J Nonlinear Sci (2010) 20: 241–275
DOI 10.1007/s00332-009-9057-y

Journal of Nonlinear Science

Finite-Temperature Coarse-Graining of One-Dimensional Models: Mathematical Analysis and Computational Approaches

X. Blanc · C. Le Bris · F. Legoll · C. Patz

Received: 25 June 2008 / Accepted: 23 October 2009 / Published online: 20 February 2010

Abstract We present a possible approach for the computation of free energies and ensemble averages of one-dimensional coarse-grained models in materials science. The approach is based upon a thermodynamic limit process, and makes use of ergodic theorems and large deviations theory. In addition to providing a possible efficient

Communicated by A. Mielke.

X. Blanc (✉)
Laboratoire J.-L. Lions, Université Pierre et Marie Curie, Boîte courrier 187, 75252 Paris Cedex 05, France
e-mail: blanc@ann.jussieu.fr
e-mail: xavier.blanc@cea.fr

C. Le Bris
CERMICS, École Nationale des Ponts et Chaussées, Université Paris-Est, 6 et 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2, France
e-mail: lebris@cermics.enpc.fr

F. Legoll
Institut Navier, LAMI, École Nationale des Ponts et Chaussées, Université Paris-Est, 6 et 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2, France
e-mail: legoll@lami.enpc.fr

X. Blanc · C. Le Bris · F. Legoll
INRIA Rocquencourt, MICMAC Team-Project, Domaine de Voluceau, B.P. 105, 78153, Le Chesnay Cedex, France

C. Patz
Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, 10117 Berlin, Germany
e-mail: patz@wias-berlin.de

computational strategy for ensemble averages, the approach allows for assessing the accuracy of approximations commonly used in practice.

Keywords Coarse-grained models · Materials science · Thermodynamic limit · Canonical averages · Free energies · Law of large numbers · Large deviations theory

1 Introduction

Computing canonical averages is a standard task of computational materials science. Consider an atomistic system consisting of N particles, at positions $u = (u^1, \ldots, u^N) \in \mathbb{R}^{3N}$. Provide this system with an energy

$$E(u) = E\big(u^1, \ldots, u^N\big). \tag{1.1}$$

The finite temperature thermodynamical properties of the material are obtained from canonical ensemble averages,

$$\langle A \rangle = \frac{\int_{\Omega^N} A(u) \exp(-\beta E(u))\, du}{\int_{\Omega^N} \exp(-\beta E(u))\, du}, \tag{1.2}$$

where $\Omega \subset \mathbb{R}^3$ is the macroscopic domain where the positions u^i vary, A is the observable of interest, and $\beta = 1/(k_B T)$ is the inverse temperature (Deák et al. 2000). The denominator of (1.2) is denoted by Z and called the *partition function*. The major computational difficulty in (1.2) is of course the N-fold integrals, where N, the number of particles, is extremely large. Indeed, for integrals of the type (1.2) to be quantitatively meaningful in practice, N does not need to approach the Avogadro number, but still needs to be extremely large (10^5, say).

The three dominant computational approaches for the evaluation of (1.2) are Monte Carlo methods, Markov chains methods, and molecular dynamics methods respectively (see, e.g. Cancès et al. 2007 for a review of sampling methods of the canonical ensemble, along with a theoretical and numerical comparison of their performances for molecular dynamics). In the present article, we use the latter type of methods, and more precisely the overdamped Langevin dynamics (also called biased random walk). The ensemble average (1.2) is calculated as the long-time average

$$\langle A \rangle = \lim_{T \to +\infty} \frac{1}{T} \int_0^T A(u_t)\, dt \tag{1.3}$$

along the trajectory generated by the stochastic differential equation

$$du_t = -\nabla_u E(u_t)\, dt + \sqrt{2/\beta}\, dB_t. \tag{1.4}$$

It is often the case that the observable A actually does not depend on the positions u^i of *all* the atoms, but only on *some* of them. Think, for instance, of nanoindentation: we are especially interested in the positions of the atoms below the indenter, in the forces applied on these atoms, ... Our aim is to design a numerical method to compute

the canonical averages of such observables in a more efficient way than the general strategy (1.3)–(1.4). This will also enable us to assess the validity of other approaches, as compared to ours.

The *Quasicontinuum Method* (QCM) is a commonly used example of approaches that allow for the calculation of the specific ensemble averages discussed above. In its original version, the method was focused on the zero temperature setting. It was originally introduced in (Tadmor and Phillips 1996; Tadmor et al. 1996), and then further developed in (Knap and Ortiz 2001; Miller et al. 1998; Miller and Tadmor 2002; Shenoy et al. 1998, 1999; Tadmor et al. 1999). It has been studied mathematically in, e.g., (Anitescu et al. 2009; Arndt and Griebel 2005; Arndt and Luskin 2008; Blanc et al. 2002, 2005, 2007a; Dobson and Luskin 2008, 2009a, 2009b; Dobson et al. 2009a, 2009b; E and Ming 2004, 2007; Lin 2003, 2007; Ortner and Süli 2008). See (Blanc et al. 2007b; Legoll 2009) for recent reviews. An extension of the original idea has recently been developed in (Dupuy et al. 2005) and carries through to the finite-temperature case, considered in the present article. See also (Curtarolo and Ceder 2002; LeSar et al. 1989) for prior studies developing ideas in the same vein.

Let us briefly detail the bottom line of coarse-graining strategies for the computation of canonical averages. For simplicity of exposition, we let the atoms vary in $\Omega = \mathbb{R}^3$. The idea is to subdivide the particles of the system into two subsets. The first subset consists of the so-called *representative atoms* (abbreviated in the QCM terminology as *repatoms*, with positions henceforth denoted by u_r). The second subset is that of atoms that are eliminated in the coarse-grained procedure. Their positions are denoted by u_c. We *assume* that the observable considered only depends on the positions u_r of the repatoms, not on those of the other atoms, u_c. More precisely, one writes

$$u = \big(u^1, \ldots, u^N\big) = (u_r, u_c), \quad u_r \in \mathbb{R}^{3N_r},\ u_c \in \mathbb{R}^{3N_c},\ N = N_r + N_c,$$

and our aim is to compute (1.2) for an observable A that only depends on u_r:

$$\langle A \rangle = Z^{-1} \int_{\mathbb{R}^{3N}} A(u_r) \exp\big(-\beta E(u)\big)\, du. \tag{1.5}$$

We observe that, owing to our assumption on A,

$$\int_{\mathbb{R}^{3N}} A(u_r) \exp\big(-\beta E(u)\big)\, du = \int_{\mathbb{R}^{3N_r}} du_r\, A(u_r) \int_{\mathbb{R}^{3N_c}} \exp\big(-\beta E(u_r, u_c)\big)\, du_c,$$

and likewise

$$Z = \int_{\mathbb{R}^{3N}} \exp\big(-\beta E(u)\big)\, du = \int_{\mathbb{R}^{3N_r}} du_r \int_{\mathbb{R}^{3N_c}} \exp\big(-\beta E(u_r, u_c)\big)\, du_c.$$

Introducing the coarse-grained potential (also called free energy)

$$E_{\mathrm{CG}}(u_r) := -\frac{1}{\beta} \ln\left[\int_{\mathbb{R}^{3N_c}} \exp\big(-\beta E(u_r, u_c)\big)\, du_c\right], \tag{1.6}$$

the expression (1.5) reads

$$\langle A\rangle = Z_r^{-1}\int_{\mathbb{R}^{3N_r}} A(u_r)\exp\bigl(-\beta E_{\mathrm{CG}}(u_r)\bigr)\,du_r, \tag{1.7}$$

with $Z_r = \int_{\mathbb{R}^{3N_r}} \exp(-\beta E_{\mathrm{CG}}(u_r))\,du_r$. Under appropriate conditions ensuring ergodicity of the system, the integral (1.7) is in turn computed from

$$\langle A\rangle = \lim_{T\to+\infty}\frac{1}{T}\int_0^T A\bigl(u_r(t)\bigr)\,dt$$

with

$$du_r = -\nabla_{u_r}E_{\mathrm{CG}}(u_r)\,dt + \sqrt{2/\beta}\,dB_t. \tag{1.8}$$

Simulating the dynamics (1.8) is a less computationally demanding task than simulating (1.4), owing to the reduced dimension N_r. This simplification comes at a price: calculating the coarse-grained free energy (1.6).

Remark 1 The present work concentrates on the computation of ensemble averages and free energies, using coarse-grained models. Practice shows that the same coarse-graining paradigm is used to simulate *actual coarse-grained dynamics* at finite temperature. We will not go in this direction, as the physical relevance of the latter approach is unclear to us.

In order to approximate the free energy (1.6), state-of-the-art finite temperature methods perform a Taylor expansion of the position of the eliminated atoms u_c. In this Taylor expansion, a linear interpolation and a harmonic approximation of the positions of the atoms are successively performed. More precisely, given the positions u_r of the repatoms, some "reference" positions $\overline{u_c}(u_r)$ of the eliminated atoms are first determined by linear interpolation between two (or more) adjacent repatoms. Then it is postulated that

$$u_c = \overline{u_c}(u_r) + \xi_c$$

where the perturbation ξ_c is small. The energy is then calculated from a Taylor expansion truncated at second order:

$$E(u_r,u_c) = E\bigl(u_r,\overline{u_c}(u_r)+\xi_c\bigr) \approx \widetilde{E}(u_r,u_c)$$

with

$$\widetilde{E}(u_r,u_c) := E\bigl(u_r,\overline{u_c}(u_r)\bigr) + \frac{\partial E}{\partial u_c}\bigl(u_r,\overline{u_c}(u_r)\bigr)\cdot\xi_c + \frac{1}{2}\xi_c\cdot\frac{\partial^2 E}{\partial u_c^2}\cdot\xi_c. \tag{1.9}$$

It follows (we skip the details of the argument and refer to the bibliography pointed out above for further details) that $E_{\mathrm{CG}}(u_r)$ is approximated by

$$E_{\mathrm{HA}}(u_r) = -\frac{1}{\beta}\ln\int_{\mathbb{R}^{3N_c}}\exp\bigl(-\beta\widetilde{E}(u_r,u_c)\bigr)\,du_c, \tag{1.10}$$

which is *analytically* computable. Without such simplifying assumptions, the actual computation of E_{CG} for practical values of N_r and N_c seems undoable. The approach has proven efficient. Reportedly, it satisfactorily treats three-dimensional problems of large size. However, from the mathematical standpoint, it is an open question to evaluate the impact of the above couple of approximations (reference positions $\overline{u_c}(u_r)$ defined by a linear interpolation, followed by an harmonic expansion around these positions). The purpose of the present article is to present an approach that, in simple cases and under specific assumptions, allows for a quantitative assessment of the validity and limits of the above couple of approximations.

Our approach is based on a thermodynamic limit. It was first outlined in (Patz 2009) for the special case of harmonic interactions. The approach is *exact* in the limit of an infinite number of eliminated atoms and, therefore, valid when this number N_c is large as compared to the number N_r of representative atoms that are kept explicit in the coarse-grained model. This regime, after all, is the regime that all effective coarse-graining strategies should target. In short, the consideration of the asymptotic limit $N_c \to +\infty$ makes tractable a computation which is not tractable for finite N_c (unless simplifications, as those mentioned above, are performed). We do not claim for originality in our theoretical considerations on the thermodynamic limit of the free energy of atomistic systems. We provide them here for consistency. However, our specific use of such theoretical considerations as a computational strategy for approximating coarse-grained ensemble averages in computational materials science seems, to the best of our knowledge, new. We were not able to find any comparable endeavor in the existing literature in which we have access.

Let us conclude this introduction by briefly describing our approach. Assume for simplicity that there is only one repatom: $N_r = 1$. Our idea to compute $\langle A \rangle$ in (1.5) is to change variables, that is, introduce $y = (y_1, \ldots, y_N) = \Phi(u)$, and recast (1.5) as

$$\langle A \rangle = \int_{\mathbb{R}^{3N}} A\left(\frac{1}{N}\sum_{i=1}^{N} y_i\right) \nu(y)\, dy$$

for some probability density $\nu(y)$ (see (2.5) below for an explicit example). We next recognize $\langle A \rangle$ as the expectation value $\mathbb{E}[A(\frac{1}{N}\sum_{i=1}^{N} Y_i)]$, where $Y = (Y_1, \ldots, Y_N)$ are random variables distributed according to the probability ν. A law of large numbers provides the limit of $\langle A \rangle$ when $N \to +\infty$ (which corresponds to $N_c \to +\infty$, since $N_r = 1$). The rate of convergence may also be evaluated using the central limit theorem.

Note that the above approach bypasses the calculation of the free energy E_{CG} to compute the ensemble average (1.5). Our strategy is hence different from the one we described above, which is based on using the formula (1.7).

It is also interesting to try and evaluate E_{CG} in the same regime. First, it is to be remarked that E_{CG} scales linearly with the number N_c of eliminated atoms. The relevant quantity is hence the free energy *per particle*

$$F_\infty(u_r) := \lim_{N_c \to +\infty} \frac{1}{N_c} E_{\mathrm{CG}}(u_r). \tag{1.11}$$

We now observe that $N_c F_\infty$ is not necessarily a good approximation of E_{CG}, for large N_c, even if F_∞ is a good approximation of E_{CG}/N_c. It is not clear to us how

to use F_∞, or the probability measure $Z_{N_c}^{-1}\exp(-\beta N_c F_\infty)$, to compute in an efficient manner an approximation of the average $\langle A\rangle$ (see Remark 8 below). Nevertheless, computing F_∞ turns out to be also interesting, but for a different reason. This energy is indeed related to the coarse-grained constitutive law of the material at finite temperature, as we explain at the end of Sect. 2.2 (see Derrida 2007; Olla 2007 for related approaches).

