is extremely sensitive to the values used for the BCs of the spline due to the ill-posedness of the naïve quartic spline formulation. This ill-posedness is also the cause of this spline's third derivative oscillations about the analytic curve. Interestingly, the quintic Hermite spline shows, with increasing number of knots, a growing sensitivity to numerical noise in the derivative estimates used to compute its spline coefficients. This can be seen by comparing figure 1(d) (which uses 10 000 knots) and figure 1(f) (which uses only 500 knots).

As a quantitative measure of the accuracy of the interpolations, we computed the normalized root-mean-square deviations (NRMSD):

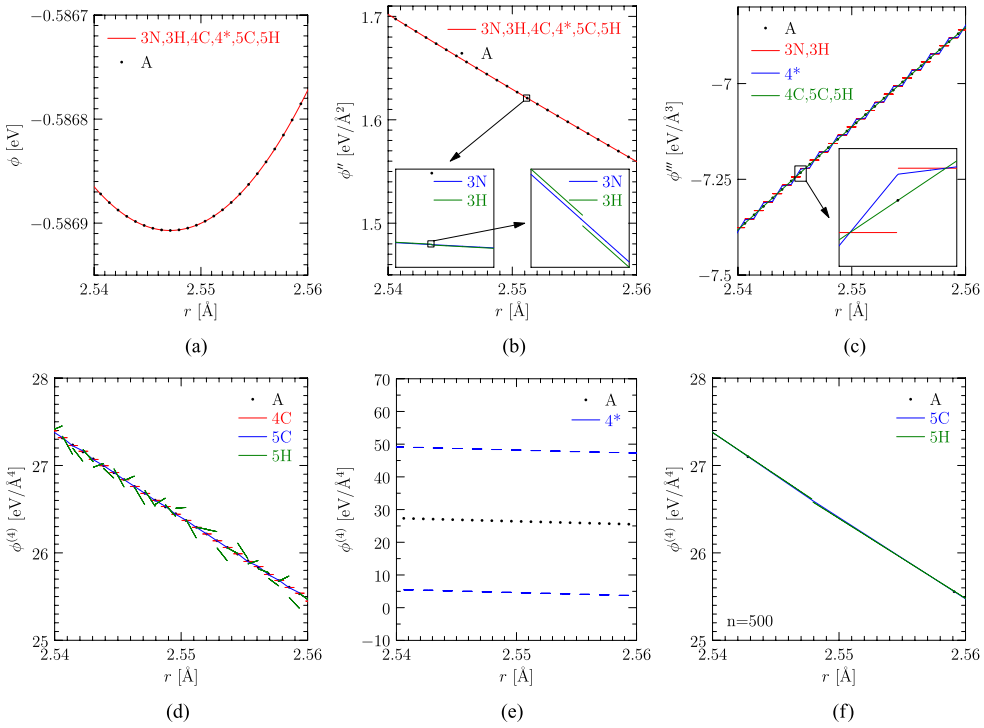

Figure 1. Spline interpolations of the modified Morse potential for copper. The number of interpolation data points is $n = 10\,000$, except for frame (f) where $n = 500$. (a) Energy values for all splines. (b) Second derivative for all splines. (c) Third derivative for all splines. (d) Fourth derivative for the clamped quartic and quintic splines. (e) Fourth derivative for the naïve quartic spline. (f) Fourth derivative for the clamped quartic and quintic splines with $n = 500$ data points.

$$e = \left[\frac{\int_a^b (f_{\text{spline}} - f_{\text{analytic}})^2 \, \mathrm{d}r}{\int_a^b (f_{\text{analytic}})^2 \, \mathrm{d}r} \right]^{1/2}, \tag{10}$$

where f_{spline} is the spline interpolation (which can be the function value, first derivative, etc) and f_{analytic} is the corresponding analytic value obtained from (2) or its derivatives; 10 000 data points are used to do the interpolations over the range $r \in [0, 8.15]$ Å, but in (10), let $a = 2.54$ Å and $b = 2.56$ Å, such that the range of r is consistent with that used in figure 1. The results are listed in table 2. As expected, clamped quartic and clamped quintic splines have the smallest NRMSD. Perhaps not surprisingly, given the discussion, the naïve quartic spline has larger NRMSD errors than the cubic splines.

3.4. Effect of noisy data on spline interpolation

In section 3.1, we discussed the locality of splines and listed our expectations for different splines in table 1. This can be an important consideration because the tabulated IP functions being interpolated are sometimes not smooth. To explore the effect of noise, we performed

Table 2. Normalized root-mean-square deviation.

