ESAIM: M2AN 52 (2018) 1315–1352
https://doi.org/10.1051/m2an/2017052

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

ON ASSESSING THE ACCURACY OF DEFECT FREE ENERGY COMPUTATIONS

Matthew Dobson[1], Manh Hong Duong[2,*] and Christoph Ortner[2]

Abstract. We develop a rigorous error analysis for coarse-graining of defect-formation free energy. For a one-dimensional constrained atomistic system, we establish the thermodynamic limit of the defect-formation free energy and obtain explicitly the rate of convergence. We then construct a sequence of coarse-grained energies with the same rate but significantly reduced computational cost. We illustrate our analytical results through explicit computations for the case of harmonic potentials and through numerical simulations.

Mathematics Subject Classification. 65G99, 74E15, 74S60.

Received April 3, 2016. Accepted September 28, 2017.

1. Introduction

Crystalline materials contain a variety of defects, such as vacancies, interstitials and dislocations. Macroscopic properties of materials are strongly dependent on the distribution of defects, in particular through the interaction between dislocations and other defects [4]. Meso-scopic models for defect interaction (*e.g.*, dislocation dynamics, point defect diffusion) usually take as input an atomistic simulation of a single, or few defects, from which the meso-scopic model parameters can be extracted. A prototypical example is the defect formation energy, which we discuss in more detail below. A great number of numerical schemes on spatial coarse-graining of the free energy have been developed in the literature, see for instance in [8,15] and references therein. However, a rigorous analysis on the accuracy of these schemes is still underdeveloped; we are only aware of the references [1,18].

In this paper, we provide such a rigorous analysis for the computations of the defect-formation free energy. We consider one-dimensional constrained atomistic systems, which model perfect and defect materials respectively, with degrees of freedom $u \in \mathbb{R}^N$. The system can be either influenced by external forces or not. In the case without external forces, free energies are respectively defined by

$$F_N(A) = -\beta^{-1} \log \int_{\mathbb{R}^{N-1}} \exp\Big[-\beta V(u)\Big]\,\mathrm{d}u_1 \dots \mathrm{d}u_{N-1}, \tag{1.1}$$

$$F_N^P(A) = -\beta^{-1} \log \int_{\mathbb{R}^{N-1}} \exp\Big[-\beta V^P(u))\Big]\,\mathrm{d}u_1 \dots \mathrm{d}u_{N-1}, \tag{1.2}$$

Keywords and phrases. Defect formation free energy, finite temperature, material defects, Cauchy–Born rule.

[1] Department of Mathematics and Statistics, UMass Amherst, 710 N Pleasant Street, Amherst, MA 01003, USA.

[2] Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

[*] Corresponding author: m.h.duong@warwick.ac.uk

Article published by EDP Sciences

© EDP Sciences, SMAI 2018

where

$$V(u) = \sum_{i=1}^{N} \psi(u_i - u_{i-1}), \quad V^P(u) = V(u) + P(u) \tag{1.3}$$

are the energies associated to the perfect and defect materials, V is the sum of bond energies $\psi(u_i - u_{i-1})$; $P : \mathbb{R}^N \to \mathbb{R}$ models the defect. For simplicity, we assume that P is a localised function and depends only on the first bond $P(u) = P(u_1 - u_0)$; the analysis may be easily adapted to the case of a defect in the bulk. Finally, $\beta > 0$ is the inverse of the temperature.

In the case with external forces, the perfect free energy is unchanged, but the deformed free energy is influenced by the external forces

$$F_N^P(A) = -\beta^{-1} \log \int_{\mathbb{R}^{N-1}} \exp\left[-\beta \sum_{i=1}^{N} \psi_i(u_i - u_{i-1}) - \beta P(u_1)\right] \mathrm{d}u_1 \ldots \mathrm{d}u_{N-1}, \tag{1.4}$$

where $\psi_i(y) = \psi(y) + h_i y$ with $\{h_i\}_{i=1}^N$ representing the external forces. The forces are included only in the defective energy in order to model a slowly decaying stress field surrounding a defect which is present in higher dimensions but not naturally present in the one dimensional case where elastic fields decay exponentially fast.

Note that the integrals (1.1), (1.2) and (1.4) are subjected to the boundary constraints

$$u_0 = 0, \quad u_N = NA \tag{1.5}$$

so that the free energies depend on N and A as shown, and $P(u) = P(u_1)$.

The main quantity of interest in this paper is the *defect-formation free energy* defined as the difference of the free energies

$$G_N(A) := F_N^P(A) - F_N(A) = -\frac{1}{\beta} \log \frac{\int_{\mathbb{R}^{N-1}} \exp(-\beta V^P(u))\, \mathrm{d}u}{\int_{\mathbb{R}^{N-1}} \exp(-\beta V(u))\, \mathrm{d}u}. \tag{1.6}$$

This quantity is used to obtain the equilibrium defect concentration [16,20] or to analyse defect clustering [12,17]. A direct computation of $G_N(A)$ is practically impossible due to the curse of dimensionality: one needs to compute integrals over $\mathbb{R}^{N-1}$, which is an extremely high-dimensional space.

As a matter of fact, N itself is an approximation parameter, the *exact* defect formation free energy is given by the thermodynamic limit, letting $N \to \infty$. Establishing this limit, and thus making precise what we mean by the "exact model" is the first result of our paper. Once we have established this, we search for an alternative scheme by which to approximate it, which yields an improved accuracy/computational cost ratio.

The computation of $\lim_N G_N$ is a problem that is interesting in its own right, but at the same time it serves as a natural benchmark problem for exploring the relative accuracy/cost of coarse-graining methods at finite temperature. We introduce and analyze a coarse-graining approach based on the use of a finite temperature Cauchy-Born energy density.

The work [1] considers a similar model as ours, but this work is focused on the scaling limit of the free energy, not the free energy difference, which is a different scale. Furthermore, it does not take defects into account. The work [18] is in spirit much closer to ours and in particular does take defects into account. The main difference to our work is that [18] considers "low" temperature *via* an asymptotic series expansion. Moreover, our coarse-grained model has some close similarities with common quasicontinuum-type models.

Technically, to prove our main results, we will link the defect–formation free energy to a ratio of the densities of certain random variables and employ techniques from statistical mechanics. The latter have been used in the literature, for example in [11, 14]. There is also a close relationship between our thermodynamic limit results and the Gibbs conditioning principle [7, 9]. We comment on this in Section 4. The connections to the defect–formation free energy, to the best of our knowledge, is new and moreover, some technical modifications of the mentioned papers were required.

1.1. Assumptions and main results

For simplicity of notation we set $\beta = 1$ throughout the paper. Moreover, we make the following standing assumptions on the bond energy ψ, the defect P and the external forces $\{h_i\}_{i=1}^N$.

Assumption 1.1. $\psi, P \in C^2(\mathbb{R})$ and there exist positive constants $\kappa_1 \leq \kappa_2$ and $\varsigma_1 \leq \varsigma_2$ such that

$$\kappa_1 \leq \psi'' \leq \kappa_2, \quad \varsigma_1 \leq (\psi + P)'' \leq \varsigma_2. \tag{1.7}$$

Assumption 1.2. $\mathbf{h} := (h_1, h_2 \ldots) \in l^1$; we can then define $H := \sum_{i=2}^{\infty} h_i$.

Step 1. Thermodynamic limit: Our first result concerns the *rate of convergence* of defect formation free energy. Its proof is given in Theorems 2.1 and 3.2.