We develop our approach in the one-dimensional setting, for simple cases of pair interactions. We first consider nearest neighbor (NN) interactions. In this case, we develop a computational strategy to approximate ensemble averages (see Sect. 2.1), and we next address the computation of free energies (see Sect. 2.2). Numerical considerations are collected in Sect. 2.3.

We next turn to next-to-nearest neighbor interactions, traditionally abbreviated as NNN. For this model, we focus on the computation of ensemble averages (see Sect. 3.1). As explained in Sect. 3.2, more complicated types of interaction potentials and "essentially one-dimensional systems" (including polymer chains) may be treated likewise, although we do not pursue in this direction.

A similar interpretation of ensemble averages as the one presented here, using a Markov chain formalism, *should* lead to an analogous strategy for two-dimensional systems. Some preliminary developments, not included in the present article, already confirm this. However, definite conclusions are yet to be obtained, both on the formal validity of the approach and on the best possible numerical efficiency accomplished. The fact that the two-dimensional case is much more difficult than the one-dimensional case is corroborated by the literature on this subject: only very simple cases, such as spin systems or harmonic interactions (with zero equilibrium length) are known to have explicit solutions in this context (see the reviews Baxter 1982; Presutti 2008). We therefore prefer to postpone considerations on the two-dimensional situation until a future publication.

2 The Nearest Neighbor (NN) Case

As mentioned in the Introduction, our approach is based on the asymptotic limit $N \longrightarrow +\infty$. We therefore first rescale the problem with the interatomic distance h, such that $Nh = L = 1$ (see Remark 2 below). The atomistic energy (1.1) in the *rescaled* NN case writes

$$E\left(u^1,\dots,u^N\right) = \sum_{i=1}^{N} W\left(\frac{u^i - u^{i-1}}{h}\right), \tag{2.1}$$

where $W : \mathbb{R} \to \mathbb{R}$ is an interatomic potential (see (2.37) for a precise example, and theorems below for precise assumptions on W). We now impose $u^0 = 0$ to avoid translation invariance, and consider that only atoms 0 and N are repatoms, while all the other atoms $i = 1,\dots,N-1$ are eliminated in the coarse-graining procedure (see Fig. 1). Our argument can be straightforwardly adapted to treat the case of $N_r > 2$ repatoms (see Fig. 2 and the end of Sect. 2.1).

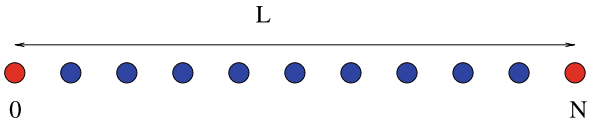

Fig. 1 (Color online) We isolate a segment between two consecutive repatoms (*in red*). All atoms in-between (*in blue*) are eliminated in the coarse-graining procedure

Fig. 2 (Color online) The repatoms (*in red*) are explicitly treated, all other atoms (*in blue*) being eliminated in the coarse-graining procedure

In this simple situation, the average (1.5) reads

$$\langle A\rangle_N = Z^{-1}\int_{\mathbb{R}^N} A\big(u^N\big)\exp\left(-\beta\sum_{i=1}^{N} W\left(\frac{u^i-u^{i-1}}{h}\right)\right)du^1\cdots du^N, \tag{2.2}$$

where we have explicitly mentioned the dependence of $\langle A\rangle$ upon N using a subscript.

Remark 2 Instead of the energy (2.1), we can work with the energy

$$\widetilde{E}\big(u^1,\ldots,u^N\big)=\sum_{i=1}^{N} W\big(u^i-u^{i-1}\big),$$

for an interatomic potential W that reaches its minimum at a value $\bar{x}$ independent of h (say $\bar{x}=1$). In that case, as $N\to\infty$, the mean size of the system diverges, and we thus expect averages of the type (1.5) to diverge as well: for instance,

$$\lim_{N\to\infty}\widetilde{Z}^{-1}\int u^N\exp\big(-\beta\widetilde{E}(u)\big)\,du=\infty.$$

With that scaling, the relevant quantity to consider is

$$\widetilde{Z}^{-1}\int A\left(\frac{u^N}{N}\right)\exp\big(-\beta\widetilde{E}(u)\big)\,du$$

which is exactly (2.2).

We introduce

$$y_i:=\frac{u^i-u^{i-1}}{h},\quad i=1,\ldots,N, \tag{2.3}$$

and next remark that

$$u^N=h\sum_{i=1}^{N} y_i=\frac{1}{N}\sum_{i=1}^{N} y_i. \tag{2.4}$$

Thus, the average (2.2) reads

$$\langle A\rangle_N = Z^{-1}\int_{\mathbb{R}^N} A\left(\frac{1}{N}\sum_{i=1}^N y_i\right)\exp\left(-\beta\sum_{i=1}^N W(y_i)\right)dy^1\cdots dy^N, \tag{2.5}$$

where now

$$Z = \int_{\mathbb{R}^N}\exp\left(-\beta\sum_{i=1}^N W(y_i)\right)dy^1\cdots dy^N.$$

Remark 3 In (2.2), we let the variables u^i vary on the whole real line. We do not constrain them to obey $u^{i-1}\le u^i$, which encodes the fact that nearest neighbors remain nearest neighbors. The argument provided here and below carries through when this constraint is accounted for: we just need to replace the interaction potential W by

$$W_c(y) = \begin{cases} W(y) & \text{when } y\ge 0,\\ +\infty & \text{otherwise.}\end{cases} \tag{2.6}$$

Likewise, we could also impose that all the u^i stay in a given macroscopic segment. If they are ordered increasingly, it is enough to impose this constraint on u^0 and u^N. This is again a simple modification of our argument.

2.1 Limit of the Average

It is evident from the expression (2.5) that

$$\langle A\rangle_N = \mathbb{E}\left[A\left(\frac{1}{N}\sum_{i=1}^N Y_i\right)\right]$$

for independent identically distributed (i.i.d.) random variables Y_i, sharing the law $z^{-1}\exp(-\beta W(y))\,dy$, with $z=\int_{\mathbb{R}}\exp(-\beta W(y))\,dy$. A simple computation thus gives the following result.

Theorem 1 *Assume that $A:\mathbb{R}\longrightarrow\mathbb{R}$ is continuous, that for some $p\ge 1$, there exists a constant $C>0$ such that*

$$\forall y\in\mathbb{R},\quad \big|A(y)\big|\le C\big(1+|y|^p\big), \tag{2.7}$$

and that

$$\int_{\mathbb{R}}\big(1+|y|^p\big)\exp\big(-\beta W(y)\big)\,dy<+\infty. \tag{2.8}$$

Introduce y^* and σ defined by

$$y^* = z^{-1}\int_{\mathbb{R}} y\exp\big(-\beta W(y)\big)\,dy\quad \textit{and} \tag{2.9}$$

$$\sigma^2 = z^{-1} \int_{\mathbb{R}} \big(y - y^*\big)^2 \exp\big(-\beta W(y)\big)\, dy,$$

with $z = \int_{\mathbb{R}} \exp(-\beta W(y))\, dy$. *Then*

$$\lim_{N \to +\infty} \langle A \rangle_N = A\big(y^*\big). \tag{2.10}$$

In addition, if A *is* C^2 *and if* (2.7)–(2.8) *hold with* $p = 2$, *then* $\sigma < \infty$ *and*

$$\langle A \rangle_N = \langle A \rangle_N^{\text{approx}} + o\left(\frac{1}{N}\right), \tag{2.11}$$

with

$$\langle A \rangle_N^{\text{approx}} := A\big(y^*\big) + \frac{\sigma^2}{2N} A''\big(y^*\big). \tag{2.12}$$

The proof of (2.10) is a direct application of the law of large numbers, and that of (2.11) is an application of the central limit theorem. We skip them. The following considerations, for more regular observables A, indeed contain the ingredients for proving (2.10)–(2.11), simply by truncating the expansion at first order.

Remark 4 Note that other growth assumptions on A are possible, along with corresponding assumptions on W. For instance, if A satisfies $|A(y)| \le C \exp(|y|)$ for some C and if W is such that $\exp(|y| - \beta W(y))$ is integrable, then (2.10) holds.

If A is more regular than stated in Theorem 1, then it is of course possible to proceed further in the expansion of $\langle A \rangle_N$ in powers of $1/N$. Indeed, assume for instance that A is C^6, that $A^{(6)}$ is globally bounded and that (2.7)–(2.8) hold with $p = 6$. Then

$$\begin{aligned} A\left(\frac{1}{N}\sum_{i=1}^N Y_i\right) &= A\left(y^* + \frac{1}{N}\sum_{i=1}^N D_i\right), \\ &= A\big(y^*\big) + A'\big(y^*\big)\frac{1}{N}\sum_{i=1}^N D_i + \frac{1}{2}A''\big(y^*\big)\left(\frac{1}{N}\sum_{i=1}^N D_i\right)^2 \\ &\quad + \frac{1}{6}A^{(3)}\big(y^*\big)\left(\frac{1}{N}\sum_{i=1}^N D_i\right)^3 + \frac{1}{24}A^{(4)}\big(y^*\big)\left(\frac{1}{N}\sum_{i=1}^N D_i\right)^4 \\ &\quad + \frac{1}{5!}A^{(5)}\big(y^*\big)\left(\frac{1}{N}\sum_{i=1}^N D_i\right)^5 + \frac{1}{6!}A^{(6)}(\xi)\left(\frac{1}{N}\sum_{i=1}^N D_i\right)^6, \end{aligned}$$

where $D_i = Y_i - y^*$ and ξ lies between y^* and $(1/N)\sum Y_i$. We now take the expectation value of this equality. Let us introduce

$$\mathcal{A}_N := A\big(y^*\big) + \frac{1}{2}A''\big(y^*\big)\frac{1}{N}\mathbb{E}\big(D_1^2\big) + \frac{1}{6}A^{(3)}\big(y^*\big)\frac{1}{N^2}\mathbb{E}\big(D_1^3\big)$$

$$+\frac{1}{24}A^{(4)}(y^*)\left(\frac{1}{N^3}\mathbb{E}(D_1^4)+\frac{N-1}{N^3}\big(\mathbb{E}(D_1^2)\big)^2\right)$$
$$+\frac{1}{5!}A^{(5)}(y^*)\left(\frac{1}{N^4}\mathbb{E}(D_1^5)+\frac{N-1}{N^4}\mathbb{E}(D_1^2)\mathbb{E}(D_1^3)\right). \tag{2.13}$$

Then

$$\big|\langle A\rangle_N-\mathcal{A}_N\big|\le\frac{1}{6!}\big\|A^{(6)}\big\|_{L^\infty}\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^N D_i\right)^6\right]. \tag{2.14}$$

We now use the fact that any i.i.d. variables D_i with mean value zero satisfy the following bounds:

$$\forall p\in\mathbb{N},\ \exists C_p>0,\quad \left|\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^N D_i\right)^p\right]\right|\le\begin{cases}\frac{C_p}{N^{\frac{p}{2}}} & \text{if } p \text{ is even;}\\ \frac{C_p}{N^{\frac{p+1}{2}}} & \text{if } p \text{ is odd.}\end{cases} \tag{2.15}$$

This is proved by developing the power p of the sum, and then using the fact that the variables are i.i.d and have mean value zero. We hence infer from (2.13), (2.14), and (2.15) that

$$\langle A\rangle_N=A(y^*)+\frac{\sigma^2}{2N}A''(y^*)+\frac{1}{N^2}\left(\frac{m_3}{6}A^{(3)}(y^*)+\frac{\sigma^4}{24}A^{(4)}(y^*)\right)+O\left(\frac{1}{N^3}\right),$$

where σ is defined by (2.9) and

$$m_3=z^{-1}\int_{\mathbb{R}}\big(y-y^*\big)^3\exp\big(-\beta W(y)\big)\,dy. \tag{2.16}$$

More generally, it is possible to expand $\langle A\rangle_N$ at any order in $1/N$, provided that A is sufficiently smooth and $\exp(-\beta W)$ sufficiently small at infinity. In view of the bounds (2.15), we can see that using a Taylor expansion of order $2p$ around y^* for A gives an expansion of $\langle A\rangle_N$ of order p.

The practical consequence of Theorem 1 is that, for computational purposes, we may take the approximation

$$\langle A\rangle_N\approx A\left(z^{-1}\int_{\mathbb{R}}y\exp\big(-\beta W(y)\big)\,dy\right). \tag{2.17}$$

As pointed out above, it is possible to improve this approximation if necessary by expanding further in powers of $1/N$.