Spline	ϕ	ϕ'	ϕ''	ϕ'''	$\phi^{(4)}$
Natural Cubic (3N)	3.3234×10^{-14}	7.8317×10^{-9}	3.9996×10^{-7}	8.6675×10^{-4}	—
Cubic Hermite (3H)	3.3235×10^{-14}	7.8319×10^{-9}	4.0007×10^{-7}	8.6679×10^{-4}	—
Clamped Quartic (4C)	1.7563×10^{-16}	1.5811×10^{-11}	3.1669×10^{-10}	3.8266×10^{-7}	8.7772×10^{-4}
Naïve Quartic (4*)	1.5246×10^{-13}	3.2462×10^{-8}	8.1550×10^{-7}	7.1508×10^{-4}	8.2501×10^{-1}
Clamped Quintic (5C)	1.7529×10^{-16}	1.5620×10^{-11}	2.5686×10^{-10}	1.7432×10^{-7}	1.5048×10^{-4}
Quintic Hermite (5H)	7.2104×10^{-16}	1.9921×10^{-11}	5.2351×10^{-9}	3.8273×10^{-6}	2.5560×10^{-3}
Natural Cubic (3N)[a]	3.3234×10^{-14}	7.8317×10^{-9}	3.9996×10^{-7}	8.6675×10^{-4}	—
Cubic Hermite (3H)[a]	3.3235×10^{-14}	7.8319×10^{-9}	4.0007×10^{-7}	8.6679×10^{-4}	—
Clamped Quartic (4C)[a]	1.7563×10^{-16}	1.5811×10^{-11}	3.1669×10^{-10}	3.8266×10^{-7}	8.7772×10^{-4}
Naïve Quartic (4*)[a]	4.7503×10^{-6}	10.114	25.408	2.2280×10^{4}	2.5705×10^{7}
Clamped Quintic (5C)[a]	1.7529×10^{-16}	1.5620×10^{-11}	2.5686×10^{-10}	1.7432×10^{-7}	1.5048×10^{-4}
Quintic Hermite (5H)[a]	7.2104×10^{-16}	1.9921×10^{-11}	5.2351×10^{-9}	3.8273×10^{-6}	2.5560×10^{-3}

[a] Denotes interpolation for noisy data (see section 3.4).

the same interpolations as in section 3.3, but with ϕ_{9900} (the function value at knot 9 900 located at $r = 8.0685$ Å) perturbed to equal that of its left neighbor (i.e. $\phi_{9900} := \phi_{9899}$). The results are presented in figure 2. The effect of the perturbation on the function value disappears approximately five knot-spacings away from the noise site for all splines, except the naïve quartic spline for which the effect persists and even amplifies as seen in figure 2(b). The second derivative of the natural cubic spline, the third derivative of the clamped quartic spline, and the fourth derivative of the clamped quintic spline are presented in figures 2(c)–(e), respectively. We see that the perturbation decays to zero at approximately 50 knot-spacings from the noise site in figures 2(c)–(e). The range of effect for lower-order derivatives is much smaller. Locality is even stronger for the Hermite splines, in which the coefficients in each segment depend only on the adjacent data points used to estimate the derivatives at the segment ends. For example, if a five-point central finite difference is used to estimate the derivatives, then a segment 3 knots away from the noise site will not be affected by the perturbation. The NRMSD values for the noisy data are shown in table 2. These data confirm that, except for the naïve quartic spline, all other splines studied in this work exhibit acceptable behavior.

4. Effect of interpolation on model predictions

To explore the effect of interpolation on the predictions of tabulated IPs that are based on analytic functional forms, we consider the case of a periodic 1D chain of N copper atoms at finite temperature. The atoms have nearest-neighbor interactions described by the modified Morse potential described in section 2. We compute the thermal expansion and finite-temperature elastic constant of the chain using both the quasi-harmonic (QH) approximation and MD simulations. To maintain focus on real effects that can be expected from well-posed spline interpolation, we delay the presentation of results for the ill-behaved naïve quartic spline to section 4.4, where we show just how badly such calculations can go wrong.