Theorem 1.3. *There exists $G_\infty \in C^\infty(\mathbb{R})$, such that, for all $A \in \mathbb{R}$*

$$|G_N(A) - G_\infty(A)| \lesssim N^{-1}.$$

Step 2. Coarse-graining: To motivate and put our work in the context of coarse-graining of thermodynamic quantities, we first recall its general set up. Let $X = \mathbb{R}^{N+M}$ be a (microscopic) phase space endowed with a probability (Boltzmann-Gibbs) measure

$$\mu(\mathrm{d}x) = Z^{-1} \exp(-E(x))\, \mathrm{d}x.$$

We want to compute the average

$$\mathcal{A} := \int_X \Phi(x)\, \mu(\mathrm{d}x)$$

of an observable $\Phi : X \to \mathbb{R}$. Observables of interest are often functions of only part of the variable: for $x = (y, z)$ with $y \in \mathbb{R}^N, z \in \mathbb{R}^M$ then $\Phi(x) = \Phi(y, z) = \Phi(y)$. In this case the average above can be computed as an integral over a lower dimensional space, $\mathbb{R}^N$ instead of $\mathbb{R}^{N+M}$ using a *coarse-grained energy* $E_{cg} : \mathbb{R}^N \to \mathbb{R}$,

$$\begin{aligned}\mathcal{A} = \int_{\mathbb{R}^{N+M}} \Phi(y)\mu(\mathrm{d}y, \mathrm{d}z) &= \int_{\mathbb{R}^N} \Phi(y) \int_{\mathbb{R}^M} Z^{-1} \exp(-E(y,z))\, \mathrm{d}z \\ &= \int_{\mathbb{R}^N} \Phi(y)\tilde{Z}^{-1} \exp(-E_{cg}(y))\, \mathrm{d}y,\end{aligned}$$

where the coarse-grained energy E^{cg} is defined *via*

$$\tilde{Z}^{-1} \exp(-E^{cg}(y)) = \int_{\mathbb{R}^M} Z^{-1} \exp(-E(y,z))\, \mathrm{d}z.$$

However, it is often computationally intractable using the above definition. Instead, to compute E_{cg} in practice, one views the problem as minimizing the Helmholtz free energy of the system with y fixed, and approximating $E(y, z)$ above. For example, in [8, 19] the authors use a local harmonic approximation along with a quasicontinuum-coarse grained mesh to compute an approximate free energy $E_{\mathrm{cb}}(y, z)$

$$E^{cg}(y) = \inf_{z \in \mathbb{R}^M} E^{cb}(y, z), \tag{1.8}$$

for some $E^{cb} : \mathbb{R}^{N+M} \to \mathbb{R}$. This paper introduces a localized Cauchy-Born approximation in the chain and justifies its use for the defect computation.

In our setting, the defect–formation free energy (1.6) can be written as an observable average as follows

$$G_N(A) = -\log \int_{\mathbb{R}^{N-1}} \exp(-P)\mu_N(\mathrm{d}u),$$

where $\mu_N(\mathrm{d}u) = Z^{-1}\exp(-V(u))\mathrm{d}u$. Since the defect P is a localised function (in this paper, $P(u) = P(u_1 - u_0)$), we can apply the strategy described above to effectively compute $G_N(A)$. Theorem 1.4 indeed shows that the ansatz (1.8) becomes rigorous in the thermodynamic limit $N \to \infty$ with explicit E^{cb}, and Theorem 1.5 computes the approximation errors. To state these Theorems, we need to recall the Cauchy-Born strain energy which will appear throughout the rest of the paper.

The (finite-temperature) Cauchy–Born strain energy function is given by [2]

$$W(A) = \sup_{\sigma\in\mathbb{R}} \left\{ \sigma A - \log \int_{\mathbb{R}} \exp(-\psi(y) + \sigma y)\,\mathrm{d}y \right\}. \tag{1.9}$$

Taking a continuum model $\int [W(u') + hu']dx$ outside the defect core $\{0, 1\}$ and then discretising it with the atomistic grid $\{1, 2, \dots N\}$ we obtain

$$E_N^{\mathrm{cb}}(u) := \sum_{i=2}^{N} \Big[W(u_i') - W(A) + h_i u_i' \Big], \quad \text{where} \quad u_i' = u_i - u_{i-1}, \tag{1.10}$$

with admissible displacements $u : \{0, \dots, N\} \to \mathbb{R}$ satisfying $u_0 = 0, u_N = AN$. After replacing $u_i = Ai + v_i$, summation by parts, and taking the formal limit $N \to \infty$, yields

$$E^{\mathrm{cb}}(u) = W'(A)(A - u_1) + AH + \sum_{i=2}^{\infty} \Big[W(A + v_i') - W(A) - W'(A)v_i' + h_i v_i' \Big]. \tag{1.11}$$

It is important to note here that E^{cb} is formulated in a way that ensures it is well-defined for arguments with $v' \in \ell^2$.

We obtain the following characterisation of $G_\infty(A)$ in terms of E^{cb}.

Theorem 1.4. *Let* $E^{\mathrm{cg}}(A, y) := \inf_{u\in\mathbb{R}^{\mathbb{N}}, u_1 = y} E^{\mathrm{cb}}(u)$*, then*

$$G_\infty(A) = -\log \frac{\int_{\mathbb{R}} \exp\big(-P(y) - \psi_1(y) - E^{\mathrm{cg}}(A, y)\big)\mathrm{d}y}{\int_{\mathbb{R}} \exp\big(-\psi(y) - E^{\mathrm{cg}}_{\mathbf{h}=\mathbf{0}}(A, y)\big)\mathrm{d}y}. \tag{1.12}$$

where $E^{\mathrm{cg}}_{\mathbf{h}=\mathbf{0}}$ *denotes the coarse-grained energy with* $h_j \equiv 0$.

In the absence of external forces, Theorem 1.4 can be derived from the Gibbs conditioning principle [7, 9]. However, it is not clear how to do so when there are external forces. We compare the technique we employ with the Gibbs conditioning principle in more details in Section 4.

Step 3. Approximation: Thus, we have replaced a limit of high-dimensional integrals by a one-dimensional integral over a coarse-grained energy functional whose evaluation requires the solution of an infinite-dimensional variational problem. In our next step, we replace $E^{\mathrm{cg}}(A, y)$ with a finite-dimensional approximation.

Let

$$E_N^{\mathrm{cg}}(A, y) := \inf_{\substack{u\in\mathbb{R}^N \\ u_1 = y, u_N = NA}} E_N^{\mathrm{cb}}(u)$$

and

$$G_N^{\mathrm{cg}}(A) := -\log \frac{\int_{\mathbb{R}} \exp\big(-P(y) - \psi_1(y) - E_N^{\mathrm{cg}}(A, y)\big)\mathrm{d}y}{\int_{\mathbb{R}} \exp\big(-\psi(y) - E^{\mathrm{cg}}_{N,\mathbf{h}=\mathbf{0}}(A, y)\big)\mathrm{d}y}.$$

Here we have chosen E_N^{cg} as the most basic approximation scheme to E^{cg}, but far more sophisticated choices could be explored. With this definition we obtain the following result.

Theorem 1.5.
(i) $G_N^{\mathrm{cg}}(A)$ *is well-defined for all* $A \in \mathbb{R}$.
(ii) *For all* $A \in \mathbb{R}$ *we have the estimate*

$$\left|G_N^{\mathrm{cg}}(A) - G_\infty(A)\right| \lesssim N^{-1}.$$

The sharpness of the results of Theorems 1.3 and 1.4 are demonstrated through explicit computations in the harmonic case $\psi(y) = \alpha|y|^2$ and $P(y) = \beta|y|^2$ in Section 5 and in numerical simulations in Section 6.

Interpretation: Statements (ii) of Theorems 1.4 and 1.5 imply that $G_\infty(A)$ can be computed from two one-dimensional integrals, but this extreme reduction of computational complexity is only due to the special one-dimensional structure of our model problem and cannot in general be reproduced.