We conclude this section by showing that our consideration of a single "segment" carries through to the case when there are 3 repatoms, of respective index 0, M_1 and M_1+M_2, with $M_1h=L_1$, $M_2h=L_2$, $Nh=L=1$ (see Fig. 3). The average to compute is

$$\langle A\rangle_N=Z^{-1}\int_{\mathbb{R}^N}A\big(u^{M_1},u^{M_1+M_2}\big)\exp\left(-\beta\sum_{i=1}^N W\left(\frac{u^i-u^{i-1}}{h}\right)\right)du^1\cdots du^N.$$

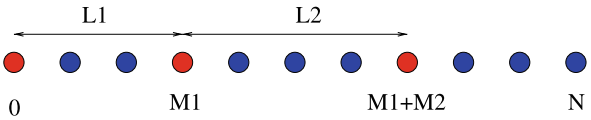

Fig. 3 When considering two consecutive segments—or more—, the argument may be readily adapted. See the text

In the regime $h \to 0$, $N, M_1, M_2 \to +\infty$ with M_1/N and M_2/N fixed, we have, using similar arguments,

$$\lim_{N\to+\infty} \langle A\rangle_N = A\big(L_1 y^*, (L_1+L_2)y^*\big).$$

The generalization to $N_r > 3$ repatoms, in the appropriate asymptotic regime, easily follows.

Remark 5 (The Small Temperature Limit) It is interesting here to consider the small temperature limit of the above expansion, that is, the limit $\beta \to +\infty$. In such a case, using the Laplace method (see Bender and Orszag 1978), it is possible to compute the limit of the terms that appear in the expansion (2.12) of $\langle A\rangle_N$. We give as an example the first and the second terms:

$$A\big(y^*\big) = A(a) + O\left(\frac{1}{\beta}\right), \qquad \frac{\sigma^2}{2}A''\big(y^*\big) = \frac{1}{2\beta}\frac{A''(a)}{W''(a)} + O\left(\frac{1}{\beta^2}\right),$$

where a is the point where W attains its minimum (in this remark, we assume for simplicity that W attains its minimum at a unique point). Hence,

$$\langle A\rangle_N = \left[A(a) + O\left(\frac{1}{\beta}\right)\right] + \frac{1}{N}\left[\frac{1}{2\beta}\frac{A''(a)}{W''(a)} + O\left(\frac{1}{\beta^2}\right)\right] + O\left(\frac{1}{N^2}\right).$$

Now, it is possible to recover these terms by expanding the energy E around the equilibrium configuration corresponding to $y_i = a$. Indeed, if we assume that $W(y) = W''(a)(y-a)^2/2$ in (2.2), then a simple explicit computation gives

$$\langle A\rangle_N = A(a) + \frac{1}{2N\beta W''(a)}A''(a) + O\left(\frac{1}{\beta^2 N^2}\right).$$

Hence, expanding the first terms of (2.12) in powers of $1/\beta$ for large β gives an expansion that agrees with that obtained using a harmonic approximation of the energy. This provides a quantitative evaluation of the latter approach in this asymptotic regime.

2.2 Limit of the Free Energy

We now look for a more demanding result. For clarity, let us return to the general coarse-grained average (1.7), which of course equals (2.2) and (2.5) in our simple NN case. Instead of searching for the limit of the average $\langle A\rangle$ for large N_c, we now look for the limit of the free energy per particle (see (1.6) and (1.11)). We discuss below (see Remark 8 and the end of this section) the interpretation of that quantity.

In the present section, u_r is in fact equal to u^N (the right-end atom) since atom 0, although a repatom, is fixed to avoid translation invariance: $u^0 = 0$. Thus, we wish to identify the behavior for N large of

$$E_{\mathrm{CG}}\big(u^N\big) = -\frac{1}{\beta}\ln\left[\int_{\mathbb{R}^{N-1}} \exp\big(-\beta E\big(u^1,\ldots,u^N\big)\big)\,du^1\cdots u^{N-1}\right]. \tag{2.18}$$

Note that E_{CG} is the free energy corresponding to integrating out $N-1$ variables. From Thermodynamics, it is expected that E_{CG} scales linearly with N. This is confirmed by the consideration of the harmonic potential $W(x) = \frac{k}{2}(x-a)^2$, for which $E_{\mathrm{CG}}(u^N) = \frac{kN}{2}(u^N - a)^2 + C(N,\beta,k)$, where

$$C(N,\beta,k) = \frac{1}{\beta}\left(N-\frac{1}{2}\right)\ln N - \frac{N-1}{2\beta}\ln\left(\frac{2\pi}{\beta k}\right)$$

does not depend on u_N (see the details in Patz 2009). Therefore, we introduce the free energy per particle

$$F_N(x) := \frac{1}{N}E_{\mathrm{CG}}(x),$$

so that

$$\langle A\rangle_N = Z_r^{-1}\int_{\mathbb{R}} A\big(u^N\big)\exp\big(-\beta N F_N\big(u^N\big)\big)\,du^N. \tag{2.19}$$

The limit behavior of F_N is provided by the large deviations principle. This claim is made precise in the following theorem.

Theorem 2 *Assume that the potential W satisfies*

$$\forall \xi\in\mathbb{R},\quad \int_{\mathbb{R}} \exp\big(\xi y - \beta W(y)\big)\,dy < +\infty, \tag{2.20}$$

and $\exp(-\beta W)\in H^1(\mathbb{R}\setminus\{0\})$. *Then the limit behavior of F_N is given by the following Legendre transform*:

$$\lim_{N\to+\infty}\left(F_N(x) + \frac{1}{\beta}\ln\frac{z}{N}\right) = F_\infty(x) \tag{2.21}$$

with

$$F_\infty(x) := \frac{1}{\beta}\sup_{\xi\in\mathbb{R}}\left(\xi x - \ln\left[z^{-1}\int_{\mathbb{R}}\exp(\xi y - \beta W(y))\,dy\right]\right) \tag{2.22}$$

and $z = \int_{\mathbb{R}}\exp\big(-\beta W(y)\big)\,dy$.

Remark 6 The assumption $\exp(-\beta W)\in H^1(\mathbb{R}\setminus\{0\})$ allows for W to be piecewise continuous, with discontinuity at the origin. This in particular allows us to deal with the type of potentials mentioned in Remark 3.

Proof Let us first rewrite the free energy $F_N(x)$ as follows:

$$\begin{aligned} F_N(x) &= -\frac{1}{\beta N}\ln\Bigg[\int_{\mathbb{R}^{N-1}}\exp\Bigg(-\beta\sum_{i=1}^{N-1}W\Big(\frac{u^i-u^{i-1}}{h}\Big) \\ &\quad -\beta W\Big(\frac{x-u^{N-1}}{h}\Big)\Bigg)du^1\cdots du^{N-1}\Bigg] \\ &= -\frac{N-1}{\beta N}\ln h-\frac{1}{\beta N}\ln\Bigg[\int_{\mathbb{R}^{N-1}}\exp\Bigg(-\beta W\Big(Nx-\sum_{i=1}^{N-1}y_i\Big) \\ &\quad -\beta\sum_{i=1}^{N-1}W(y_i)\Bigg)dy_1\cdots dy_{N-1}\Bigg] \\ &= -\frac{1}{\beta}\ln h-\frac{1}{\beta}\ln z-\frac{1}{\beta N}\ln\mu_N(x), \end{aligned}$$

where μ_N is the law of the random variable $(1/N)\sum_{i=1}^N Y_i$ and Y_i is a sequence of i.i.d. random variables with law $\mu(y)=z^{-1}\exp(-\beta W(y))$. Observe that

$$\mu_N(x)=\frac{N}{z^N}\int_{\mathbb{R}^{N-1}}\exp\Bigg(-\beta W\Big(Nx-\sum_{i=1}^{N-1}y_i\Big)-\beta\sum_{i=1}^{N-1}W(y_i)\Bigg)dy_1\cdots dy_{N-1}.$$

We also have

$$\mu_N(x)=N\mu^{*N}(Nx),$$

where μ^{*N} denotes the $(N-1)$-fold convolution product of μ ($\mu^{*2}=\mu*\mu$).

The sequence of measures μ_N satisfies a large deviations property (see, for instance, Ellis 1985a, 1985b, 1995; Varadhan 1984). We are going to use it in order to compute the limit of $\frac{1}{N}\ln\mu_N$. We first prove a lower bound, which is a simple consequence of the results of (Varadhan 1984). The upper bound is more involved: we need to reproduce the corresponding proof of (Varadhan 1984), and use a refined version of the central limit theorem (see Lions and Toscani 1995).

We introduce the function

$$G_N(x)=-\frac{1}{\beta N}\ln\mu_N(x), \tag{2.23}$$

which satisfies, in view of the above computation,

$$F_N(x)=-\frac{1}{\beta}\ln\frac{z}{N}+G_N(x). \tag{2.24}$$

First step: lower bound. We write

$$\mu_{N+1}(x)=(N+1)\int_{\mathbb{R}}\mu\big(N(x-t)+x\big)\mu_N(t)\,dt. \tag{2.25}$$

Let us define

$$J_N^x(t) = -\frac{1}{N}\ln\mu\big(N(x-t)+x\big).$$

This function clearly satisfies the following convergence:

$$\liminf_{u\to t, N\to+\infty} J_N^x(u) = J_\infty^x(t) := \begin{cases} +\infty & \text{if } t \neq x, \\ 0 & \text{if } t = x. \end{cases}$$

Hence, we may apply Theorem 2.3 of (Varadhan 1984), which implies that

$$\begin{aligned}\liminf_{N\to+\infty}\left(-\frac{1}{N}\ln\int_{\mathbb{R}}\exp\big(-NJ_N^x(t)\big)\mu_N(t)\,dt\right) &\geq \inf_{t\in\mathbb{R}}\big(J_\infty^x(t)+\beta F_\infty(t)\big)\\ &= \beta F_\infty(x), \end{aligned} \tag{2.26}$$

where F_∞ is defined by (2.22). Since the left-hand side of (2.26) is equal to $\frac{\beta(N+1)}{N}G_{N+1}(x) + \frac{\ln(N+1)}{N}$, we infer from the above bound that

$$\liminf_{N\to+\infty} G_N(x) \geq F_\infty(x). \tag{2.27}$$

Second step: upper bound. We now aim at bounding G_N from above. We recall that the function we maximize in (2.22) is concave, so there exists a unique $\xi_x \in \mathbb{R}$ such that

$$F_\infty(x) = \frac{1}{\beta}\left(\xi_x x - \ln\left[z^{-1}\int_{\mathbb{R}}\exp\big(\xi_x y - \beta W(y)\big)\,dy\right]\right).$$

The Euler–Lagrange equation of the maximization problem implies

$$x = \frac{\int_{\mathbb{R}} y\exp(\xi_x y - \beta W(y))\,dy}{\int_{\mathbb{R}}\exp(\xi_x y - \beta W(y))\,dy}. \tag{2.28}$$

We introduce the notations

$$\widetilde{\mu}(t) = \frac{\exp(\xi_x t - \beta W(t))}{\int_{\mathbb{R}}\exp(\xi_x t - \beta W(t))\,dt} \quad \text{and} \quad M(\xi) = z^{-1}\int_{\mathbb{R}}\exp\big(\xi t - \beta W(t)\big)\,dt,$$

and compute

$$\begin{aligned}\mu_N(x) &= N\int_{\mathbb{R}^{N-1}}\mu\left(Nx - \sum_{i=1}^{N-1} y_i\right)\mu(y_1)\cdots\mu(y_{N-1})\,dy_1\cdots dy_{N-1}\\ &= NM(\xi_x)^{N-1}\int_{\mathbb{R}^{N-1}}\mu\left(Nx - \sum_{i=1}^{N-1} y_i\right)\exp\left(-\xi_x\sum_{i=1}^{N-1} y_i\right)\\ &\quad\times\widetilde{\mu}(y_1)\cdots\widetilde{\mu}(y_{N-1})\,dy_1\cdots dy_{N-1}\\ &\geq NM(\xi_x)^{N-1}\int_{|Nx-\sum y_i|\leq\delta}\mu\left(Nx - \sum_{i=1}^{N-1} y_i\right)\end{aligned}$$

$$\times \exp\left(-\xi_x \sum_{i=1}^{N-1} y_i\right) \widetilde{\mu}(y_1)\cdots\widetilde{\mu}(y_{N-1})\,dy_1\cdots dy_{N-1}$$

$$\geq NM(\xi_x)^{N-1}\Big(\inf_{[-\delta,\delta]}\mu\Big)\exp\big(-\xi_x Nx - |\xi_x|\delta\big)$$

$$\times \int_{|Nx-\sum y_i|\leq\delta} \widetilde{\mu}(y_1)\cdots\widetilde{\mu}(y_{N-1})\,dy_1\cdots dy_{N-1},$$

for any $\delta > 0$. Hence,

$$G_N(x) \leq -\frac{1}{\beta N}\ln N - \frac{N-1}{\beta N}\ln\big(M(\xi_x)\big) + \frac{\xi_x x}{\beta} + |\xi_x|\frac{\delta}{\beta N}$$
$$- \frac{1}{\beta N}\ln\Big(\inf_{[-\delta,\delta]}\mu\Big) - \frac{1}{\beta N}\ln\mathbb{P}\left(\left|\frac{1}{N}\sum_{i=1}^{N-1} Y_i - x\right| \leq \frac{\delta}{N}\right), \quad (2.29)$$

where the random variables Y_i are i.i.d. of law $\widetilde{\mu}$. Equation (2.28) implies that $\mathbb{E}(Y_i) = x$. According to the hypotheses on W, we have $\widetilde{\mu} \in H^1(\mathbb{R}\setminus\{0\})$, hence we may apply Theorem 5.1 of (Lions and Toscani 1995). It implies that the law θ_N of the variable $(\sum_{i=1}^N Y_i - Nx)/\sqrt{N}$ converges in $H^1(\mathbb{R})$ to some normal law. In particular, we have convergence in L^∞, hence