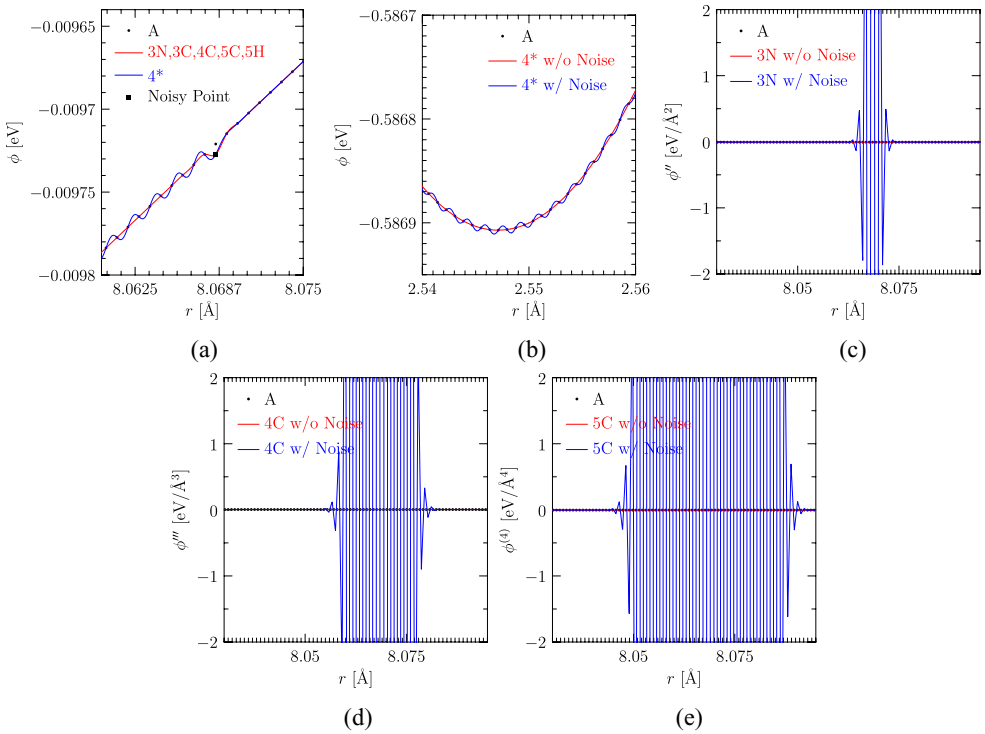

Figure 2. Interpolation of noisy data to test the locality of the splines. The number of interpolation data points is $n = 10\,000$. (a) Function values for all splines in the vicinity of noise site. (b) Function value for the naïve quartic spline far from the location of the noise. (c) Second derivative of the natural cubic spline. (d) Third derivative of the clamped quartic spline. (e) Fourth derivative of the clamped quintic spline.

4.1. Thermal expansion

The QH approximation for the thermomechanical properties of a 1D chain is described in Chapter 11 of another work [4]. For atoms interacting via a pair potential, the temperature dependence of the equilibrium lattice constant is obtained by solving[4],

$$\phi'(a) + \frac{k_B T}{2}\frac{\phi'''(a)}{\phi''(a)} = 0, \tag{11}$$

where a is the lattice constant at temperature T, and k_B is Boltzmann's constant. The resulting lattice constant versus temperature plots for the tabulated modified Morse potential with different interpolations and number of knots are presented in figure 3. As shown, the lattice constants computed from the clamped quartic and quintic spline potentials coincide well with those obtained using the analytic potential. However, the cubic splines show significant anomalies. Due to their discontinuous third-order derivatives, the $T(a)$ relationship defined

[4] In practice, we solve (11) for T to obtain $T = -2\phi'(a)\phi''(a)/[k_B\phi'''(a)]$. Then, we use this expression to compute (a, T) pairs over the range $a \in [2.5471, 2.7450]$ Å. These pairs are plotted in figure 3 to obtain the lattice constant versus temperature relationship.

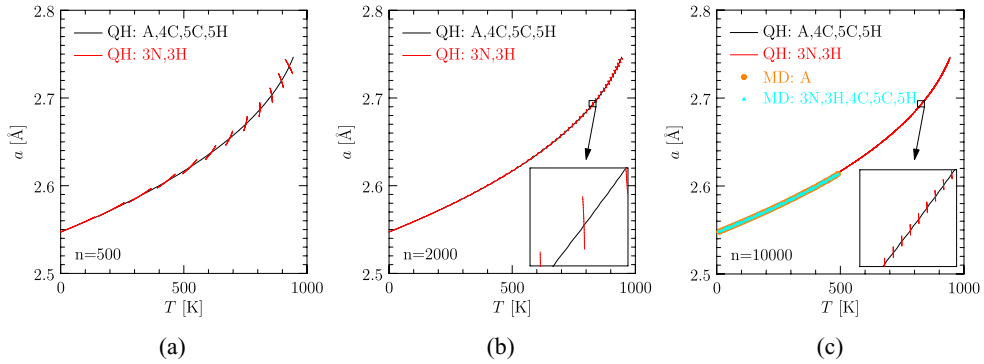

Figure 3. Equilibrium lattice constant as a function of temperature computed using the QH approximation with a tabulated modified Morse potential using different splines with (a) $n = 500$ knots, (b) $n = 2000$ knots, and (c) $n = 10\,000$ knots. In addition to the lattice dynamics QH approximation results, frame (c) also shows the MD results computed using the analytic and spline potentials.

by (11) is also discontinuous. This results in significant ranges of temperature over which the QH equilibrium lattice constant is *not defined* for the cubic splines. It is also interesting to note that for high temperatures where a QH equilibrium lattice constant is defined, we see a non-physical trend where the lattice constant rapidly decreases with increasing temperature. As can be seen in figure 3, these anomalies exist regardless of the number of knots used in the cubic spline.