The structure in our construction that can be expected more generally though is that $G_\infty(A)$ can be *approximated by* a low-dimensional canonical average with respect to a coarse-grained energy that is obtained by a variational problem in the exterior of the computational domain.

In our case the coarse-grained measure is one–dimensional but in general one may still expect it to be relatively low–dimensional. A Langevin or other type of Markov–Chain type algorithm can now be employed to compute $G_\infty(A)$; *cf.* Section 6.

Of course, the evaluation of $E^{\mathrm{cg}}(y)$ is in general impossible, and an approximation needs to be performed. For example, $E_N^{\mathrm{cg}}(A, y)$ (and its derivatives) is computable with a reasonably low $O(N)$ cost. Note that W itself may be costly to evaluate, but it could be easily precomputed to high accuracy *e.g. via* Taylor expansions or spline techniques. The $O(N)$ cost could be reduced further if we employ a quasi-continuum style coarse-graining of E_N^{cg}.

1.2. Organisation of the paper

The rest of the paper is structured as follows. In Section 2 we study the case without external forces. Extension to the case with external forces is shown in Section 3. In Section 4 we compare our work with Gibbs conditioning principle. In Section 5, we provide explicit computations for the harmonic case. Finally, in Section 6, we present some numerical simulations.

2. The case without external forces

In this section, we analyse the case without external forces.

2.1. Thermodynamic limit

In this section, we prove Theorem 1.3 for the case without external forces by establishing the existence of the thermodynamic limit G_∞ and the rate of convergence of G_N to G_∞. We give an expression for G_∞ in terms of the Cauchy-Born strain energy (1.9), which arises here as in the Gibbs conditioning principle, see Section 4. The main result of this section is the following theorem.

Theorem 2.1. *Suppose that Assumption 1.1 is satisfied. Then the thermodynamic limit is given by*

$$G_\infty(A) = -\log \frac{\int_{\mathbb{R}} \exp\left[-(\psi + P)(y) + W'(A)y\right]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp\left[-\psi(y) + W'(A)y\right]\,\mathrm{d}y}. \tag{2.1}$$

Moreover, for all $A \in \mathbb{R}$, *we have the estimate*

$$|G_N(A) - G_\infty(A)| \lesssim N^{-1}. \tag{2.2}$$

Proof. The proof is split into three steps that are Proposition 2.3, Proposition 2.7 and Proposition 2.10 below. □

We start with the following auxiliary lemma that links the free energy to the density of an average of independent random variables. This lemma will be applied in Proposition 2.3 and Theorem 3.2 later on.

Lemma 2.2. *Suppose that* $\tilde{\psi}_i \in C^2(\mathbb{R})$ *and* $0 < \kappa_1 \le \tilde{\psi}_i'' \le \kappa_2$ *for* $i = 1, \ldots, N$. *We define*

$$\tilde{W}_N(A) = \sup_{\sigma\in\mathbb{R}} \left\{ \sigma A - \frac{1}{N}\sum_{i=1}^{N} \log \int_{\mathbb{R}} \exp(-\tilde{\psi}_i(y) + \sigma y)\,\mathrm{d}y \right\}, \tag{2.3}$$

$$\tilde{F}_N(A) = -\log \int_{\mathbb{R}^{N-1}} \exp\left[-\sum_{i=1}^{N} \tilde{\psi}_i(u_i - u_{i-1}) \right] \mathrm{d}u_1 \ldots \mathrm{d}u_{N-1}, \tag{2.4}$$

with $u_0 = 0, u_N = NA$.

Let σ^* *be the maximizer in* (2.3). *We define the one dimensional probability measures*

$$\tilde{\mu}_i^{\sigma^*}(\mathrm{d}y) = Z_i^{-1} \exp(\sigma^* y - \tilde{\psi}_i(y))\,\mathrm{d}y, \tag{2.5}$$

where Z_i *is the normalising constant. Let* $\tilde{X}_i$ *be independent random variables distributed according to* $\tilde{\mu}_i^{\sigma*}$ *and let* $\tilde{m}_i$ *be the mean of* $\tilde{X}_i$. *Let* $\tilde{g}_{N,A}$ *be the density of* $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\tilde{X}_i - \tilde{m}_i)$. *Then it holds that*

$$\tilde{F}_N(A) = \frac{1}{2}\log N + N\tilde{W}_N(A) - \log \tilde{g}_{N,A}(0). \tag{2.6}$$

Proof. This proof is adapted from Lemma 8 of [14] (see also [11], Eq. (125)). By change of variables $y_i = u_i - u_{i-1}$, for $i = 1, \ldots, N-1$, we can re-write $\tilde{F}_N(A)$ as

$$\tilde{F}_N(A) = -\log \int_{\mathbb{R}^{N-1}} \exp\left[-\sum_{i=1}^{N-1} \tilde{\psi}_i(y_i) - \tilde{\psi}_N\left(NA - \sum_{i=1}^{N-1} y_i \right) \right] \mathrm{d}y_1 \ldots \mathrm{d}y_{N-1}. \tag{2.7}$$

We define

$$\tilde{\varphi}_{N,i}(\sigma) = \log \int_{\mathbb{R}} \exp[-\tilde{\psi}_i(y) + \sigma y]\,\mathrm{d}y,$$

$$\tilde{\varphi}_N(\sigma) := \frac{1}{N}\log \int_{\mathbb{R}^N} \exp\left[-\sum_{i=1}^{N} \tilde{\psi}_i(y_i) + \sigma \sum_{i=1}^{N} y_i \right] \mathrm{d}y_1 \ldots \mathrm{d}y_N$$

then

$$\begin{cases} \tilde{W}_N(A) = \sigma^* A - \tilde{\varphi}_N(\sigma^*), \\ A = \dfrac{\mathrm{d}}{\mathrm{d}\sigma}\tilde{\varphi}_N(\sigma)\Big|_{\sigma=\sigma^*}. \end{cases} \tag{2.8}$$

We have

$$\begin{aligned} \tilde{\varphi}_N(\sigma) &= \frac{1}{N}\log \int_{\mathbb{R}^N} \exp\left[-\sum_{i=1}^{N} \tilde{\psi}_i(y_i) + \sigma \sum_{i=1}^{N} y_i \right] \mathrm{d}y_1 \ldots \mathrm{d}y_N \\ &= \frac{1}{N}\sum_{i=1}^{N} \log \int_{\mathbb{R}} \exp\left[-\tilde{\psi}_i(y_i) + \sigma y_i \right] \mathrm{d}y_i \\ &= \frac{1}{N}\sum_{i=1}^{N} \tilde{\varphi}_{N,i}(\sigma). \end{aligned} \tag{2.9}$$

A straightforward calculation gives

$$\tilde{m}_i = \int_{\mathbb{R}} y_i \tilde{\mu}_i^{\sigma^*}(\mathrm{d}y_i) = \frac{\mathrm{d}}{\mathrm{d}\sigma}\tilde{\varphi}_{N,i}(\sigma)\Big|_{\sigma=\sigma^*}. \tag{2.10}$$

Substituting (2.9) and (2.10) into (2.8), we obtain

$$A = \frac{\mathrm{d}}{\mathrm{d}\sigma}\tilde{\varphi}_N(\sigma)\Bigg|_{\sigma=\sigma^*} = \frac{1}{N}\sum_{i=1}^N \frac{\mathrm{d}}{\mathrm{d}\sigma}\tilde{\varphi}_{N,i}(\sigma)\Bigg|_{\sigma=\sigma^*} = \frac{1}{N}\sum_{i=1}^N \tilde{m}_i. \tag{2.11}$$

Since $\tilde{X}_i$ are independent, the density of the sum $\sum_{i=1}^N \tilde{X}_i$ is given by the convolution

$$\tilde{f}_{\sum_{i=1}^N X_i}(\xi) = (\tilde{\mu}_1^{\sigma^*} * \ldots * \tilde{\mu}_N^{\sigma^*})(\xi).$$