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{i=1}^{N-1} Y_i - x\right| \leq \frac{\delta}{N}\right) = \int_{\frac{x-\delta}{\sqrt{N-1}}}^{\frac{x+\delta}{\sqrt{N-1}}} \theta_{N-1}(t)\,dt \geq \frac{2\gamma\delta}{\sqrt{N-1}},$$

for N large enough, where $\gamma > 0$ does not depend on N. Inserting this inequality into (2.29), we find

$$G_N(x) \leq -\frac{1}{\beta N}\ln N - \frac{N-1}{\beta N}\ln\big(M(\xi_x)\big) + \frac{\xi_x x}{\beta} + |\xi_x|\frac{\delta}{\beta N}$$
$$- \frac{1}{\beta N}\ln\Big(\inf_{[-\delta,\delta]}\mu\Big) - \frac{1}{\beta N}\ln\left(\frac{2\gamma\delta}{\sqrt{N-1}}\right). \quad (2.30)$$

Hence,

$$\limsup_{N\to+\infty} G_N(x) \leq -\frac{1}{\beta}\ln\big(M(\xi_x)\big) + \frac{\xi_x x}{\beta},$$

which implies, according to the definition of M and ξ_x, that

$$\limsup_{N\to+\infty} G_N(x) \leq F_\infty(x). \quad (2.31)$$

Estimates (2.27) and (2.31) imply $\lim_{N\to+\infty} G_N(x) = F_\infty(x)$. In view of (2.24), this implies (2.21). □

Remark 7 (The Small Temperature Limit) As in Remark 5, it is possible to compute the expansion of $F_\infty(x)$ as $\beta \to +\infty$. Using the Laplace method, and assuming that W is convex, one finds that

$$F_\infty(x) = W(x) + \frac{1}{2\beta} \ln W''(x) + O\left(\frac{1}{\beta^2}\right).$$

Let us now consider another strategy to find an approximation of F_N. In the spirit of the quasicontinuum method, we expand $E(u^1, \dots, u^N)$ around the equilibrium configuration $\overline{u}^i = iu^N/N$, for a given u^N. More precisely, we set $u^i = \overline{u}^i + \xi_i$, assume that ξ_i is small, and expand the energy at second order with respect to ξ_i, as explained in the Introduction. This yields the approximate energy $\widetilde{E}$ defined by (1.9), that we next insert in (2.18) (as we did in (1.10)). Due to the harmonic approximation, the resulting coarse-grained energy that we denote E_{HA}, is analytically computable and reads

$$E_{\mathrm{HA}}(x) = NW(x) + \frac{N-1}{2\beta} \ln W''(x) + \frac{N-1}{2\beta} \ln \frac{\beta}{2\pi} + \frac{1}{2\beta} \ln N. \tag{2.32}$$

Hence,

$$F_{\mathrm{HA}}(x) := \lim_{N\to+\infty} \frac{1}{N} E_{\mathrm{HA}}(x) = W(x) + \frac{1}{2\beta} \ln W''(x) + \frac{1}{2\beta} \ln \frac{\beta}{2\pi}. \tag{2.33}$$

Thus, up to an additive constant, $F_{\mathrm{HA}}(x)$ corresponds to the first-order approximation (in powers of $1/\beta$) of $F_\infty(x)$.

Slightly improving the proof of Theorem 2 above, it is possible to prove the convergence of the derivative of the free energy, a quantity which is indeed practically relevant (e.g., for the simulation of (1.8)):

Corollary 1 *Assume that the hypotheses of Theorem 2 are satisfied. Then we have*

$$F_N(x) + \frac{1}{\beta} \ln\left(\frac{z}{N}\right) \longrightarrow F_\infty(x) \quad \text{in } L^p_{\mathrm{loc}}, \ \forall p \in [1, +\infty).$$

In particular, this implies that F_N' converges to F_∞' in $W^{-1,p}_{\mathrm{loc}}$.

Proof According to Theorem 2, we already know the pointwise convergence of $G_N(x) = F_N(x) + \beta^{-1} \ln(z/N)$. We therefore only need to prove that G_N is bounded in L^∞_{loc} to prove our claim.

Lower bound: We go back to (2.25), and point out that $\mu \le 1/z$. Hence,

$$\mu_{N+1}(x) \le \frac{N+1}{z} \int_{\mathbb{R}} \mu_N = \frac{N+1}{z},$$

which implies, using (2.23) that

$$G_{N+1}(x) \ge -\frac{1}{\beta(N+1)} \ln \frac{N+1}{z},$$

which is bounded from below independently of N.

Upper bound: We return to (2.30), and notice that according to the definition of ξ_x, the function $x \mapsto \xi_x$ is continuous. In addition, the constant γ in (2.30) is a continuous function of ξ_x. Therefore, (2.30) provides an upper bound on G_N.

As a conclusion, G_N is bounded in L^∞_{loc}, which allows to conclude. □

Remark 8 Considering the above theoretical results, it could be tempting to approach the average (2.19), that is,

$$\langle A \rangle_N = Z_r^{-1} \int_{\mathbb{R}} A\big(u^N\big) \exp\big(-\beta N F_N\big(u^N\big)\big)\, du^N,$$

by

$$Z_\infty^{-1} \int_{\mathbb{R}} A\big(u^N\big) \exp\big(-\beta N F_\infty\big(u^N\big)\big)\, du^N, \tag{2.34}$$

with $Z_\infty = \int_{\mathbb{R}} \exp(-\beta N F_\infty(u^N))\, du^N$. Note that F_N has been replaced by F_∞ in the exponential factor. This strategy is not efficient since this approximation does not provide the expansion (2.11)–(2.12) of $\langle A \rangle_N$ in powers of $1/N$. Indeed, it is possible to use the Laplace method to compute the expansion of (2.34) as $N \to +\infty$. It reads

$$A\big(y^*\big) + \frac{1}{2N}\left(\sigma^2 A''\big(y^*\big) + \frac{m_3}{\sigma^2} A'\big(y^*\big)\right) + o\left(\frac{1}{N}\right),$$

where σ is defined by (2.9) and m_3 is defined by (2.16). This expansion coincides with (2.11)–(2.12) only for the first term, that is $A(y^*)$. The second one differs, unless $m_3 A'(y^*) = 0$.

To improve the approximation (2.34), one may use the precised large deviations principle (see Dembo and Zeitouni 1993, Theorem 3.7.4 or Bahadur and Ranga Rao 1960). In such a case, one replaces (2.34) by

$$\widetilde{Z}_\infty^{-1} \int_{\mathbb{R}} A\big(u^N\big) \sqrt{F''_\infty\big(u^N\big)} \exp\big(-\beta N F_\infty\big(u^N\big)\big)\, du^N, \tag{2.35}$$

with $\widetilde{Z}_\infty = \int_{\mathbb{R}} \sqrt{F''_\infty(u^N)} \exp(-\beta N F_\infty(u^N))\, du^N$. This quantity is well defined since F_∞ is a convex function. Then it is seen that the expansion of (2.35) in powers of $1/N$ agrees with (2.11)–(2.12) up to the second order term. Note however that using (2.35) leads to a much more expensive computation than using (2.12), since it requires the evaluation of F_∞ and its second derivative.

The above convergence of the free energy F_N is useful for the computation of the free energy of a chain of atoms with a *prescribed* length. Indeed, consider a chain on which we impose

$$u^N = \ell,$$

for a fixed ℓ. We aim at computing the free energy F_N as a function of ℓ, in the limit $N \to +\infty$. We have

$$F_N(\ell) = -\frac{1}{\beta N} \ln\left[\int_{\mathbb{R}^{N-1}} \exp\left(-\beta \sum_{i=1}^{N} W\left(\frac{u^i - u^{i-1}}{h}\right)\right) du^1 \cdots du^{N-1}\right],$$

where $u^N = \ell$. The limit of F_N is provided by Theorem 2.

Another interest of the approach is to provide an approximation of $F_N'(\ell)$, a quantity related to the constitutive law of the material under consideration, at the *finite temperature* $1/\beta$. Indeed, note that

$$F_N'(\ell) = \langle A_N \rangle_N^\ell, \tag{2.36}$$

where $\langle\cdot\rangle_N^\ell$ is the average with respect to the Gibbs measure $d\mu_\ell(u^1, \ldots, u^{N-1})$ associated to the energy $\sum_{i=1}^{N-1} W(\frac{u^i - u^{i-1}}{h}) + W(\frac{\ell - u^{N-1}}{h})$ (recall that $h = 1/N$), and the observable A_N is defined by

$$A_N\big(u^1, \ldots, u^{N-1}\big) = W'\big(N\big(\ell - u^{N-1}\big)\big).$$

Note that $W'(N(\ell - u^{N-1}))$ can be interpreted as the force in the spring between atom $N-1$ and N, when the position of the latter is fixed at the value ℓ. Hence, $F_N'(\ell)$ can be interpreted as the average force between atoms $N-1$ and N, when the position of atom N is prescribed at $u^N = \ell$, and the relation $\ell \mapsto F_N'(\ell)$ can be considered as the constitutive relation (at a given temperature) of the chain, providing the stress $F_N'(\ell)$ as a function of the strain ℓ. Corollary 1 provides the convergence of $F_N'(\ell)$ to $F_\infty'(\ell)$.

Remark 9 Note that in (2.36) the observable A_N depends on N. Hence, the results of Sect. 2.1 (obtained using the law of large numbers and not involving the large deviations principle) do not apply to compute the large N limit of $\langle A_N \rangle_N^\ell$. In addition, the Gibbs measure $d\mu_\ell$ is not of the form considered previously, since the atom N has a prescribed position.

2.3 Numerical Tests

For our numerical tests, we choose the pair interaction potential

$$W(x) = \frac{1}{2}(x-1)^4 + \frac{1}{2}x^2 \tag{2.37}$$

shown on Fig. 4. Note that $W(x)$ grows fast enough to $+\infty$ when $|x| \to +\infty$, such that assumptions (2.8) and (2.20) are satisfied. Note also that we have made no assumption on the convexity of W in Theorems 1 and 2. We consider here a convex potential. At the end of this section, we will consider a nonconvex example (see (2.38)), and show that we obtain similar conclusions.

We first consider the computation of ensemble averages, and we again restrict ourselves to the case of two repatoms $u^0 = 0$ and u^N. This is just for simplicity and

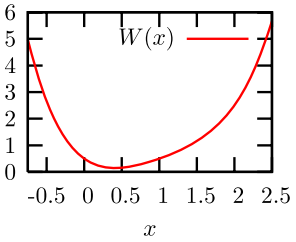

Fig. 4 The potential W (defined by (2.37)) chosen for the tests

for the sake of demonstrating the feasibility and the interest of our approach. The case of N_r repatoms may be treated likewise.

We choose an observable $A(x)$, and we compare the following four quantities:

(i) the exact average $\langle A\rangle_N$ defined by (2.2). Following (1.3)–(1.4), this quantity is computed as the long-time average of $A(u^N(t))$ along the full system dynamics

$$du_t = -\nabla_u E(u_t)\,dt + \sqrt{2/\beta}\,dB_t \quad \text{in } \mathbb{R}^N.$$

This equation is numerically integrated with the forward Euler scheme (also called the Euler–Maruyama scheme), with a small time step Δt:

$$u_{n+1} = u_n - \Delta t \nabla_u E(u_n) + \sqrt{2\Delta t/\beta} G_n$$

where G_n is a N-dimensional vector of random variables distributed according to a Gaussian normal law. In practice, we have simulated many independent realizations of this SDE, in order to compute error bars for $\langle A\rangle_N$.

(ii) a harmonic type approximation of $\langle A\rangle_N$, based on the "interpolation + harmonic expansion" procedure outlined above. That is, we introduce $E_{\rm HA}$ defined by (2.32), and we approximate $\langle A\rangle_N$ by

$$\langle A\rangle_N^{\rm HA} := \frac{\int_{\mathbb{R}} A(x)\exp[-\beta E_{\rm HA}(x)]\,dx}{\int_{\mathbb{R}} \exp[-\beta E_{\rm HA}(x)]\,dx}.$$

(iii) a Law of Large Numbers (LLN) type approximation of $\langle A\rangle_N$, which consists in approximating $\langle A\rangle_N$ by $A(y^*)$, following Theorem 1.

(iv) a refined approximation, which consists in approximating $\langle A\rangle_N$ by $\langle A\rangle_N^{\rm approx}$ defined by (2.12), following Theorem 1.

Note that only one-dimensional integrals are needed for approximations (ii), (iii), and (iv). They can be computed with a high accuracy.

We plot on Fig. 5 these four quantities, for increasing values of N (the temperature is fixed at $1/\beta = 1$), for the observable $A(x) = \exp(x)$. On Fig. 6, we compare the same quantities, now as functions of the temperature, for $N = 100$ and for $N = 10$. We here work with $A(x) = x^2$, for which $\langle A\rangle_N = \langle A\rangle_N^{\rm approx}$.