Figure 3(c) also shows the results from MD simulations. To obtain these results, a series of finite chain ($N = 1000$) constant temperature simulations were performed to obtain the equilibrium lattice constant at different temperatures. Using a finite chain ensures that the average force (stress) in the chain is zero. Temperature was maintained using a Langevin thermostat with damping coefficient of 0.05 ps^{-1}. The equations of motion were integrated using a velocity-Verlet algorithm with a time step of $\Delta t = 1$ fs. To compute the lattice constant at a given temperature, the chain was evolved for 10^7 MD steps and the equilibrium chain length was computed as $L_{\text{eq}} = \langle r_N - r_1 \rangle$ (where $\langle \cdot \rangle$ indicates a time average). The equilibrium lattice constant follows as $a = L_{\text{eq}}/(N - 1)$. The temperature was then incremented by $\Delta T = 10$ K, and the calculation was repeated at the new temperature. For each calculation, the initial atom spacing was taken equal to the equilibrium lattice constant at the previous temperature to minimize stress waves in the chain.

The MD results for the analytic potential and the various spline potentials, shown in figure 3(c), are in excellent agreement with the analytic QH calculations.[5] This is not surprising because we saw in figure 1 that all well-behaved splines accurately represent the analytic results up to derivatives of second order, and higher-order derivatives (although non-smooth and/or discontinuous in some cases) provide a reasonably accurate representation of the corresponding analytic curves. The time averaged nature of equilibrium material properties computed via MD simulations provides a 'smeared' result that effectively smooths out the splines' higher-order derivatives. This is true as long as the temperature of the MD simulation is large enough so that more than just a couple of spline segments are sampled. We therefore do not expect to see differences due to interpolation in standard MD simulations as long as

[5] The MD data end at $T = 490$ K, just before the 1D chain starts to 'melt' (corresponding to individual atoms separating from the bulk of the chain).

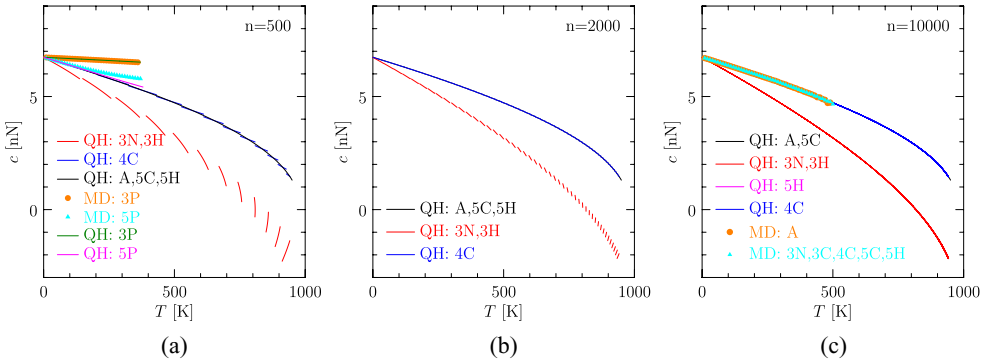

Figure 4. Stress-free spatial elastic constant (tangent modulus) as a function of temperature computed using the QH approximation with a tabulated modified Morse potential using different splines with (a) $n = 500$ knots, (b) $n = 2000$ knots, and (c) $n = 10\,000$ knots. In addition to the lattice dynamics QH approximation results, frame (a) also shows the QH and MD results computed using pure cubic and quintic polynomial potentials, and frame (c) also shows the MD results computed using the analytic and spline potentials.

the splines used have a sufficient number of knots to accurately represent, on average, how the energy and its derivative change with atomic separation. See section 4.3 for a discussion of what happens when MD simulations sample only one spline segment.

4.2. Finite-temperature elastic constants

The QH approximation for the temperature-dependent stress-free spatial elastic constant (tangent modulus) of a 1D chain of atoms interacting via a nearest-neighbor pair potential is [4],

$$c = a\left[\phi''(a) + \frac{k_B T}{2}\frac{\phi^{(4)}(a)\phi''(a) - (\phi'''(a))^2}{(\phi''(a))^2}\right], \tag{12}$$

where $a = a(T)$ is the stress-free equilibrium lattice constant at temperature T. The value of c as a function of temperature for the spline and analytic potentials is presented in figure 4. We see that the results obtained for the clamped quintic spline potential agree well with the analytic potential results, even when the number of knots is on the order of $n = 500$. The results obtained using the quintic Hermite spline potential agree with the analytic results when the number of data points is not too large (e.g. $n = 500$ and $n = 2000$). However, due to the increased sensitivity to numerical noise from the use of numerical differentiation, the results of the quintic Hermite spline degrade somewhat when $n = 10\,000$ data points are used. This is in agreement with figures 1(d) and (f), where the quintic Hermite spline predicts worse fourth-order derivative values with more data points. Thus, at least for the quintic Hermite spline, increasing the number of knots does not necessarily lead to better accuracy. The results for the clamped quartic spline are interesting. Although this spline is only C^3 continuous, it is able to follow the analytic curve, on average, and shows better results with more data points.