Using the definition of convolution, we can compute the above density explicitly as follows

$$\tilde{f}_{\sum_{i=1}^N \tilde{X}_i}(\xi) = \int_{\mathbb{R}^{N-1}} \exp\left[-\sum_{i=1}^N \tilde{\varphi}_{N,i}(\sigma^*) + \sigma^*\xi - \tilde{\psi}_N(\xi - \sum_{i=1}^{N-1} y_i) - \sum_{i=1}^{N-1}\tilde{\psi}_i(y_i)\right] \mathrm{d}y_1 \ldots \mathrm{d}y_{N-1}.$$

We recall that if Y has density $f(y)\mathrm{d}y$ then, for $\alpha > 0, \beta \in \mathbb{R}$, $\alpha Y + \beta$ has density $\frac{1}{\alpha}f(\frac{y-\beta}{\alpha})$. Hence, we obtain

$$\begin{aligned}
\tilde{g}_{N,A}(\xi) &= f_{\frac{1}{\sqrt{N}}\sum_{i=1}^N(\tilde{X}_i - m_i)}(\xi) \\
&= \sqrt{N}\int_{\mathbb{R}^{N-1}} \exp\left[-\sum_{i=1}^N \tilde{\varphi}_{N,i}(\sigma^*) + \sigma^*\left(\xi\sqrt{N} + \sum_{i=1}^N \tilde{m}_i\right)\right. \\
&\quad \left. -\tilde{\psi}_N\left(\sqrt{N}\xi - \sum_{i=1}^{N-1} y_i + \sum_{i=1}^N \tilde{m}_i\right) - \sum_{i=1}^{N-1}\tilde{\psi}_i(y_i)\right] \mathrm{d}y_1 \ldots \mathrm{d}y_{N-1}.
\end{aligned}$$

In particular, using (2.9), (2.7) and (2.11), we get

$$\begin{aligned}
\tilde{g}_{N,A}(0) &= \sqrt{N}\int_{\mathbb{R}^{N-1}} \exp\left[-\sum_{i=1}^N \tilde{\varphi}_{N,i}(\sigma^*) + \sigma^*\sum_{i=1}^N \tilde{m}_i - \tilde{\psi}_N\left(-\sum_{i=1}^{N-1} y_i + \sum_{i=1}^N \tilde{m}_i\right) - \sum_{i=1}^{N-1}\tilde{\psi}_i(y_i)\right] \mathrm{d}y_1 \ldots \mathrm{d}y_{N-1} \\
&= \sqrt{N}\int_{\mathbb{R}^{N-1}} \exp\left[-N\tilde{\varphi}_N(\sigma^*) + \sigma^* NA - \tilde{\psi}_N\left(NA - \sum_{i=1}^{N-1} y_i\right) - \sum_{i=1}^{N-1}\tilde{\psi}_i(y_i)\right] \mathrm{d}y_1 \ldots \mathrm{d}y_{N-1} \\
&= \sqrt{N}\exp[N(\sigma^* A - \tilde{\varphi}_N(\sigma^*))]\int_{\mathbb{R}^{N-1}} \exp\left[-\tilde{\psi}_N\left(NA - \sum_{i=1}^{N-1} y_i\right) - \sum_{i=1}^{N-1}\tilde{\psi}_i(y_i)\right] \mathrm{d}y_1 \ldots \mathrm{d}y_{N-1} \\
&= \sqrt{N}\exp[N(\sigma^* A - \tilde{\varphi}_N(\sigma^*))]\exp[-\tilde{F}_N(A)].
\end{aligned}$$

It follows from (2.8) and the above equality that

$$\log \tilde{g}_{N,A}(0) = \frac{1}{2}\log N + N\tilde{W}_N(A) - \tilde{F}_N(A),$$

which is equivalent to (2.6) as claimed. □

The following proposition provides an analytical expression of the defect–formation free energy in terms of densities of averages of independent random variables.

Proposition 2.3. *Recall that the Cauchy–Born energy is given by*

$$W(A) = \sup_{\sigma\in\mathbb{R}} \left\{ \sigma A - \log \int_{\mathbb{R}} \exp(-\psi(y) + \sigma y)\,\mathrm{d}y \right\}. \tag{2.12}$$

We define an analogous function that is associated to the defect material

$$W_N^P(A) = \sup_{\sigma\in\mathbb{R}} \left\{ \sigma A - \frac{1}{N}\Big(\log \int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma y]\,\mathrm{d}y + (N-1)\log \int_{\mathbb{R}} \exp(-\psi(y)+\sigma y)\,\mathrm{d}y \Big) \right\}. \tag{2.13}$$

Let σ_0 and σ_P^N be the maximisers in definitions of (2.12) and (2.13) respectively. We define the one-dimensional probability measures

$$\mu^{\sigma_0}(\mathrm{d}y) = Z_\mu^{-1}\exp(\sigma_0 y - \psi(y))\,\mathrm{d}y, \quad \text{and} \tag{2.14}$$

$$\nu^{\sigma_P^N}(\mathrm{d}y) = Z_\nu^{-1}\exp(\sigma_P^N y - (\psi+P)(y))\,\mathrm{d}y, \quad \mu^{\sigma_P^N}(\mathrm{d}y) = Z_{\mu_P}^{-1}\exp(\sigma_P^N y - \psi(y))\,\mathrm{d}y, \tag{2.15}$$

where Z_μ, Z_ν and Z_{μ_P} are normalising constants. Let $m, m_{P,1}$ and $m_{P,2}$ be respectively the means of $\mu^{\sigma_0}, \nu^{\sigma_P^N}(\mathrm{d}y)$ and $\mu^{\sigma_P^N}(\mathrm{d}y)$.

Let $\{X_i\}_{i=1,\ldots,N}$ and $\{Y_i\}_{i=1,\ldots,N}$ be independent random variables, where $\{X_i\}_{i=1,\ldots,N}$ distributed according to $\mu^{\sigma_0}(\mathrm{d}y)$, $\{Y_1\}$ distributed according to $\nu^{\sigma_P^N}(\mathrm{d}y)$, and $\{Y_i\}_{i=2,\ldots,N}$ distributed according to $\mu^{\sigma_P^N}(\mathrm{d}y)$. Let $g_{N,A}$ and $g_{N,A}^P$ be respectively the density of $\frac{1}{\sqrt{N}}\sum_{i=1}^N (X_i - m)$ and $\frac{1}{\sqrt{N}}\sum_{i=1}^N (Y_i - m_{P,i})$ (with $m_{P,2} = \ldots = m_{P,N}$).

Then it holds that

$$F_N^P(A) - F_N(A) = N[W_N^P(A) - W(A)] + \log\frac{g_{N,A}(0)}{g_{N,A}^P(0)}. \tag{2.16}$$

Proof. Applying Lemma 2.2 for the cases $\tilde\psi_i = \psi \quad (i=1,\ldots,N)$ and $\tilde\psi_1 = \psi + P, \quad \tilde\psi_i = \psi \quad (i=2,\ldots,N)$, we obtain the following relations respectively

$$F_N(A) = \frac{1}{2}\log N + NW_N(A) - \log g_{N,A}(0),$$

$$F_N^P(A) = \frac{1}{2}\log N + NW_N^P(A) - \log g_{N,A}^P(0).$$

The assertion (2.16) immediately follows from these two relations. □

The next step is to passing to the limit $N\to\infty$ for each term in the relation (2.16). We will need some auxiliary lemmas. We define

$$\Psi(\sigma) := \frac{\int_{\mathbb{R}} y\exp(-\psi(y)+\sigma y)\,\mathrm{d}y}{\int_{\mathbb{R}}\exp(-\psi(y)+\sigma y)\,\mathrm{d}y}, \tag{2.17}$$

$$\Phi(\sigma) := \frac{\int_{\mathbb{R}} y\exp[-(\psi+P)(y)+\sigma y]\,\mathrm{d}y}{\int_{\mathbb{R}}\exp[-(\psi+P)(y)+\sigma y]\,\mathrm{d}y} - \frac{\int_{\mathbb{R}} y\exp(-\psi(y)+\sigma y)\,\mathrm{d}y}{\int_{\mathbb{R}}\exp(-\psi(y)+\sigma y)\,\mathrm{d}y}.$$

The following lemma on boundedness of derivatives of Ψ and Φ will be used several times in the sequel.