As expected, the thermodynamic limit strategies (iii) and (iv) better agree with the full atom calculation, whatever the temperature, provided the number of eliminated atoms is large (note that the strategy (iv) is very accurate even for the small value $N = 10$, at the temperature $1/\beta = 1$). Approximation (ii) is clearly ineffective for high

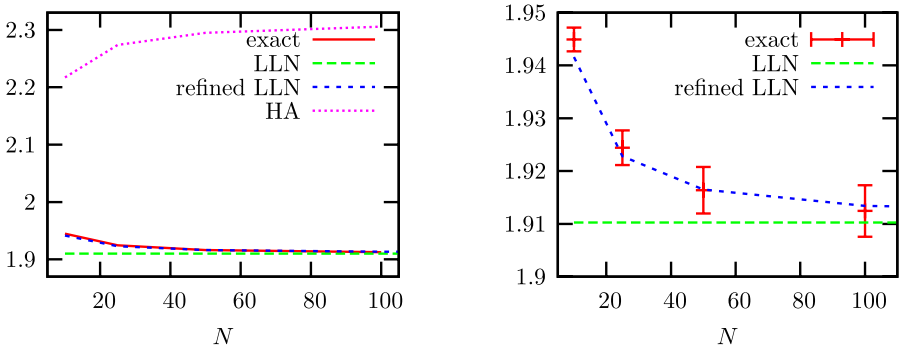

Fig. 5 Convergence, as N increases, of $\langle A(u^N)\rangle_N$ (exact), of $\langle A(u^N)\rangle_N^{\text{approx}}$ (refined LLN) and of $\langle A(u^N)\rangle_N^{\text{HA}}$ (HA) and comparison to $A(y^*)$ (LLN) (temperature $1/\beta = 1$, observable $A(x) = \exp(x)$; we have performed computations for $N = 10$, 25, 50, and 100; on the *right graph*, we only plot the most accurate results with *error bars*)

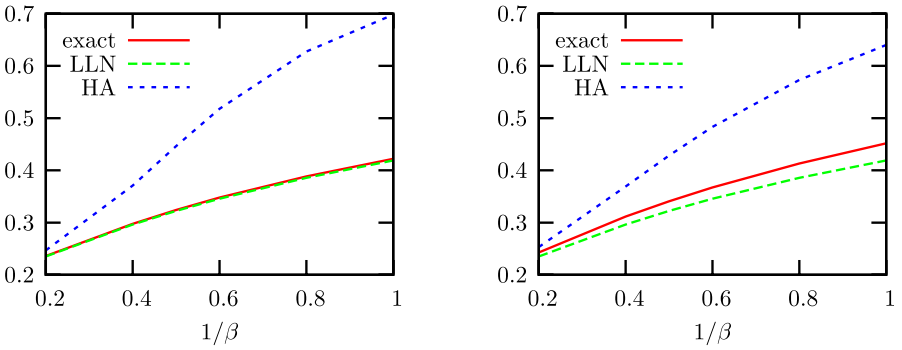

Fig. 6 We plot $\langle A(u^N)\rangle_N = \langle A(u^N)\rangle_N^{\text{approx}}$ (exact), $\langle A(u^N)\rangle_N^{\text{HA}}$ (HA) and $A(y^*)$ (LLN) as functions of the temperature $1/\beta$: on the *left graph*, $N = 100$; on the *right graph*, $N = 10$ (observable $A(x) = x^2$)

temperatures. On the other hand, for a sufficiently small temperature and a sufficiently small number of eliminated atoms, this approximation is close to the full atom result. However, even for the small values $N = 10$ and $1/\beta = 0.2$, our asymptotic result $\langle A(u^N)\rangle_N^{\text{approx}} = 1.6299$ (for $A(x) = \exp(x)$) is closer to the exact result $\langle A(u^N)\rangle_N = 1.6303 \pm 0.0008$ than the harmonic approximation result $\langle A(u^N)\rangle_N^{\text{HA}} = 1.6469$.

Remark 10 As in Remark 3, we emphasize that the computations reported on here do not account for constraints on the positions of atoms. Analogous computations, that account for constraints, may be performed, using the potential W_c defined by (2.6)–(2.37). They provide similar conclusions, as can be seen on Fig. 7, which is very similar to Fig. 6.

We now turn to the case of the free energy F_N. We are going to compare different approximations of its derivative. The full atom value $F_N'(x)$ is computed as the ensemble average (2.36). We compare this quantity with

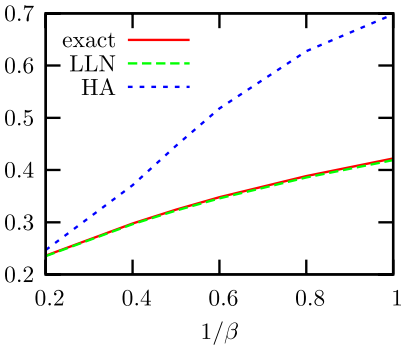

Fig. 7 We plot $\langle A(u^N)\rangle_N = \langle A(u^N)\rangle_N^{\text{approx}}$ (exact), $\langle A(u^N)\rangle_N^{\text{HA}}$ (HA) and $A(y^*)$ (LLN) as functions of the temperature $1/\beta$ ($N = 100$, observable $A(x) = x^2$). The potential energy is of type (2.6): it is equal to $W(x)$ defined by (2.37) if $x > 0$, and $+\infty$ otherwise

(i) its large N limit $F'_\infty(x)$, where F_∞ is defined by (2.22), on the one hand,
(ii) and, on the other hand, its harmonic type approximation $F'_{\text{HA}}(x)$, where F_{HA} is defined by (2.33). We have

$$F'_{\text{HA}}(x) = W'(x) + \frac{1}{2\beta}\frac{W'''(x)}{W''(x)}.$$

We briefly detail how we compute $F'_\infty(x)$. Let ξ_x be the unique real number at which the supremum in (2.22) is attained. We have

$$F'_\infty(x) = \frac{\xi_x}{\beta}.$$

The Euler–Lagrange equation solved by ξ_x is (2.28), that we recast as

$$z^{-1}\int_{\mathbb{R}} (x-y)\exp(\xi_x y)\exp\bigl(-\beta W(y)\bigr)\,dy = 0.$$

Let us introduce $G(y,\xi) = (x-y)\exp(\xi y)$. We hence look for ξ_x such that $\mathbb{E}_\mu[G(Y,\xi_x)] = 0$, where the scalar random variable Y is distributed according to the probability measure $\mu(y) = z^{-1}\exp(-\beta W(y))$. The Robbins–Monroe algorithm (Kushner and Clark 1978) can be used to compute ξ_x, hence $F'_\infty(x)$.

We first study the convergence of $F'_N(x)$ to $F'_\infty(x)$ as N increases, for a fixed chain length $x = 1.4$ and a fixed temperature $1/\beta = 1$. Results are shown on Fig. 8. We indeed observe that $F'_N(x) \to F'_\infty(x)$ when $N \to +\infty$.

We now compare the two approximations (i) and (ii) of $F'_N(x)$, for $N = 100$ and $1/\beta = 1$. Results are shown on Fig. 9. We observe that $F'_\infty(x)$ is a very good approximation of $F'_N(x)$. As expected, the temperature is too high for the harmonic approximation to provide an accurate approximation of $F'_N(x)$.

On Fig. 10, we plot $F'_\infty(x)$ for several temperatures, as well as its zero temperature limit, which is $W'(x)$ (see Remark 7).

Up to here, we have used the convex potential (2.37). For the sake of completeness, we now briefly consider the case of a nonconvex potential W. We choose the toy-model

$$W(x) = \bigl(x^2-1\bigr)^2, \tag{2.38}$$

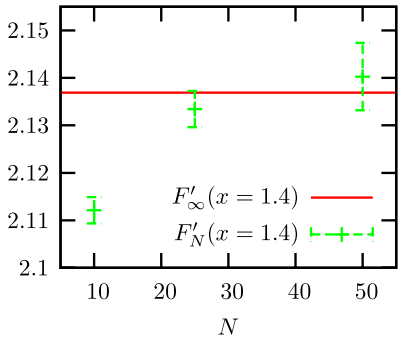

Fig. 8 Convergence of $F'_N(x)$ (shown with *error bars*) to $F'_\infty(x)$ as N increases (temperature $1/\beta = 1$, fixed chain length $x = 1.4$)

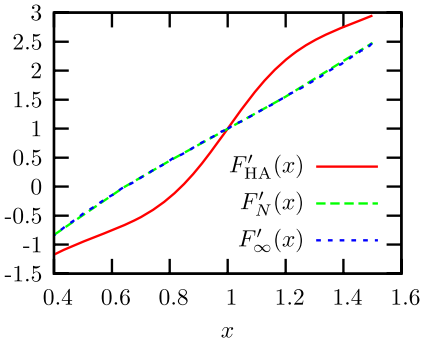

Fig. 9 We plot $F'_N(x)$, $F'_\infty(x)$, and $F'_{\mathrm{HA}}(x)$, for the temperature $1/\beta = 1$ and $N = 100$. On the scale of the figure, $F'_N(x)$ and $F'_\infty(x)$ are on top of each other

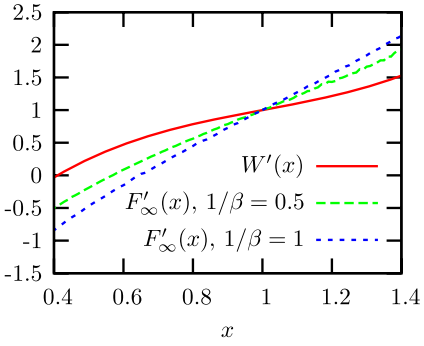

Fig. 10 $F'_\infty(x)$ for different temperatures

which corresponds to a double-well potential. We first study the convergence of $F'_N(x)$ to $F'_\infty(x)$ as N increases, for a fixed chain length $x = 0.5$ and a fixed temperature $1/\beta = 1$. Results are shown on Fig. 11. As for the convex potential case, we observe that $F'_N(x) \to F'_\infty(x)$ when $N \to +\infty$. On Fig. 12, we plot $F'_\infty(x)$ for several temperatures. Numerical results are consistent with the small temperature limit $\lim_{T\to 0} F'_\infty(x) = (W^*)'(x)$, where W^* is the convex envelop of W.

Note that, in view of its definition (2.22), F_∞ is always a convex function. Hence, as in the zero temperature case, we observe, in this one-dimensional setting, that the macroscopic constitutive law $\ell \mapsto F_\infty(\ell)$ is a convex function.

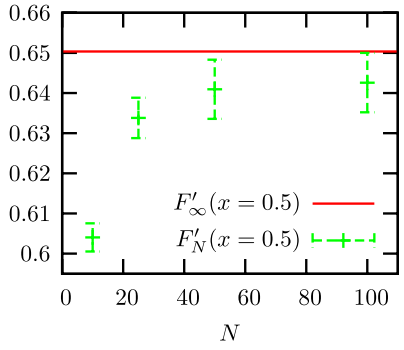

Fig. 11 Convergence of $F'_N(x)$ (shown with *error bars*) to $F'_\infty(x)$ as N increases (temperature $1/\beta = 1$, fixed chain length $x = 0.5$), in the case of the double-well potential (2.38)

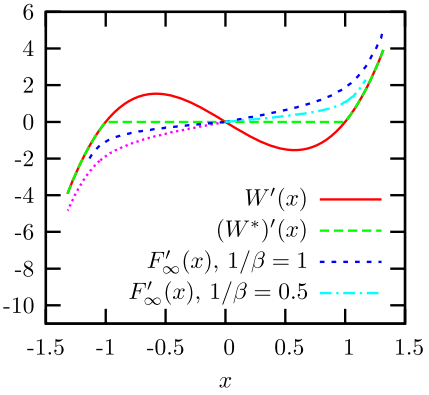

Fig. 12 $F'_\infty(x)$ for different temperatures, in the case of the double-well potential (2.38)

3 The NNN Case and Some Extensions

In this section, we first consider the case of a NNN interacting system. The analysis is detailed in Sect. 3.1. In Sect. 3.2, we point out some possible extensions, first the NNNN case (still for one-dimensional systems) and second the case of linear polymer chains, where atoms sample the physical space $\mathbb{R}^3$.

3.1 The Next-to-Nearest Neighbor (NNN) Case

We now consider the next-to-nearest neighbor case. It turns out that, for the computation of ensemble averages as well as for other questions, this case is significantly more intricate than the NN case. Our strategy, based on the law of large numbers, will be similar to that used for the NN case, but the object manipulated are not independent random variables any longer. Markov chains are the right notion formalizing the situation mathematically.