Unlike the higher-order quartic and quintic splines, the cubic splines (natural and Hermite) produce a quantitatively different behavior for the temperature-dependent elastic constant. In fact, they predict entirely the wrong initial slope at $T = 0$. Increasing the number

of knots decreases the discontinuity in the cubic spline results and the curve appears to become smoother (by reducing the error in the piecewise constant third derivative), but this cannot resolve the basic error because the elastic constant depends on the fourth derivative of ϕ, which is identically zero for cubic splines. Further, it is worth noting that at high temperatures the spline potentials give *qualitatively* incorrect results corresponding to a negative elastic constant.

The QH results for the stress-free elastic constant are compared with MD simulation results in figure 4(c) (the MD results in figure 4(a) are a special case that is discussed later). There are two common approaches to computing elastic constants in MD. The direct method in which the elastic constants are obtained from a stress-strain curve, and the stress or strain fluctuation method where they are obtained as a suitable ensemble average over the vibrating atoms [43–45]. We use the stress fluctuation method, which is performed under constant length L and constant temperature conditions. For the 1D case, the spatial elastic constant is [4]

$$c = \frac{1}{L}\left[2Nk_{\mathrm{B}}T + L\langle c^0\rangle - \frac{L^2}{k_{\mathrm{B}}T}\mathrm{Cov}(\sigma^{\mathrm{V,inst}}, \sigma^{\mathrm{V,inst}})\right], \tag{13}$$

where L is the length of the chain and Cov is the covariance operator. In (13), c^0 is the Born term given by

$$c^0 = \frac{1}{L}\sum_{i=1}^{N}\sum_{j=i+1}^{N}[\phi''(r_{ij})(r_{ij})^2 - \phi'(r_{ij})r_{ij}], \tag{14}$$

and $r_{ij} = x_j - x_i$. In (13), $\sigma^{\mathrm{V,inst}}$ is the potential part of the instantaneous stress,

$$\sigma^{\mathrm{V,inst}} = \frac{1}{L}\sum_{i=1}^{N}\sum_{j=i+1}^{N}\phi'(r_{ij})r_{ij}. \tag{15}$$

Equations (14) and (15) are written for the case of a finite chain of atoms. The expressions for a periodic chain are similar. The MD results are plotted in figure 4(c) and are seen to be in close agreement with the QH prediction using the analytic potential. As expected (see explanation in section 4.1), the MD results for all splines and knot numbers were in close agreement. Note that this includes the MD results obtained from the cubic splines, which do not agree with their QH counterparts.

4.3. Fitted spline interatomic potentials

So far, we have been concerned with IPs whose tabulated data are generated directly from an analytic functional form with parameters that have been fitted to reproduce various material properties. However, as mentioned in section 1, in some cases (which are becoming more common) an analytic form does not exist and the fitting procedure directly generates the tabulated data. In such cases, the number of data points is small[6], typically 10–30. This leads to a situation where, at least for low temperatures and in the worst-case scenario, only one spline segment is involved in a QH or MD computation.

[6] Fitting with more spline knots is difficult due to a lack of experimental and first principles data representative of a full range of interatomic spacings, and because over-fitting can occur [46, 47].

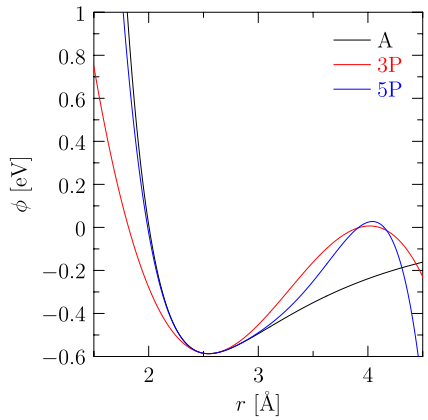

Figure 5. Pure cubic polynomial and quintic polynomial potentials fit to reproduce the Morse potential energy and its derivatives at the minimum r_0.

To illustrate what can happen in this situation, a pure cubic polynomial IP (3P) and a pure quintic polynomial IP (5P), shown in figure 5, were defined such that the pure cubic IP reproduces the modified Morse potential energy and its first and second derivatives at the pair equilibrium distance r_0 and the pure quintic IP (5P) reproduces the modified Morse potential energy and its first through third derivatives at r_0. These IPs were used to compute the QH and MD elastic constant and the results are presented in figure 4(a). Here, we see that the MD results always match the corresponding QH results (at least for low temperatures). This shows that when an IP is based on a spline interpolation and a small number of knots is used for fitting the IP, the results (MD as well as QH) obtained from such an IP cannot be expected to match those of an analytic IP to which it was fit. Thus, caution is warranted when creating a spline-based IP by directly fitting the spline data[7].