Lemma 2.4. *It holds that*

$$\frac{1}{\kappa_2} \le \frac{\mathrm{d}}{\mathrm{d}\sigma}\Psi(\sigma) \le \frac{1}{\kappa_1} \quad \text{and} \quad \left|\frac{\mathrm{d}}{\mathrm{d}\sigma}\Phi(\sigma)\right| \le C, \tag{2.18}$$

for some positive constant C.

Proof. We first prove the first part of (2.18). The following proof is simplified from Lemma 2.4 of [3]. In [3], (Lem. 2.4) the author has actually proved a stronger result than we need here. We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\sigma}\Psi(\sigma) &= \frac{\left(\int_{\mathbb{R}} y^2 \exp(-\psi(y)+\sigma y)\,\mathrm{d}y\right)\left(\int_{\mathbb{R}} \exp(-\psi(y)+\sigma y)\,\mathrm{d}y\right) - \left(\int_{\mathbb{R}} y \exp(-\psi(y)+\sigma y)\,\mathrm{d}y\right)^2}{\left(\int_{\mathbb{R}} \exp(-\psi(y)+\sigma y)\,\mathrm{d}y\right)^2}\\ &= \int_{\mathbb{R}} \Big(y - m_\sigma\Big)^2 \mu_\sigma(\mathrm{d}y), \end{aligned} \tag{2.19}$$

where

$$\mu_\sigma(\mathrm{d}y) := \frac{\exp(-\psi(y)+\sigma y)}{\int_{\mathbb{R}} \exp(-\psi(y)+\sigma y)\,\mathrm{d}y}\,\mathrm{d}y \in \mathcal{P}(\mathbb{R}), \quad \text{and} \quad m_\sigma = \int_{\mathbb{R}} y\mu_\sigma(\mathrm{d}y).$$

Using this equality, we now estimate $\frac{\mathrm{d}}{\mathrm{d}\sigma}\Psi(\sigma)$ using assumptions on ψ. For the upper bound: since $\psi'' \geq \kappa_1$, μ_σ satisfies the Poincare inequality with constant κ_1 uniformly in σ. Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\Psi(\sigma) \leq \frac{1}{\kappa_1}\int \Big|\frac{\mathrm{d}}{\mathrm{d}y}y\Big|^2 \mu_\sigma(\mathrm{d}y) = \frac{1}{\kappa_1}\cdot$$

For the lower bound: using the inequality $g^2 \geq 2fg - f^2$ for all functions f and g, with $g = y - m_\sigma$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\Psi(\sigma) \geq \int [2f(y-m_\sigma) - f^2]\,\mu_\sigma(\mathrm{d}y).$$

By taking $f = \beta(\psi' - \sigma)$ for $\beta \in \mathbb{R}$, and applying integration by parts, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\Psi(\sigma) \geq 2\beta - \beta^2 \int \psi''(y)\mu_\sigma(\mathrm{d}y).$$

Now maximizing over β, by choosing $\beta = \frac{1}{\int \psi''(y)\mu_\sigma(\mathrm{d}y)}$, we get

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\Psi(\sigma) \geq \frac{1}{\int \psi''(y)\mu_\sigma(\mathrm{d}y)} \geq \frac{1}{\kappa_2},$$

where we have used the assumption that $\psi'' \leq \kappa_2$.

The second estimate in (2.18) is proved similarly. We have

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}\Phi(\sigma) = \int_{\mathbb{R}} (y - m_\sigma^P)^2\,\mathrm{d}\mu_\sigma^P(\mathrm{d}x) - \int_{\mathbb{R}} (y - m_\sigma)^2\,\mathrm{d}\mu_\sigma(\mathrm{d}x), \quad \text{where}$$

$$\mu_\sigma^P = \frac{\exp[-(\psi+P)(y)+\sigma y]}{\int_{\mathbb{R}} \exp[-(\psi+P)(y)+\sigma y)\,\mathrm{d}y}\,\mathrm{d}y, \quad \text{and} \quad m_\sigma^P = \int_{\mathbb{R}} y\mu_\sigma^P(\mathrm{d}y).$$

Since $\psi + P$ satisfies a similar assumption as ψ, we obtain

$$\frac{1}{\varsigma_2} \leq \int_{\mathbb{R}} (y - m_\sigma^P)^2\,\mathrm{d}\mu_\sigma^P(\mathrm{d}x) \leq \frac{1}{\varsigma_1}\cdot$$

As a consequence, we get

$$\frac{1}{\varsigma_2} - \frac{1}{\kappa_1} \leq \frac{\mathrm{d}}{\mathrm{d}\sigma}\Phi(\sigma) \leq \frac{1}{\varsigma_1} - \frac{1}{\kappa_2},$$

which implies the second estimate in (2.18). □

We recall that σ_0 and σ_N^P are respectively maximisers in (2.12) and (2.13). The following lemma provides an estimate for $|\sigma_N^P - \sigma_0|$.

Lemma 2.5. *There exists a positive constant C such that, for N sufficiently large,*

$$|\sigma_P^N - \sigma_0| \leq \frac{C}{N}. \tag{2.20}$$

Proof. Set $F := \Psi + \frac{1}{N}\Phi$. Then we have

$$A = \Psi(\sigma_0) = F(\sigma_P^N), \quad \text{and} \quad F'(\sigma) = \Psi'(\sigma) + \frac{1}{N}\Phi'(\sigma).$$

This, together with Lemma 2.4, imply that for sufficiently large N and for all $\sigma \in \mathbb{R}$

$$\frac{0.5}{\kappa_2} \leq |F'(\sigma)| \leq \frac{2}{\kappa_1}.$$

By the mean value theorem, there exists $\theta \in \mathbb{R}$ such that

$$F'(\theta)(\sigma_P^N - \sigma_0) = F(\sigma_P^N) - F_P(\sigma_0) = F_0(\sigma_0) - \left(F_0(\sigma_0) + \frac{1}{N}\Phi(\sigma_0)\right) = -\frac{1}{N}\Phi(\sigma_0).$$

Hence

$$|\sigma_P^N - \sigma_0^N| = \frac{1}{N}\left|\frac{\Phi(\sigma_0)}{F'(\theta)}\right| \leq \frac{C}{N},$$

for some constant $C > 0$ and for N sufficiently large. □

The following estimate is elementary but will be used at various places later.

Lemma 2.6. *For any $z \in \mathbb{C}$, we have*

$$|\mathrm{e}^z - 1| leq |z| \mathrm{e}^{|z|}. \tag{2.21}$$

Proof. We have

$$|\mathrm{e}^z - 1| = \left|\mathrm{e}^{tz}\Big|_0^1\right| = \left|\int_0^1 z\mathrm{e}^{tz}\,\mathrm{d}t\right| \leq |z| \int_0^1 |\mathrm{e}^{tz}|\,\mathrm{d}t = |z| \int_0^1 \mathrm{e}^{t\mathrm{Rel}(z)}\,\mathrm{d}t \leq |z| \int_0^1 \mathrm{e}^{|z|}\,\mathrm{d}t = |z|\mathrm{e}^{|z|}.$$ □

The second ingredient of the proof of Theorem 2.1 is the following proposition.