We begin by introducing the rescaled atomistic energy, similarly to (2.1):

$$E\big(u^1,\dots,u^N\big) = \sum_{i=1}^{N} W_1\left(\frac{u^i - u^{i-1}}{h}\right) + \sum_{i=1}^{N-1} W_2\left(\frac{u^{i+1} - u^{i-1}}{h}\right).$$

As above, we set $u^0 = 0$, and we introduce the change of variables (2.3), replacing $(u^i - u^{i-1})/h$ by the interatomic distances y_i. Recall from (2.4) that $u^N = \frac{1}{N}\sum_{i=1}^N y_i$. The ensemble average $\langle A\rangle_N$ of an observable that depends only on the right-end atom therefore writes

$$\begin{aligned}\langle A\rangle_N &= Z^{-1}\int_{\mathbb{R}^N} A\big(u^N\big)\exp\big(-\beta E\big(u^1,\dots,u^N\big)\big)\,du^1\cdots du^N\\ &= Z^{-1}\int_{\mathbb{R}^N} A\left(\frac{1}{N}\sum_{i=1}^N y_i\right)\\ &\qquad\times e^{-\beta(\sum_{i=1}^N W_1(y_i)+\sum_{i=1}^{N-1} W_2(y_i+y_{i+1}))}\,dy_1\cdots dy_N. \end{aligned}\tag{3.1}$$

The key ingredient is now to see the above expression, as N goes to infinity, as an asymptotics for a *discrete-time Markov chain*. The asymptotics of Markov chains being a mathematical problem much more involved than that of i.i.d. sequences, we restrict ourselves to the computation of the average of an observable. The asymptotic behavior of the free energy may be studied, applying a large deviations principle for Markov chains (see, for instance den Hollander 2000, Theorem IV.3). We will not pursue in this direction.

Section 3.1.1 deals with the case of two repatoms (namely $u^0 = 0$ and u^N), while Sect. 3.1.2 indicates the changes in order to deal with more than two repatoms. Numerical results will be reported in Sect. 3.1.3.

3.1.1 Limit of the Average, the Case of Two Repatoms

In order to compute $\lim_{N\to+\infty}\langle A\rangle_N$, we introduce the notation

$$f(x,y) := \exp\big(-\beta W_2(x+y)\big)\exp\big(-\beta W_1(y)\big).$$

Equation (3.1) rewrites

$$\begin{aligned}\langle A\rangle_N = Z^{-1}\int_{\mathbb{R}^N} A\left(\frac{1}{N}\sum_{i=1}^N y_i\right)e^{-\beta W_1(y_1)}\\ \times f(y_1,y_2)\cdots f(y_{N-1},y_N)\,dy_1\cdots dy_N.\end{aligned}\tag{3.2}$$

Our method consists in considering the sequence of variables $(y_1,\dots,y_N)$ in (3.2) as a realization of a Markov chain with kernel $f(\cdot,\cdot)$. However, the slight technical difficulty at this stage is that the kernel f is not normalized, since in general

$$\int_{\mathbb{R}} f(y_1,y_2)\,dy_2 = \int_{\mathbb{R}}\exp\big(-\beta W_2(y_1+y_2)\big)\exp\big(-\beta W_1(y_2)\big)\,dy_2 \neq 1.$$

A standard trick of probability theory allows to circumvent this difficulty. Introduce

$$\overline{f}(x,y) := \exp\left[-\beta W_2(x+y) - \frac{\beta}{2}W_1(x) - \frac{\beta}{2}W_1(y)\right].$$

Note that $\overline{f}$ is a symmetric function (whereas f is not), hence the operator

$$P\phi(y) = \int_{\mathbb{R}} \overline{f}(y,z)\phi(z)\,dz \tag{3.3}$$

is self-adjoint on $L^2(\mathbb{R})$. Consider then the variational problem

$$\max_{\psi\in L^2(\mathbb{R})} \left\{ \int_{\mathbb{R}^2} \psi(y)\psi(z)\overline{f}(y,z)\,dy\,dz;\ \int_{\mathbb{R}} \psi^2(y)\,dy = 1 \right\}. \tag{3.4}$$

Using standard tools of spectral theory of self-adjoint compact operators (Dunford and Schwartz 1963), it is possible to prove that this problem has a maximizer (we denote it by ψ_1), and that up to changing ψ_1 into $-\psi_1$, the maximizer is unique. In addition, one can choose ψ_1 such that $\psi_1 > 0$. The Euler–Lagrange equation of (3.4) reads

$$\lambda\psi_1(y) = \int_{\mathbb{R}} \overline{f}(y,z)\psi_1(z)\,dz$$

for some λ, and we recognize an eigenvalue problem, $P\psi_1 = \lambda\psi_1$, for the operator (3.3). Multiplying the above equation by $\psi_1(y)$ and integrating, we obtain

$$\lambda = \int_{\mathbb{R}^2} \psi_1(y)\psi_1(z)\overline{f}(y,z)\,dy\,dz > 0.$$

We now define

$$g(y,z) := \frac{\psi_1(z)}{\lambda\psi_1(y)}\overline{f}(y,z). \tag{3.5}$$

By construction,

$$\int_{\mathbb{R}} g(y,z)\,dz = 1, \qquad \int_{\mathbb{R}} \psi_1^2(y)g(y,z)\,dy = \psi_1^2(z).$$

The average (3.2) now reads

$$\langle A\rangle_N = Z_g^{-1} \int_{\mathbb{R}^N} A\left(\frac{1}{N}\sum_{i=1}^N y_i\right)\psi_1(y_1)e^{-\frac{\beta}{2}W_1(y_1)} \times g(y_1,y_2)\cdots g(y_{N-1},y_N)\frac{e^{-\frac{\beta}{2}W_1(y_N)}}{\psi_1(y_N)}\,dy_1\cdots dy_N,$$

with

$$Z_g = \int_{\mathbb{R}^N} \psi_1(y_1)e^{-\frac{\beta}{2}W_1(y_1)}g(y_1,y_2)\cdots g(y_{N-1},y_N)\frac{e^{-\frac{\beta}{2}W_1(y_N)}}{\psi_1(y_N)}\,dy_1\cdots dy_N,$$

and where $(y_1,\ldots,y_N)$ may now be seen as a realization of a *normalized* Markov chain of kernel g, with invariant probability measure ψ_1^2. We assume that $\overline{f}$ decays fast enough at infinity (which ensures for instance that (3.4) is well posed) and that

it is positive. This latter assumption ensures that the Markov chain satisfies the following accessibility condition: for any $x \in \mathbb{R}$, and for any measurable set $\mathcal{B} \subset \mathbb{R}$ of positive Lebesgue measure, we have $\int_{\mathcal{B}} g(x,y)\,dy > 0$. Under this property, combined with the existence of an invariant probability measure ($\psi_1^2(y)\,dy$ in the present case), it is known (see Meyn and Tweedie 1993, Theorem 17.1.7) that the invariant measure is unique, and that the Markov chain satisfies a law of large numbers with respect to it. We now state a direct corollary of this general result that applies to our context.

Theorem 3 *Assume that A is continuous, and satisfies the following conditions*:

$$\exists p \geq 0,\ \exists C > 0,\ \forall x \in \mathbb{R}, \quad \big|A(x)\big| \leq C\big(1 + |x|^p\big).$$

Under the assumptions that $W_1, W_2 \in L^1_{\mathrm{loc}}(\mathbb{R})$ are bounded from below that $e^{-\beta W_1}, e^{-\beta W_2} \in W^{1,1}_{\mathrm{loc}}(\mathbb{R})$ and that

$$\forall q \geq 0, \quad \int_{\mathbb{R}} |x|^q e^{-\beta W_1(x)}\,dx < +\infty \quad \text{and} \quad \int_{\mathbb{R}} |x|^q e^{-\beta W_2(x)}\,dx < +\infty, \tag{3.6}$$

the ergodic theorem for Markov chains (Meyn and Tweedie 1993) *yields*

$$\lim_{N\to+\infty} \langle A\rangle_N = A\big(y^*\big) \quad \text{where } y^* := \int_{\mathbb{R}} y\psi_1^2(y)\,dy.$$

Remark 11 Note that, for the result to hold true, (3.6) is not needed. The existence of the moment of order p is sufficient. However, assumption (3.6) will be useful for Theorem 4 below.

Remark 12 It might sound a little strange that ψ_1 is the eigenvector of the transition operator P defined by (3.3), whereas the invariant measure of the chain is ψ_1^2. This is explained by the following fact: the expectation value of y_i is equal to

$$\mathbb{E}(y_i) = \int_{\mathbb{R}} x\big(P^{i-1}\varphi_0\big)(x)\big(P^{N-i}\varphi_1\big)(x)\,dx,$$

for some initial laws φ_0 and φ_1, where P^i is recursively defined by $P^i\varphi = P[P^{i-1}\varphi]$ (recall that P is the operator defined by (3.3)). Hence, if $1 \ll i \ll N$, then both $P^{i-1}\varphi_0$ and $P^{N-i}\varphi_1$ converge to the eigenvector of P associated with the largest eigenvalue, that is, ψ_1 (hence the appearance of ψ_1^2 rather than ψ_1). This is explained in more details in Sect. 3.2.1 below in the case of a non self-adjoint transition operator P.

Here again, it is possible to compute the next terms in the expansion of $\langle A\rangle_N$ in powers of $1/N$. However, the computations are much more intricate than in the i.i.d. case. The terms of the expansion here contain covariance terms, together with terms containing the initial state of the Markov chain. As an example, we give the first term of the expansion in the following theorem.

Theorem 4 *Assume that A is of class C^3. Then under the assumptions of Theorem 3, we have*

$$\langle A\rangle_N = A\big(y^*\big) + \frac{1}{N}A'\big(y^*\big)\sum_{i\geq 1}\mathbb{E}\big(Y_i - y^*\big) + \frac{\sigma^2}{2N}A''\big(y^*\big) + O\left(\frac{1}{N^2}\right), \tag{3.7}$$

with

$$\sigma^2 = \int_{\mathbb{R}}\big(x-y^*\big)^2\psi_1^2(x)\,dx + 2\sum_{i\geq 2}\mathbb{E}\big(\big(\widetilde{Y}_i - y^*\big)\big(\widetilde{Y}_1 - y^*\big)\big), \tag{3.8}$$

where $(Y_i)_{i\geq 1}$ and $(\widetilde{Y}_i)_{i\geq 1}$ are Markov chains of initial law $Z_1^{-1}\psi_1 e^{-\frac{\beta}{2}W_1}$ and ψ_1^2 respectively, and of transition kernel g. Moreover, the series appearing in (3.7) and (3.8) converge exponentially fast.

Remark 13 Let us mention that σ^2 defined by (3.8) is exactly the variance appearing in the central limit theorem for Markov chains (Meyn and Tweedie 1993, Theorem 17.0.1).

In addition, we see that in the special case of i.i.d. random variables, the second term of (3.8) vanishes, together with the term proportional to $A'(y^*)$ in (3.7). We then recover estimate (2.11)–(2.12).

Remark 14 The assumptions of Theorems 3 and 4 are not sharp. However, they allow for simple proofs, and for a wide variety of interaction potentials W_1 and W_2.

Remark 15 Again, as in Remark 3, constraints on the positions of the atoms may be accounted for.

Note that Theorem 4 suggests a strategy for numerically computing the terms of (3.7). Indeed, it is possible to compute numerically ψ_1 by discretizing (3.4). Numerical integration then allows to compute y^* and the variance $\int_{\mathbb{R}}(x-y^*)^2\psi_1^2(x)\,dx$. The computation of the infinite sums in (3.7)–(3.8) is then performed using a simulation of the corresponding Markov chains and taking the expectation value. Note that the law of Y_i converges exponentially fast (when $i\to\infty$) to the invariant measure ψ_1^2 due to the existence of a spectral gap for the transition operator. Hence, the terms $\mathbb{E}(Y_i - y^*)$ and $\mathbb{E}((\widetilde{Y}_i - y^*)(\widetilde{Y}_1 - y^*))$ that appear in both sums decay exponentially fast, and only a few terms are needed in practice. We will observe in Sect. 3.1.3 that, on our test example, $A(y^*)$ is already a good approximation of $\langle A\rangle_N$. Hence, we have not implemented the strategy just described.

3.1.2 More than Two Repatoms

We explain in this section how the results of Sect. 3.1.1 can be adapted to the case when more than two repatoms are considered.