4.4. Results for the naïve quartic spine

We now return to the results for the naïve quartic spline. The temperature-dependent lattice constant and stress-free elastic constant curves, along with the corresponding analytic curves, are shown in figures 6(a)–(f), respectively.

The lattice constant curve computed using the naïve quartic spline potential shows reasonable agreement with that obtained from the analytic potential when the number of knots is 500 or 2000. However, when 10 000 knots are used, the spline results show anomalous behavior. In particular, there are temperature ranges over which there exist two equilibrium lattice constant values (see inset of figure 6(c)), which is non-physical and qualitatively different from results for the analytic potential. Finally, we see that the naïve quartic spline IP produces complete nonsense for the stress-free elastic constant, shown in figures 6(d)–(f). Note that in each frame there is just one curve for the naïve quartic that rapidly oscillates about the analytic curve.

[7] Note that refining such an IP by generating more knots from the original fitted spline will *not* improve its quality. We created cubic and quintic splines using 2000 data points from the pure cubic and pure quintic IPs. The QH and MD elastic constant values computed using these many-knotted spline IPs are the same as those obtained for the pure cubic and pure quintic IPs, respectively.

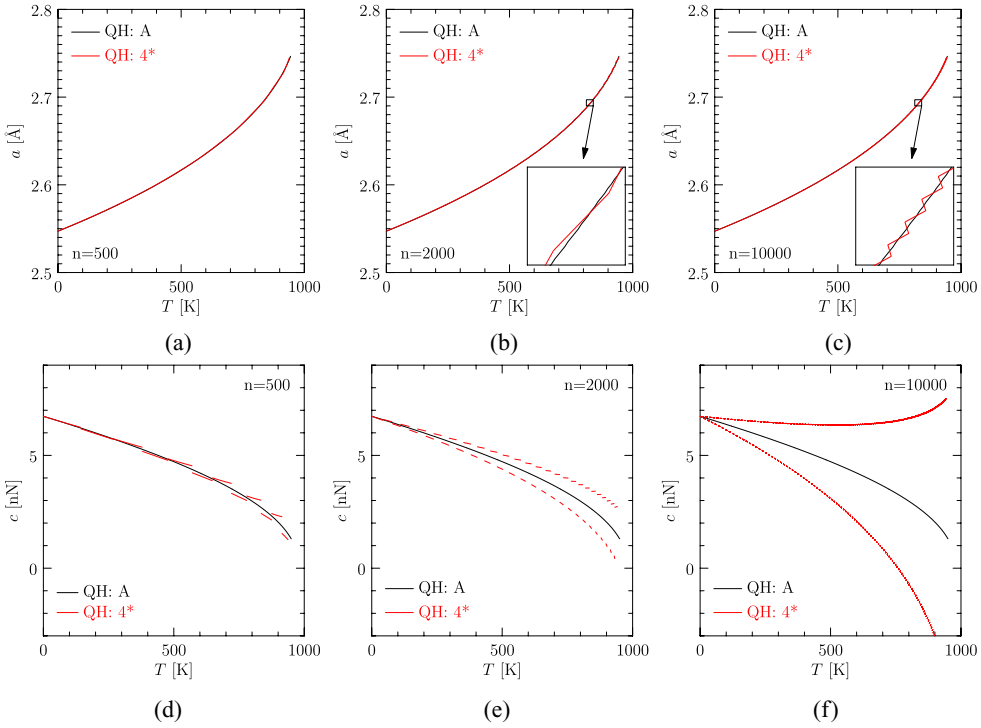

Figure 6. Temperature-dependent lattice constant (a)–(c) and stress-free elastic constant (d)–(f), computed using the QH approximation and the naïve quartic spline with $n = 500$, $n = 2000$, and $n = 10\,000$ data points.

5. Summary and conclusions

Interatomic potentials (IPs) that are used with MD codes are often tabulated and interpolated using polynomials or splines for reasons of efficiency and interoperability. In this article, we explored the effect this has on IPs and their predictions. First, we studied the ability of a variety of cubic, quartic, and quintic splines with different numbers of data points to reproduce a modified Morse potential for copper and its derivatives. An important issue is the 'locality' of the interpolation, which is related to how quickly the effect of noisy data decays. Based on our investigations, we can make the following general recommendations:

(a) Natural cubic, clamped quartic, and clamped quintic splines are able to reproduce the modified Morse potential and its derivatives up to an order of one degree lower than that of the spline. These splines exhibit an acceptable level of locality.
(b) Hermite splines by construction have strong locality. However, this comes at a cost of reduced continuity. In particular, the cubic and quintic Hermite splines have relatively poor second and fourth derivatives, respectively, and should be avoided when these are required.
(c) Even degree splines should not be formulated using the 'naïve approach,' which leads to low-quality interpolation and extreme sensitivity to changes in the data. Instead, the well-posed formulation, presented in section 3.1, should be used.