Proposition 2.7. *It holds that*

$$\lim_{N\to\infty} N[W_N^P(A) - W(A)] = -\log \frac{\int \exp[-(\psi + P)(y) - W'(A)y]\,\mathrm{d}y}{\int \exp[-\psi(y) - W'(A)y]\,\mathrm{d}y}. \tag{2.22}$$

Moreover, it hods that

$$\left| N[W_N^P(A) - W(A)] + \log \frac{\int \exp[-(\psi + P)(y) - W'(A)y]\,\mathrm{d}y}{\int \exp[-\psi(y) - W'(A)y]\,\mathrm{d}y} \right| \leq \frac{C}{N}.$$

Proof. We recall that σ_0 and σ_P^N are respectively the maximisers in the definitions of $W(A)$ and $W_N^P(A)$, so that

$$W(A) = \sup_{\sigma\in\mathbb{R}} \left\{\sigma A - \log \int_{\mathbb{R}} \exp(-\psi(y) + \sigma y)\,\mathrm{d}y\right\} \tag{2.23}$$

$$= \sigma_0 A - \log \int_{\mathbb{R}} \exp(-\psi(y) + \sigma_0 y)\,\mathrm{d}y, \tag{2.24}$$

where σ_0 satisfies

$$A = \frac{\int_{\mathbb{R}} y \exp(-\psi(y) + \sigma_0 y)\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + \sigma_0 y)\,\mathrm{d}y} = \Psi(\sigma_0). \tag{2.25}$$

By properties of the Legendre transform, we also have $W'(A) = \sigma_0$, which is explicitly shown in (2.53). Similarly

$$W_N^P(A) = \sup_{\sigma\in\mathbb{R}} \left\{ \sigma A - \frac{1}{N}\left(\log \int_{\mathbb{R}} \exp[-(\psi + P)(y) + \sigma y]\,\mathrm{d}y + (N-1)\log \int_{\mathbb{R}} \exp(-\psi(y) + \sigma y)\,\mathrm{d}y \right) \right\} \tag{2.26}$$

$$= \sigma_P^N A - \log \int_{\mathbb{R}} \exp[-\psi(y) + \sigma_P^N y]\,\mathrm{d}y - \frac{1}{N}\log \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_P^N y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + \sigma_P^N y)\,\mathrm{d}y}, \tag{2.27}$$

where σ_P^N solves

$$A = \frac{1}{N}\frac{\int_{\mathbb{R}} y\exp[-(\psi+P)(y) + \sigma y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma y]\,\mathrm{d}y} + \frac{N-1}{N}\frac{\int_{\mathbb{R}} y\exp(-\psi(y) + \sigma y)\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + \sigma y)\,\mathrm{d}y}. \tag{2.28}$$

Using these supremum representations we will estimate lower and upper bounds for $N[W_N^P(A) - W(A)]$. For an upper bound: it follows from (2.23) that

$$W_N(A) \geq \sigma_P^N A - \log \int_{\mathbb{R}} \exp(-\psi(y) + \sigma_P^N y)\,\mathrm{d}y.$$

This, together with (2.27), we get

$$N[W_N^P(A) - W(A)] \leq -\log \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_P^N y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + \sigma_P^N y)\,\mathrm{d}y}.$$

Similarly, using (2.26) and (2.24), we obtain

$$N[W_N^P(A) - W(A)] \geq -\log \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_0 y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + \sigma_0 y)\,\mathrm{d}y}.$$

Bringing these bounds together,

$$-\log \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_0 y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi_1(y) + \sigma_0 y)\,\mathrm{d}y} \leq N[W_N^P(A) - W(A)] \leq -\log \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_P^N y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + \sigma_P^N y)\,\mathrm{d}y}. \tag{2.29}$$

We now estimate the right-hand side of the last expression. We have

$$\begin{aligned}
&\frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_P^N y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + \sigma_P^N y)\,\mathrm{d}y} \\
&= \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_0 y + (\sigma_P^N - \sigma_0)y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_0 y)\,\mathrm{d}y} \times \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_0 y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + \sigma_0 y)\,\mathrm{d}y} \\
&\quad \times \frac{\int_{\mathbb{R}} \exp[-\psi(y) + \sigma_0 y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) + \sigma_P^N y]\,\mathrm{d}y} \\
&= \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + \sigma_0 y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi_1(y) + \sigma_0 y)\,\mathrm{d}y} \times \left\langle \exp[(\sigma_P^N - \sigma_0)y] \right\rangle_{\nu^{\sigma_0}} \times \left\langle \exp[(\sigma_P^N - \sigma_0)y] \right\rangle_{\mu^{\sigma_0}}^{-1},
\end{aligned}$$

where

$$\nu^{\sigma_0}(y)\mathrm{d}y = \frac{\exp[-(\psi+P)(y)+\sigma_0 y]}{\int_{\mathbb{R}}\exp[-(\psi+P)(y)+\sigma_0 y)]\,\mathrm{d}y}\,\mathrm{d}y \quad \text{and} \quad \mu^{\sigma_0}(y)\mathrm{d}y = \frac{\exp[-\psi(y)+\sigma_0 y]}{\int_{\mathbb{R}}\exp[-\psi(y)+\sigma_0 y]\,\mathrm{d}y}\cdot$$

Taking the logarithm of the above equality, we deduce

$$\begin{aligned} &-\log\frac{\int_{\mathbb{R}}\exp[-(\psi+P)(y)+\sigma_P^N y]\,\mathrm{d}y}{\int_{\mathbb{R}}\exp(-\psi(y)+\sigma_P^N y)\,\mathrm{d}y} \\ &= -\log\frac{\int_{\mathbb{R}}\exp[-(\psi+P)(y)+\sigma_0 y]\,\mathrm{d}y}{\int_{\mathbb{R}}\exp(-\psi(y)+\sigma_0 y)\,\mathrm{d}y} + \log\left\langle\exp[(\sigma_P^N-\sigma_0)y]\right\rangle_{\nu^{\sigma_0}} - \log\left\langle\exp[(\sigma_P^N-\sigma_0)y]\right\rangle_{\mu^{\sigma_0}}. \end{aligned} \tag{2.30}$$

We now show that the last two terms in the right-hand side of (2.30) are of order $O(N^{-1})$. Using the estimate $|e^t-1|\le|t|e^{|t|}$ (Lem. 2.6) and Lemma 2.5, we have

$$|\exp[(\sigma_P^N-\sigma_0)y]-1| \le |(\sigma_P^N-\sigma_0)y|\exp(|(\sigma_P^N-\sigma_0)y|) \le \frac{C}{N}|y|\exp(C|y|).$$

Therefore

$$\begin{aligned} \left|\left\langle\exp[(\sigma_P^N-\sigma_0)y]\right\rangle_{\nu^{\sigma_0}}-1\right| = \left|\left\langle\exp[(\sigma_P^N-\sigma_0)y]-1\right\rangle_{\nu^{\sigma_0}}\right| &\le \left\langle|\exp[(\sigma_P^N-\sigma_0)y]-1|\right\rangle_{\nu^{\sigma_0}} \\ &\le \frac{C}{N}\langle|y|\exp(C|y|)\rangle_{\nu^{\sigma_0}}. \end{aligned}$$