We thus consider the following setting: we have $N+M+1$ atoms of positions u^i, $0\leq i\leq M+N$, and u^0, u^N and u^{N+M} are the repatoms. We assume that N and M are such that $N/(N+M) = L_1$ is fixed. To remove translation invariance, we set

$u^0 = 0$. Since the atoms on the right of u^N will not play the same role as those on the left, we denote their distance differently:

$$y_i := \frac{u^i - u^{i-1}}{h} \quad \forall 1 \le i \le N,$$
$$z_i := \frac{u^{i+1+N} - u^{i+N}}{h} \quad \forall 0 \le i \le M-1,$$

where $h = 1/(N+M)$. We assume that the observable A is a function of $u^N - u^0$ and $u^{N+M} - u^N$. Hence, the expectation value of A reads

$$\begin{aligned}\langle A\rangle_{N,M} = Z^{-1} \int_{\mathbb{R}^{N+M}} A\left(\frac{1}{N+M}\sum_{i=1}^{N} y_i, \frac{1}{N+M}\sum_{i=0}^{M-1} z_i\right) e^{-\frac{\beta}{2}W_1(y_1)} \\ \times \prod_{i=1}^{N-1} \overline{f}(y_i, y_{i+1}) e^{-\frac{\beta}{2}W_1(y_N)} e^{-\beta W_2(y_N+z_0)} e^{-\frac{\beta}{2}W_1(z_0)} \\ \times \prod_{i=0}^{M-2} \overline{f}(z_i, z_{i+1}) e^{-\frac{\beta}{2}W_1(z_{M-1})}\, dy\, dz, \end{aligned} \tag{3.9}$$

where, as before, we have set $\overline{f}(x,y) = \exp[-\beta W_2(x+y) - \frac{\beta}{2}W_1(x) - \frac{\beta}{2}W_1(y)]$. Here again, we may use ψ_1 that solves (3.4) in order to rewrite (3.9) as the expectation value of a function of two independent Markov chains. Indeed, ψ_1 and g being defined as before (see (3.4) and (3.5)), we have

$$\begin{aligned}\langle A\rangle_{N,M} &= Z^{-1} \int_{\mathbb{R}^{N+M}} A\left(\frac{1}{N+M}\sum_{i=1}^{N} y_i, \frac{1}{N+M}\sum_{i=0}^{M-1} z_i\right) \psi_1(y_1) e^{-\frac{\beta}{2}W_1(y_1)} \\ &\quad \times \prod_{i=1}^{N-1} g(y_i, y_{i+1}) \frac{e^{-\frac{\beta}{2}W_1(y_N)}}{\psi_1(y_N)} e^{-\beta W_2(y_N+z_0)} \psi_1(z_0) e^{-\frac{\beta}{2}W_1(z_0)} \\ &\quad \times \prod_{i=0}^{M-2} g(z_i, z_{i+1}) \frac{e^{-\frac{\beta}{2}W_1(z_{M-1})}}{\psi_1(z_{M-1})}\, dy\, dz \\ &= \mathbb{E}\left[A\left(\frac{1}{N+M}\sum_{i=1}^{N} Y_i, \frac{1}{N+M}\sum_{i=0}^{M-1} Z_i\right) \right. \\ &\quad \left. \times \frac{e^{-\frac{\beta}{2}W_1(Y_N) - \frac{\beta}{2}W_1(Z_{M-1}) - \beta W_2(Y_N+Z_0)}}{\psi_1(Y_N)\psi_1(Z_{M-1})} \right] \\ &\quad \times \left(\mathbb{E}\left[\frac{e^{-\frac{\beta}{2}W_1(Y_N) - \frac{\beta}{2}W_1(Z_{M-1}) - \beta W_2(Y_N+Z_0)}}{\psi_1(Y_N)\psi_1(Z_{M-1})} \right] \right)^{-1}, \end{aligned} \tag{3.10}$$

where the sequences $(Y_i)_{i\geq 1}$ and $(Z_i)_{i\geq 0}$ are two independent realizations of a Markov chain of initial law $\psi_1 e^{-\frac{\beta}{2}W_1}$, and of transition kernel g. These Markov chains have exactly the same properties as the chain of Sect. 3.1.1. Hence, we may use again the ergodic theorem as before to prove that

$$\frac{1}{N+M}\sum_{i=1}^{N} Y_i \longrightarrow L_1 y^*, \qquad \frac{1}{N+M}\sum_{i=0}^{M-1} Z_i \longrightarrow L_2 y^*,$$

almost surely, with $L_1 = N/(N+M)$ and $L_2 = M/(N+M)$. Thus, the expectation values in (3.10) simplify, allowing to prove the following theorem.

Theorem 5 *Assume that A, W_1 and W_2 satisfy the assumptions of Theorem 3. Assume in addition that $L_1 = \frac{N}{N+M}$ is fixed, and set $L_2 = \frac{M}{N+M} = 1 - L_1$. Then we have*

$$\lim_{N,M\to+\infty} \langle A\rangle_{N,M} = A\big(L_1 y^*, L_2 y^*\big) \quad \text{with } y^* = \int_{\mathbb{R}} y\psi_1^2(y)\,dy.$$

Here again, it is possible to use an expansion in powers of $1/N$ and $1/M$ of the expectation value $\langle A\rangle_{N,M}$. For simplicity, we restrict ourselves to the expansion at first order, and consider the case $N = M$. We assume that the hypotheses of Theorem 4 and Theorem 5 are satisfied. We then have

$$\begin{aligned}\langle A\rangle_{N,N} = {} & A\big(L_1 y^*, L_2 y^*\big) + \frac{\sigma^2}{2N}\Delta A\big(L_1 y^*, L_2 y^*\big) \\ & + \frac{1}{N}\big[\partial_1 A\big(L_1 y^*, L_2 y^*\big) + \partial_2 A\big(L_1 y^*, L_2 y^*\big)\big]\sum_{i\geq 1}\mathbb{E}\big(Y_i - y^*\big) \\ & + O\left(\frac{1}{N^2}\right),\end{aligned}$$

with $L_1 = L_2 = 1/2$ and where σ is defined by (3.8). We have not implemented this formula, since on our test example, $A(L_1 y^*, L_2 y^*)$ is already a good approximation of $\langle A\rangle_{N,N}$.

3.1.3 Numerical Results

For the NNN model, we choose the potentials

$$W_1(x) = \frac{1}{2}(x-1)^4 + \frac{1}{2}x^2 \quad \text{and} \quad W_2(x) = \frac{1}{4}(x-2.1)^4.$$

Note that other choices are possible, such as $W_1 \equiv W_2$, or $W_2(x) = W_1(x/2)$ (such that the equilibrium distances of W_1 and W_2 are compatible). We have chosen W_2 such that we observe a significant dependence of ensemble averages (for instance of the mean length $\langle u^N\rangle_N$ of the chain) with respect to temperature.

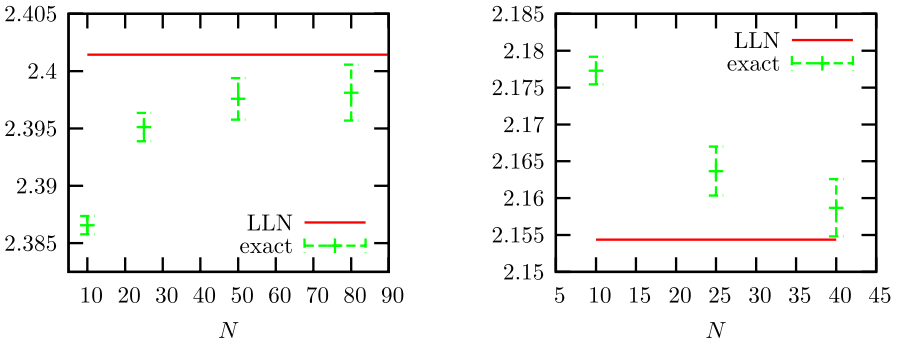

Fig. 13 *Left graph*: convergence, as N increases, of $\langle A(u^N)\rangle_N$ (exact) to $A(y^*)$ (LLN), at the temperature $1/\beta = 1$, for $A(x) = \exp(x)$. *Right graph*: convergence, as N increases, of $\langle A(u^N, u^{2N} - u^N)\rangle_{2N}$ (exact) to $A(y^*/2, y^*/2)$ (LLN), at the temperature $1/\beta = 1$, for $A(x, y) = \exp(2x(x + y))$

It is important that W_1 and W_2 grow fast enough at infinity, such that assumptions of Theorem 3 are satisfied (in particular assumption (3.6)). As in the NN case, we do not need any convexity assumption on W_1 and W_2.

We consider two cases:

- the chain consists of $N + 1$ atoms, there are two repatoms $u^0 = 0$ and u^N, and the observable only depends on the right-end atom u^N. We aim at computing $\langle A(u^N)\rangle_N$. This is the situation of Sect. 3.1.1.
- the chain consists of $2N + 1$ atoms, there are three repatoms $u^0 = 0$, u^N and u^{2N}, and the observable depends on u^N and $u^{2N} - u^N$. We aim at computing $\langle A(u^N, u^{2N} - u^N)\rangle_{2N}$. This is a situation covered by Sect. 3.1.2.

Theorems 3 and 5, respectively, provide the asymptotics $\lim_{N\to+\infty}\langle A(u^N)\rangle_N = A(y^*)$ and $\lim_{N\to+\infty}\langle A(u^N, u^{2N} - u^N)\rangle_{2N} = A(\frac{1}{2}y^*, \frac{1}{2}y^*)$.

We first study the convergence of ensemble averages at the temperature $1/\beta = 1$, as N increases. Results are shown on Fig. 13, for a particular choice of observable (we have performed the same tests with other observables, with similar conclusions). We indeed observe that the ensemble averages of the full atom system converge to their Law of Large Numbers (LLN) limit, in both cases of two and three repatoms. Note that the exact result for $N = 10$ is already very well approximated by the asymptotic limit, namely $A(y^*)$ in the two repatoms case, $A(y^*/2, y^*/2)$ in the three repatoms case.

We next study the averages as functions of the temperature, for $N = 100$. Results are shown on Fig. 14. We observe an excellent agreement between the full atom value and the asymptotic limit, in both cases of two and three repatoms, whatever the temperature.

3.2 Extensions

In this section, we briefly explain that our strategy carries out to more general cases.

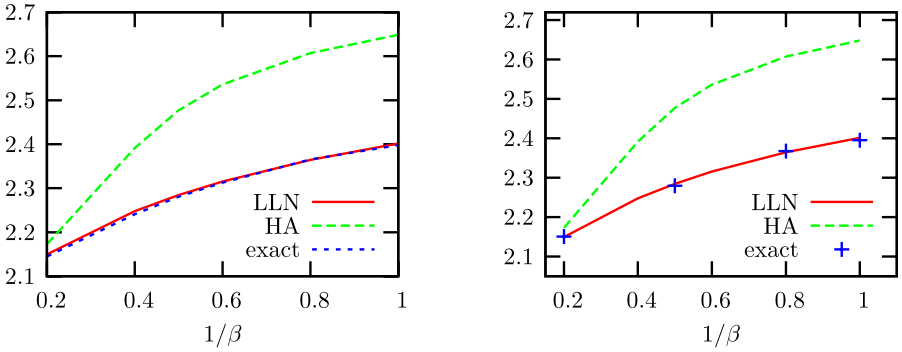

Fig. 14 *Left graph*: we plot $\langle A(u^N)\rangle_N$ (exact), $\langle A(u^N)\rangle_N^{\mathrm{HA}}$ (HA) and $A(y^*)$ (LLN) as functions of the temperature $1/\beta$ ($N = 100$, $A(x) = \exp(x)$). *Right graph*: we plot $\langle A(u^N, u^{2N} - u^N)\rangle_{2N}$ (exact), $\langle A(u^N, u^{2N} - u^N)\rangle_{2N}^{\mathrm{HA}}$ (HA) and $A(y^*/2, y^*/2)$ (LLN) as functions of the temperature $1/\beta$ ($N = 100$, $A(x, y) = \exp(2x)$)

3.2.1 The NNNN Case

The case of any finite range interaction may be treated in the same way as we treated the NNN case in Sect. 3.1.1. Indeed, consider for instance the case of next to next to nearest neighbor interaction (NNNN). In such a case, we are lead to consider (we go back here to the case of 2 repatoms for the sake of clarity):

$$\begin{aligned}\langle A\rangle_N &= Z^{-1}\int_{\mathbb{R}^N} A\left(\frac{1}{N}\sum_{i=1}^N y_i\right)\exp\Bigg[-\beta\sum_{i=1}^N W_1(y_i)\\ &\qquad -\beta\sum_{i=1}^{N-1} W_2(y_i+y_{i+1}) - \beta\sum_{i=1}^{N-2} W_3(y_i+y_{i+1}+y_{i+2})\Bigg]dy_1\cdots dy_N\\ &= Z^{-1}\int_{\mathbb{R}^N} A\left(\frac{1}{N}\sum_{i=1}^N y_i\right) b(y_{N-1}, y_N)\prod_{i=1}^{N-2} f(y_i, y_{i+1}, y_{i+2})\,dy_1\cdots dy_N,\end{aligned}$$

where we have set $f(x, y, z) = \exp[-\beta W_1(x) - \beta W_2(x+y) - \beta W_3(x+y+z)]$, and used the notation $b(y_{N-1}, y_N) = \exp[-\beta W_1(y_{N-1}) - \beta W_1(y_N) - \beta W_2(y_{N-1}+y_N)]$ for the boundary term. We assume, for the sake of simplicity, that N is even, i.e.,

$$N = 2M,$$

and define the new variables

$$\xi_i = (y_{2i-1}, y_{2i}), \quad 1 \le i \le M.$$

Hence, we have

$$\langle A\rangle_N = Z^{-1}\int_{\mathbb{R}^{2M}} A\left(\frac{1}{2M}\sum_{i=1}^M \xi_i\cdot(1,1)\right) b(\xi_M)\prod_{i=1}^{M-1}\widetilde{f}[\xi_i, \xi_{i+1}]\,d\xi_1\cdots d\xi_M, \quad (3.11)$$

where

$$\widetilde{f}\big[(x,y),(z,t)\big] = f(x,y,z)f(y,z,t).$$

Hence, this change of variables allows to manipulate again a Markov chain, but in dimension 2. However, the renormalization trick we have used in the NNN case (see Sect. 3.1.1) cannot be used here, because it relies on the fact that the transition operator is self-adjoint. It is nevertheless possible to use the above structure in the following way: define the operator (on $L^2(\mathbb{R}^2)$)

$$[P\varphi](z,t) = \int_{\mathbb{R}^2} \widetilde{f}\big[(x,y),(z,t)\big]\varphi(x,y)\,dx\,dy,$$

together with its adjoint

$$\big[P^*\varphi\big](z,t) = \int_{\mathbb{R}^2} \widetilde{f}\big[(z,t),(x,y)\big]\varphi(x,y)\,dx\,dy.$$