We also explored the effect of the spline interpolation on various properties of a 1D chain of copper atoms interacting via nearest-neighbor pairwise interactions, including thermal

expansion, and the stress-free spatial elastic constant at finite temperature. The properties were computed using both QH approximation expressions and MD simulations. We found that the MD results are insensitive to the interpolation used. This is not surprising because MD sampling smooths out the higher-order derivatives of the splines and all of the studied splines provide good average estimates for the pair potential energy and its derivatives. This conclusion is correct as long as an MD simulation samples an interpolated IP over multiple knots carrying *distinct* information. This can fail for IPs that are constructed by directly fitting the knot values of a coarse spline, even if a refined spline is constructed by adding knots between those of the original source spline through interpolation. In particular, if the MD sampling is limited to a single interval of the original source spline, then pathologies similar to those experienced by QH (summarized later) can also happen in MD as shown in section 4.3.

QH computations, which do not use sampling, show strong dependence on the interpolation method and number of knots. In particular, cubic splines cannot be used to compute the QH estimate for the finite-temperature elastic constant, which depends on the fourth derivative of the potential. As expected from our recommendations, the naïve quartic spline behaves quite poorly. The results for the elastic constant show that the accuracy obtained with the naïve quartic spline is significantly *reduced* with increasing knot numbers.

The sensitivity of QH expressions to interpolation is representative of a class of computational methods that make use of higher-order derivatives. These include: (i) lattice dynamics methods for computing vibrational spectra of crystals, phase transformations, and so on [48]; (ii) multi-scale methods that include phonon stability analysis [49, 50] or harmonic free energy expressions [51, 52]; and (iii) branch-following and bifurcation methods that require smooth second derivatives [53–55]. In all of these cases, either the analytic potential or an interpolation of sufficient order to ensure the required level of continuity must be used.

This discussion suggests that the interpolation used with a tabulated IP *must* be considered part of the potential definition. This is the philosophy adopted by the KIM project (https://openkim.org), which is an online resource for archiving and testing IPs [19, 20]. In KIM, a tabulated IP (i.e. a data file in standard format containing the functions defining the IP as a table of argument–value pairs) is called a 'parameterized model'. The computer routine that interpolates the tabulated data and computes the energy and forces for a collection of atoms is called a 'model driver'. Both of these together define the IP. Models and model drivers in KIM have unique identifiers that can be cited in publications to facilitate research reproducibility. As part of the KIM project, an application programming interface called the 'KIM API' [18] has been developed in consultation with the atomistic simulation community. The KIM API makes it possible to use KIM models with KIM-compliant simulation codes, such as LAMMPS, IMD, DL_POLY, and GULP.

Finally, given the importance of interpolation effects, and the requirement for higher-order splines for many lattice dynamics and related computations, we have developed two new quintic spline model drivers (clamped and Hermite) for pair functional potentials (EAM, FS, EMT), which have been uploaded to KIM [56, 57]. These model drivers use the original EAM DYNAMO parameter file format [58] and can be used seamlessly with any of the codes mentioned. For example, they can be used to replace the built-in cubic Hermite spline interpolation in LAMMPS.

Acknowledgments

The authors thank M Baskes and J Gale for helpful comments. The research was partly supported through the National Science Foundation (NSF) under grants No. PHY-0941493 and DMR-1408211.