Since $(\psi+P)(y)$ is bounded from below and above by a quadratic potential, it implies that the term

$$\langle|y|\exp(C|y|)\rangle_{\nu^{\sigma_0}} = \frac{1}{\int_{\mathbb{R}}\exp[-(\psi+P)(y)+\sigma_0 y]\,\mathrm{d}y}\int|y|\exp[-(\psi+P)(y)+\sigma_0 y+C|y|]\,\mathrm{d}y.$$

is finite. Therefore $\left|\left\langle\exp[(\sigma_P^N-\sigma_0)y]\right\rangle_{\nu^{\sigma_0}}-1\right|\le\frac{C}{N}$, which implies that

$$\left|\log\left\langle\exp[(\sigma_P^N-\sigma_0)y]\right\rangle_{\nu^{\sigma_0}}\right|\le\frac{C}{N}\cdot$$

Similarly, we obtain the following estimate for the last term in (2.30)

$$\left|\log\left\langle\exp[(\sigma_P^N-\sigma_0)y]\right\rangle_{\mu^{\sigma_0}}\right|\le\frac{C}{N}\cdot$$

Substituting these above estimates to (2.30), we achieve the following estimate for the upper bound in (2.29)

$$\left|-\log\frac{\int_{\mathbb{R}}\exp[-(\psi+P)(y)+\sigma_P^N y]\,\mathrm{d}y}{\int_{\mathbb{R}}\exp(-\psi(y)+\sigma_P^N y)\,\mathrm{d}y}+\log\frac{\int_{\mathbb{R}}\exp[-(\psi+P)(y)+\sigma_0 y]\,\mathrm{d}y}{\int_{\mathbb{R}}\exp(-\psi(y)+\sigma_0 y)\,\mathrm{d}y}\right|\le\frac{C}{N}\cdot$$

Therefore, it follows from (2.29) that

$$\left|N[W_N^P(A)-W(A)]+\log\frac{\int\exp[-(\psi+P)(y)-\sigma_0 y]\,\mathrm{d}y}{\int\exp[-\psi(y)-\sigma_0 y]\,\mathrm{d}y}\right|\le\frac{C}{N}\cdot$$

This completes the proof of the proposition. □

Next, we estimate the last term in (2.16). We will need two auxiliary lemmas.

Let $h(m,\xi):=\langle\exp(i\xi(x-m))\rangle_{\mu^\sigma}$, where $\mu^\sigma(x)\,\mathrm{d}x = Z_\sigma^{-1}\exp(\sigma x-\psi(x))\,\mathrm{d}x$.

Lemma 2.8. *For any $\delta > 0$, it holds that*

$$|h(m,\xi)| \leq 1 - \frac{1}{2}\sqrt{C_\sigma}\left(1 - \exp\left(-\frac{\delta^2}{2\kappa_2}\right)\right) \quad \textit{for} \quad |\xi| \geq \delta, \tag{2.31}$$

where $C_\sigma = \exp\left(\frac{\sigma^2}{4}\frac{\kappa_1-\kappa_2}{\kappa_1\kappa_2}\right)\sqrt{\frac{\kappa_1}{\kappa_2}}$

Note that since $0 < \kappa_1 < \kappa_2$, we have $0 < C_\sigma < \sqrt{\frac{\kappa_1}{\kappa_2}} < 1$, which is independent of σ.

Proof. The proof of this lemma is adapted from that of Lemma 39, (i) from [11]. Since $\kappa_1 x^2 \leq \psi(x) \leq \kappa_2 x^2$, we have

$$\mu^\sigma(x) \geq \frac{\exp(\sigma x - \kappa_2 x^2)}{\int_{\mathbb{R}} \exp(\sigma y - \kappa_1 y^2)\,\mathrm{d}y} = \frac{\exp(\sigma x - \kappa_2 x^2)}{\int_{\mathbb{R}} \exp(\sigma y - \kappa_2 y^2)\,\mathrm{d}y}\frac{\int_{\mathbb{R}} \exp(\sigma y - \kappa_2 y^2)\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(\sigma y - \kappa_1 y^2)\,\mathrm{d}y} = n_\sigma(x) C_\sigma,$$

where

$$n_\sigma(x) = \frac{\exp(\sigma x - \kappa_2 x^2)}{\int_{\mathbb{R}} \exp(\sigma y - \kappa_2 y^2)\,\mathrm{d}y}, \quad C_\sigma = \frac{\int_{\mathbb{R}} \exp(\sigma y - \kappa_2 y^2)\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(\sigma y - \kappa_1 y^2)\,\mathrm{d}y} = \exp\left(\frac{\sigma^2}{4}\frac{\kappa_1 - \kappa_2}{\kappa_1\kappa_2}\right)\sqrt{\frac{\kappa_1}{\kappa_2}}.$$

Note that $0 < C_\sigma < 1$ for all σ. The following identity is the same as (157) of [11]

$$|h(m,\xi)|^2 = 1 - \mathrm{Var}(\cos(\xi x)) - \mathrm{Var}(\sin(\xi x)). \tag{2.32}$$

Next we estimate $\mathrm{Var}(\cos(\xi x))$.

$$\begin{aligned}
\mathrm{Var}(\cos(\xi x)) &= \int_{\mathbb{R}} \left(\cos(\xi x) - \int_{\mathbb{R}} \cos(\xi y)\mu_\sigma\,\mathrm{d}y\right)^2 \mu_\sigma\,\mathrm{d}y \\
&\geq C_\sigma \int_{\mathbb{R}} \left(\cos(\xi x) - \int_{\mathbb{R}} \cos(\xi y)\mu_\sigma\,\mathrm{d}y\right)^2 n_\sigma(x) \\
&\geq C_\sigma \left[\int_{\mathbb{R}} \cos(\xi x)^2 n_\sigma(dx) - \left(\int_{\mathbb{R}} \cos(\xi x) n_\sigma(\mathrm{d}x)\right)^2\right].
\end{aligned} \tag{2.33}$$

The second integral on the right-hand side can be computed explicitly as follows:

$$\begin{aligned}
&\left(\int_{\mathbb{R}} \cos(\xi y) n_\sigma(\mathrm{d}y)\right)^2 \\
&= \frac{1}{4}\left(\sqrt{\frac{\kappa_2}{\pi}}\exp\left(-\frac{\sigma^2}{4\kappa_2}\right)\int_{\mathbb{R}} [\exp(i\xi x) + \exp(-i\xi x)]\exp(-\kappa_2 x^2 + \sigma x)\,\mathrm{d}x\right)^2 \\
&= \frac{1}{4}\left(\sqrt{\frac{\kappa_2}{\pi}}\exp\left(\frac{i\sigma\xi}{2\kappa_2}\right)\int_{\mathbb{R}} \exp(i\xi y)\exp(-\kappa_2 y^2)\,\mathrm{d}y + \sqrt{\frac{\kappa_2}{\pi}}\exp\left(-\frac{i\sigma\xi}{2\kappa_2}\right)\int_{\mathbb{R}} \exp(-i\xi y)\exp(-\kappa_2 y^2)\,\mathrm{d}y\right)^2 \\
&= \frac{1}{4}\left(\exp\left(-\frac{\xi^2}{4\kappa_2}\right)\exp\left(\frac{i\sigma\xi}{2\kappa_2}\right) + \exp\left(-\frac{\xi^2}{4\kappa_2}\right)\exp\left(-\frac{i\sigma\xi}{2\kappa_2}\right)\right)^2 \\
&= \frac{1}{4}\exp\left(-\frac{\xi^2}{2\kappa_2}\right)\left(\exp\left(\frac{i\sigma\xi}{\kappa_2}\right) + \exp\left(-\frac{i\sigma\xi}{\kappa_2}\right) + 2\right) \\
&= \frac{1}{2}\exp\left(-\frac{\xi^2}{2\kappa_2}\right)\left(1 + \cos\left(\frac{\sigma\xi}{\kappa_2}\right)\right).
\end{aligned}$$

The first integral can be computed similarly:

$$\int_{\mathbb{R}} \cos^2(\xi x) n_\sigma(\mathrm{d}x) = \frac{1}{2}\left(1 + \cos\left(\frac{\sigma\xi}{\kappa_2}\right)\exp\left(-\frac{\xi^2}{\kappa_2}\right)\right)\cdot$$