Before studying the average (3.11), let us consider the average

$$\langle B\rangle = Z^{-1}\int_{\mathbb{R}^{2M}} B(\xi_i)b(\xi_M)\prod_{i=1}^{M-1}\widetilde{f}[\xi_i,\xi_{i+1}]\,d\xi_1\cdots d\xi_M$$

for a continuous and bounded function B. Then

$$\langle B\rangle = Z^{-1}\int_{\mathbb{R}^2} B(\xi)\big[P^{i-2}\varphi\big](\xi)\big[\big(P^*\big)^{M-i-1}\psi\big](\xi)\,d\xi, \tag{3.12}$$

where

$$\varphi(z,t) = \int_{\mathbb{R}^2}\widetilde{f}\big[(x,y),(z,t)\big]\,dx\,dy,$$
$$\psi(t,z) = \int_{\mathbb{R}^2} b(y,x)\widetilde{f}\big[(t,z),(y,x)\big]\,dx\,dy.$$

We assume that the operators P and P^* have a simple and isolated largest eigenvalue (which can be proved for many interactions, using for instance Krein–Rutman theorem, Schaefer and Wolff 1999). Let us denote by ϕ and ϕ^* the corresponding eigenvectors in $L^2(\mathbb{R}^2)$, namely

$$P\phi = \lambda\phi, \qquad P^*\phi^* = \lambda\phi^*,$$

where $\lambda = \sup\mathrm{Spectrum}(P) = \sup\mathrm{Spectrum}(P^*)$. If $1 \ll i \ll M$, we infer from (3.12) that

$$\langle B\rangle \longrightarrow Z_\infty^{-1}\int_{\mathbb{R}^2} B(\xi)\phi(\xi)\phi^*(\xi)\,d\xi,$$

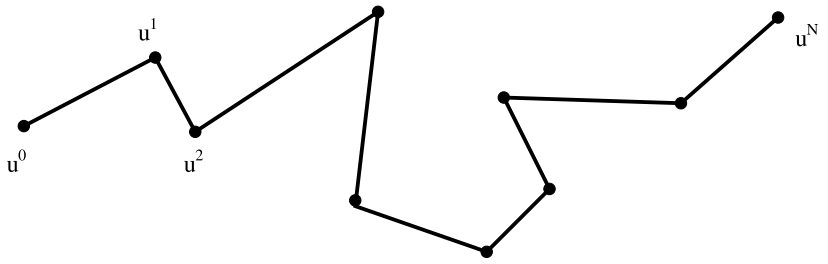

Fig. 15 An example of polymer chain. The corresponding model is the same as the one studied in this article, except that the positions u^i are in $\mathbb{R}^3$

where $Z_\infty = \int_{\mathbb{R}^2} \phi\,\phi^*$. This argument may be adapted to prove that the expectation value $\langle A\rangle_N$ defined by (3.11) converges:

$$\langle A\rangle_N \underset{N\to+\infty}{\longrightarrow} A\big(y^*\big),$$

where

$$y^* = \frac{\int_{\mathbb{R}^2}(\xi_1+\xi_2)\phi(\xi_1,\xi_2)\phi^*(\xi_1,\xi_2)\,d\xi_1\,d\xi_2}{\int_{\mathbb{R}^2}\phi(\xi_1,\xi_2)\phi^*(\xi_1,\xi_2)\,d\xi_1\,d\xi_2}.$$

3.2.2 Polymer Chains

The considerations of Sect. 2 and Sect. 3.1 may be easily generalized to the case when the positions u^i of the atoms are not restricted to be in the real line, but are vectors of $\mathbb{R}^2$ or $\mathbb{R}^3$. The only important thing here is that they are indexed by a one-dimensional parameter (here, $1 \le i \le N$). This is the case for instance if one considers a polymer chain (see Fig. 15). In such a case, NN or NNN approximations are commonly used, in order to compute the average length of the chain (see Bird et al. 1987 for instance). Our approximation strategy carries out to this case.

Acknowledgements Claude Le Bris and Frédéric Legoll would like to acknowledge stimulating discussions with Mitchell Luskin and Ellad Tadmor, University of Minnesota, Minneapolis. The hospitality of the Institute for Mathematics and its Applications (Minneapolis) and of the Weierstrass-Institut für Angewandte Analysis und Stochastik (Berlin) is gratefully acknowledged. Mathias Rousset provided some instrumental input from probability theory. This work has been initiated during the one-year stay (2005–2006) of the fourth author, Carsten Patz, at the Ecole Nationale des Ponts et Chaussées, as a visiting graduate student.

The research of Claude Le Bris and Frédéric Legoll is supported in part by the MEGAS nonthematic program (Agence Nationale de la Recherche, France) and by the INRIA, under the grant "Action de Recherche Collaborative" HYBRID.

References

Anitescu, M., Negrut, D., Zapol, P., El-Azab, A.: A note on the regularity of reduced models obtained by nonlocal quasi-continuum-like approach. Math. Program. **118**(2), 207–236 (2009)

Arndt, M., Griebel, M.: Derivation of higher order gradient continuum models from atomistic models for crystalline solids. SIAM J. Multiscale Model. Simul. **4**(2), 531–562 (2005)

Arndt, M., Luskin, M.: Error estimation and atomistic-continuum adaptivity for the quasicontinuum approximation of a Frenkel–Kantorova model. SIAM J. Multiscale Model. Simul. **7**(1), 147–170 (2008)
Bahadur, R.R., Ranga Rao, R.: On deviations of the sample mean. Ann. Math. Stat. **31**, 1015–1027 (1960)
Baxter, R.J.: Exactly Solved Models in Statistical Mechanics. Academic Press, London (1982)
Bender, C.M., Orszag, S.A.: Advanced mathematical methods for scientists and engineers. In: International Series in Pure and Applied Mathematics. McGraw-Hill, New York (1978)
Blanc, X., Le Bris, C., Lions, P.-L.: From molecular models to continuum mechanics. Arch. Ration. Mech. Anal. **164**, 341–381 (2002)
Blanc, X., Le Bris, C., Legoll, F.: Analysis of a prototypical multiscale method coupling atomistic and continuum mechanics. Math. Model. Numer. Anal. **39**(4), 797–826 (2005)
Blanc, X., Le Bris, C., Legoll, F.: Analysis of a prototypical multiscale method coupling atomistic and continuum mechanics: the convex case. Acta Math. Appl. Sin. **23**(2), 209–216 (2007a)
Blanc, X., Le Bris, C., Lions, P.-L.: Atomistic to continuum limits for computational materials science. Math. Model. Numer. Anal. **41**(2), 391–426 (2007b)
Bird, R.B., Curtiss, C.F., Armstrong, R.C., Hassager, O.: Dynamics of Polymeric Liquids, Vol. 2: Kinetic Theory, 2nd edn. Wiley, New York (1987)
Cancès, E., Legoll, F., Stoltz, G.: Theoretical and numerical comparison of some sampling methods for molecular dynamics. Math. Model. Numer. Anal. **41**(2), 351–389 (2007)
Curtarolo, S., Ceder, G.: Dynamics of an inhomogeneously coarse grained multiscale system. Phys. Rev. Lett. **88**(25), 255504 (2002)
Deák, P., Frauenheim, Th., Pederson, M.R. (eds.): Computer Simulation of Materials at Atomic Level. Wiley, New York (2000)
Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. Jones and Bartlett Publishers, Boston (1993)
den Hollander, F.: Large Deviations. Fields Institute Monographs, vol. 14. American Mathematical Society, Providence (2000)
Derrida, B.: Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. J. Stat. Mech. P07023 (2007)
Dobson, M., Luskin, M.: Analysis of a force-based quasicontinuum approximation. Math. Model. Numer. Anal. **42**(1), 113–139 (2008)
Dobson, M., Luskin, M.: An analysis of the effect of ghost force oscillation on quasicontinuum error. Math. Model. Numer. Anal. **43**(3), 591–604 (2009a)
Dobson, M., Luskin, M.: An optimal order error analysis of the one-dimensional quasicontinuum approximation. SIAM J. Numer. Anal. **47**(4), 2455–2475 (2009b)
Dobson, M., Luskin, M., Ortner, C.: Stability, instability, and error of the force-based quasicontinuum approximation. arXiv preprint 0903.0610 (2009a)
Dobson, M., Luskin, M., Ortner, C.: Sharp stability estimates for the accurate prediction of instabilities by the quasicontinuum method. arXiv preprint 0905.2914 (2009b)
Dunford, N., Schwartz, J.T.: Linear Operators, vol. 2: Spectral Theory: Self Adjoint Operators in Hilbert Space. Wiley, New York (1963)
Dupuy, L.M., Tadmor, E.B., Miller, R.E., Phillips, R.: Finite temperature Quasicontinuum: molecular dynamics without all the atoms. Phys. Rev. Lett. **95**, 060202 (2005)
E, W., Ming, P.B.: Analysis of multiscale methods. J. Comput. Math. **22**(2), 210–219 (2004)
E, W., Ming, P.B.: Cauchy–Born rule and stability of crystals: static problems. Arch. Ration Mech. Anal. **183**(2), 241–297 (2007)
Ellis, R.S.: Entropy, Large Deviations, and Statistical Mechanics. Grundlehren der Mathematischen Wissenschaften, vol. 271. Springer, New York (1985a)
Ellis, R.S.: Large deviations and statistical mechanics. In: Particle Systems, Random Media and Large Deviations, Brunswick, Maine, 1984. Contemp. Math., vol. 41, pp. 101–123. American Mathematical Society, Providence (1985b)
Ellis, R.S.: An overview of the theory of large deviations and applications to statistical mechanics. Scand. Actuar. J. **1**, 97–142 (1995), Harald Cramer Symposium (Stockholm, 1993)
Knap, J., Ortiz, M.: An analysis of the quasicontinuum method. J. Mech. Phys. Solids **49**(9), 1899–1923 (2001)
Kushner, H.J., Clark, D.S.: Stochastic Approximation Methods for Constrained and Unconstrained Systems, Applied Mathematical Sciences, vol. 26. Springer, New York (1978)
Legoll, F.: Multiscale methods coupling atomistic and continuum mechanics: some examples of mathematical analysis. In: Emmrich, E., Wittbold, P. (eds.) Analytical and Numerical Aspects of Partial Differential Equations, pp. 193–245. de Gruyter, Berlin (2009)

LeSar, R., Najafabadi, R., Srolovitz, D.J.: Finite-temperature defect properties from free-energy minimization. Phys. Rev. Lett. **63**, 624–627 (1989)
Lin, P.: Theoretical and numerical analysis of the quasi-continuum approximation of a material particle model. Math. Comput. **72**, 657–675 (2003)
Lin, P.: Convergence analysis of a quasi-continuum approximation for a two-dimensional material. SIAM J. Numer. Anal. **45**(1), 313–332 (2007)
Lions, P.-L., Toscani, G.: A strengthened central limit theorem for smooth densities. J. Funct. Anal. **128**, 148–176 (1995)
Meyn, S.P., Tweedie, R.L.: Markov Chains and Stochastic Stability. Springer, Berlin (1993)
Miller, R., Tadmor, E.B.: The quasicontinuum method: overview, applications and current directions. J. Comput.-Aided Mater. Des. **9**, 203–239 (2002)
Miller, R., Tadmor, E.B., Phillips, R., Ortiz, M.: Quasicontinuum simulation of fracture at the atomic scale. Model. Simul. Mater. Sci. Eng. **6**, 607–638 (1998)
Olla, S.: Non-equilibrium macroscopic behaviour of chain of interacting oscillators. In: Lecture Notes from a Course at IHP (Paris), Fall 2007
Ortner, C., Süli, E.: Analysis of a quasicontinuum method in one dimension. Math. Model. Numer. Anal. **42**(1), 57–91 (2008)
Patz, C.: PhD dissertation (2009, in preparation)
Presutti, E.: Scaling Limits in Statistical Mechanics and Microstructures in Continuum Mechanics. Springer, Berlin (2008)
Schaefer, H., Wolff, M.P.: Topological Vector Spaces, 2nd edn. Graduate Texts in Mathematics, vol. 3. Springer, New York (1999)
Shenoy, V.B., Miller, R., Tadmor, E.B., Phillips, R., Ortiz, M.: Quasicontinuum models of interfacial structure and deformation. Phys. Rev. Lett. **80**(4), 742–745 (1998)
Shenoy, V.B., Miller, R., Tadmor, E.B., Rodney, D., Phillips, R., Ortiz, M.: An adaptive finite element approach to atomic-scale mechanics—the quasicontinuum method. J. Mech. Phys. Solids **47**, 611–642 (1999)
Tadmor, E.B., Phillips, R.: Mixed atomistic and continuum models of deformation in solids. Langmuir **12**, 4529–4534 (1996)
Tadmor, E.B., Ortiz, M., Phillips, R.: Quasicontinuum analysis of defects in solids. Philos. Mag. A **73**, 1529–1563 (1996)
Tadmor, E.B., Smith, G.S., Bernstein, N., Kaxiras, E.: Mixed finite element and atomistic formulation for complex crystals. Phys. Rev. B **59**(1), 235–245 (1999)
Varadhan, S.R.S.: Large Deviations and Applications. SIAM, Philadelphia (1984)