References

[1] Jones J E 1924 *Proc. R. Soc.* A **106** 441–62
[2] Jones J E 1924 *Proc. R. Soc.* A **106** 463–77
[3] Morse P M 1929 *Phys. Rev.* **34** 57–64
[4] Tadmor E B and Miller R E 2011 *Modeling Materials: Continuum, Atomistic and Multiscale Techniques* (Cambridge: Cambridge University Press)
[5] Mishin Y, Farkas D, Mehl M J and Papaconstantopoulos D A 1999 *Phys. Rev.* B **59** 3393–407
[6] Ercolessi F and Adams J B 1994 *Europhys. Lett.* **26** 583–8
[7] Baskes M 2014 private communication
[8] Daw M S and Baskes M I 1984 *Phys. Rev.* B **29** 6443–53
[9] Daw M S 1989 *Phys. Rev.* B **39** 7441–52
[10] Daw M S, Foiles S M and Baskes M I 1993 *Mater. Sci. Rep.* **9** 251–310
[11] Finnis M and Sinclair J 1984 *Phil. Mag.* A **50** 45–55
[12] Jacobsen K W, Nørskov J K and Puska M J 1987 *Phys. Rev.* B **35** 7423–42
[13] Baskes M I 1992 *Phys. Rev.* B **46** 2727–42
[14] Mishin Y, Mehl M J and Papaconstantopoulos D A 2005 *Acta Mater.* **53** 4029–41
[15] Apostol F and Mishin Y 2011 *Phys. Rev.* B **83** 054116
[16] Ercolessi F 1994 Interatomic potentials: The glue model http://www.fisica.uniud.it/~ercolessi/potentials/
[17] Becker C A NIST Interatomic Potentials Repository Project http://www.ctcms.nist.gov/potentials/
[18] Elliott R S and Tadmor E B kim-api-v1.7.0 https://openkim.org/kim-api/
[19] Tadmor E B, Elliott R S, Sethna J P, Miller R E and Becker C A 2011 *JOM* **63** 17
[20] Tadmor E B, Elliott R S, Phillpot S R and Sinnott S B 2013 *Curr. Opin. Solid State Mater. Sci.* **17** 298–304
[21] Open Knowledgebase of Interatomic Models (KIM) Website https://openkim.org
[22] Plimpton S 1995 *J. Comput. Phys.* **117** 1–19
[23] Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) http://lammps.sandia.gov
[24] Stadler J, Mikulla R and Trebin H R 1997 *Int. J. Mod. Phys.* C **8** 1131–40
[25] IMD: The ITAP Molecular Dynamics Program http://imd.itap.physik.uni-stuttgart.de
[26] Smith W and Forester T R 1996 *J. Mol. Graph.* **14** 136–41
[27] DL_POLY Molecular Simulation Package http://www.stfc.ac.uk/SCD/44516.aspx
[28] Gale J D 1997 *J. Chem. Soc. Faraday Trans.* **93** 629–37
[29] GULP website http://nanochemistry.curtin.edu.au/gulp
[30] Bowman J M 1978 *J. Chem. Phys.* **68** 608–10
[31] Brenner D W, Shenderova O A, Harrison J A, Stuart S J, Ni B and Sinnott S B 2002 *J. Phys.: Condens. Matter* **14** 783–802
[32] Gale J and Buckley A 2014 private communication
[33] MacDonald R A and MacDonald W M 1981 *Phys. Rev.* B **24** 1715–24
[34] Tadmor E B Modified Morse pair potential for copper due to MacDonald and MacDonald https://openkim.org/cite/MO_034823476734_000
[35] LeVeque R J 2007 *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems* (Philadelphia: SIAM)
[36] de Boor C 2001 *A Practical Guide to Splines* revised edn (New York: Springer)
[37] Ahlberg J H, Nilson E N and Walsh J L 1967 *The Theory of Splines and Their Applications* (New York: Academic)
[38] Judd K L 1998 *Numerical Methods in Economics* (Cambridge, MA: MIT Press)
[39] Cheney W and Kincaid D 2008 *Numerical Mathematics and Computing* 6 edn (Belmont, CA: Thomson Brooks/Cole)
[40] Hadamard J 1902 *Princeton Univ. Bull.* **13** 49–52
[41] Elliott R S EAM Model Driver with cubic Hermite spline interpolation https://openkim.org/cite/MD_120291908751_001
[42] Horner W G 1819 *Phil. Trans. R. Soc.* **109** 308–35
[43] Ray J R, Moody M C and Rahman A 1986 *Phys. Rev.* B **33** 895–9
[44] Ray J R 1988 *Comput. Phys. Rep.* **8** 109–51
[45] Gao G, Van Workum K, Schall J D and Harrison J A 2006 *J. Phys.: Condens. Matter* **18** S1737–50

[46] Payne M, Robertson I, Thomson D and Heine V 1996 *Phil. Mag.* B **73** 191–9
[47] Robertson I, Heine V and Payne M 1993 *Phys. Rev. Lett.* **70** 1944–9
[48] Dove M T 2005 *Introduction to Lattice Dynamics* (Cambridge: Cambridge University Press)
[49] Dobson M, Elliott R S, Luskin M and Tadmor E B 2007 *J. Comput.-Aided Mater. Des.* **14** 219–37
[50] Sorkin V, Elliott R S and Tadmor E B 2014 *Model. Simul. Mater. Sci. Eng.* **22** 055001
[51] Tadmor E B, Legoll F, Kim W K, Dupuy L M and Miller R E 2013 *Appl. Mech. Rev.* **65** 010803
[52] Kim W K, Luskin M, Perez D, Voter A F and Tadmor E B 2014 *J. Mech. Phys. Solids* **63** 94–112
[53] Elliott R S, Triantafyllidis N and Shaw J A 2006 *J. Mech. Phys. Solids* **54** 161–92
[54] Elliott R S, Shaw J A and Triantafyllidis N 2006 *J. Mech. Phys. Solids* **54** 193–232
[55] Pattamatta S, Elliott R S and Tadmor E B 2014 *Proc. Natl Acad. Sci. USA* **11** E1678–86
[56] Wen M EAM Model Driver with clamped quintic spline interpolation https://openkim.org/cite/MD_532469991695_000
[57] Wen M EAM Model Driver with quintic Hermite spline interpolation https://openkim.org/cite/MD_029719603993_000
[58] DYNAMO file formats in LAMMPS http://lammps.sandia.gov/doc/pair_eam.html