Therefore,

$$\begin{aligned} &\int_{\mathbb{R}} \cos(\xi x)^2 n_\sigma(\mathrm{d}x) - \left(\int_{\mathbb{R}} \cos(\xi x) n_\sigma(\mathrm{d}x)\right)^2 \\ &= \frac{1}{2}\left(1 - \exp\left(-\frac{\xi^2}{2\kappa_2}\right)\right)\left(1 - \cos\left(\frac{\sigma\xi}{\kappa_2}\right)\exp\left(-\frac{\xi^2}{2\kappa_2}\right)\right) \\ &\quad \geq \frac{1}{2}\left(1 - \exp\left(\frac{-\xi^2}{2\kappa_2}\right)\right)^2 \cdot \end{aligned}$$

Substituting these computations into (2.33) we obtain

$$\mathrm{Var}(\cos(\xi x)) \geq \frac{1}{2} C_\sigma \left(1 - \exp\left(-\frac{\xi^2}{2\kappa_2}\right)\right)^2 \cdot$$

By repeating the computation, we obtain that the same inequality holds for $\mathrm{Var}(\sin(\xi x))$. Therefore,

$$|h(m,\xi)|^2 \leq 1 - C_\sigma \left(1 - \exp\left(-\frac{\xi^2}{2\kappa_2}\right)\right)^2 \cdot$$

If $|\xi| \geq \delta$, then

$$|h(m,\xi)|^2 \leq 1 - C_\sigma \left(1 - \exp\left(-\frac{\delta^2}{2\kappa_2}\right)\right)^2 \cdot$$

Since $\sqrt{1-x} \leq 1 - \frac{1}{2}x$, it follows that

$$|h(m,\xi)| \leq 1 - \frac{1}{2}\sqrt{C_\sigma}\left(1 - \exp\left(-\frac{\delta^2}{2\kappa_2}\right)\right) \quad \text{for} \quad |\xi| \geq \delta.$$

This concludes the proof. □

Define $\Lambda(\sigma) := \mathrm{Var}(\mu_\sigma) = \int_{\mathbb{R}} \left(x - \int_{\mathbb{R}} x\, \mu_\sigma(\mathrm{d}x)\right)^2 \mu_\sigma(\mathrm{d}x)$.

Lemma 2.9. *There exists $C > 0$ such that, for any $\sigma_1, \sigma_2 \in \mathbb{R}$,*

$$|\Lambda(\sigma_1) - \Lambda(\sigma_2)| \leq C|\sigma_1 - \sigma_2|.$$

Proof. It follows from (2.19) that $\Lambda(\sigma) = \Psi'(\sigma)$. According to [11], (Lem. 41) we have

$$|\Psi''(\sigma)| \leq C,$$

for some constant $C > 0$. As a consequence, we obtain that

$$|\Lambda(\sigma_1) - \Lambda(\sigma_2)| = |\Psi'(\sigma_1) - \Psi'(\sigma_2)| \leq C|\sigma_1 - \sigma_2|\cdot$$

This finishes the proof.

We are now ready to estimate the last term in the right-hand side of (2.16). □

Proposition 2.10. *There exists $C > 0$ such that*

$$\left| \log \frac{g^P_{N,A}(0)}{g_{N,A}(0)} \right| \le \frac{C}{N}. \tag{2.34}$$

Proof. We recall the general setting in Lemma 2.2.

$$\tilde{\mu}_j^{\sigma^*}(\mathrm{d}y) = \exp\left[-\tilde{\varphi}_{N,j}(\sigma^*) + \sigma^* y - \tilde{\psi}_j(y) \right] \mathrm{d}y,$$

where

$$\tilde{\varphi}_j(\sigma) = \log \int_{\mathbb{R}} \exp[-\tilde{\psi}_j(y) + \sigma\, y]\, \mathrm{d}y$$

For each $j = 1, \dots, N$, let $\tilde{m}_j$ and $\tilde{\varsigma}_j^2$ be the mean and variance of $\tilde{\mu}_j^{\sigma^*}$, *i.e.*,

$$\tilde{m}_j = \int_{\mathbb{R}} y \tilde{\mu}_j^{\sigma^*}(\mathrm{d}y) \quad \text{and} \quad \tilde{\varsigma}_j^2 = \int_{\mathbb{R}} (y - \tilde{m}_j)^2\, \tilde{\mu}_j^{\sigma^*}(\mathrm{d}y).$$

Then $\tilde{g}_{N,A}$ has been defined to be the density of $\frac{1}{\sqrt{N}} \sum_{j=1}^N (\tilde{X}_j - m_j)$, where $\tilde{X}_j$ are independent random variables distributed according to $\tilde{\mu}_j^{\sigma^*}$.

Define $\tilde{y}_j = y_j - \tilde{m}_j$. The value of $\tilde{g}_{N,A}$ at 0 can be expressed as (*cf.* *e.g.*, [11], (Eq. (127)), [14], (Eq. (72)))

$$\tilde{g}_{N,A}(0) = \frac{1}{2\pi} \int_{\mathbb{R}} \prod_{j=1}^N \left\langle \exp\left(i\, \tilde{y}_j \frac{1}{\sqrt{N}}\, \xi \right) \right\rangle_j \mathrm{d}\xi,$$

where $\langle \cdot \rangle_j$ denotes the average with respect to $\tilde{\mu}_j^{\sigma^*}$. For some $\delta > 0$ sufficiently small, we split the above integral into two terms

$$\begin{aligned}
\int_{\mathbb{R}} \prod_{j=1}^N \left\langle \exp\left(i\, \tilde{y}_j \frac{1}{\sqrt{N}}\, \xi \right) \right\rangle_j \mathrm{d}\xi &= \int_{\left\{ \left| \frac{1}{\sqrt{N}} \xi \right| \le \delta \right\}} \prod_{j=1}^N \left\langle \exp\left(i\, \tilde{y}_j \frac{1}{\sqrt{N}}\, \xi \right) \right\rangle_j \mathrm{d}\xi \\
&\quad + \int_{\left\{ \left| \frac{1}{\sqrt{N}} \xi \right| \ge \delta \right\}} \prod_{j=1}^N \left\langle \exp\left(i\, \tilde{y}_j \frac{1}{\sqrt{N}}\, \xi \right) \right\rangle_j \mathrm{d}\xi \\
&= \mathrm{I} + \mathrm{II},
\end{aligned}$$

so that

$$\tilde{g}_{N,A}(0) = \frac{1}{2\pi} (\mathrm{I} + \mathrm{II}). \tag{2.35}$$

According to ([14], Proof of Thm. 4), the following estimates hold

$$0 < C_1 \le |\mathrm{I}| \le \mathrm{C}_2, \quad \text{and} \quad |\mathrm{II}| \le \mathrm{C}_3\, \mathrm{N}\, \lambda^{\mathrm{N}-2}, \tag{2.36}$$

for some positive constants C_1, C_2, C_3 and $0 < \lambda < 1$ depending only on δ. The constant λ is the upper bound of $\left| \langle \exp(i \tilde{y}_j \xi) \rangle_j \right|$. Moreover, there exists a complex-valued function $h_j(\xi)$ such that for $0 < |\xi|$ sufficiently small,

$$\langle \exp(i \tilde{y}_j \xi) \rangle_j = \exp(-h_j(\xi)) \quad \text{with} \quad \left| h_j(\xi) - \frac{1}{2} \tilde{\varsigma}_j^2 \xi^2 \right| \le C |\xi|^3. \tag{2.37}$$

We are now ready to prove Proposition 2.10. Applying (2.35), (2.36) and (2.37) for the perfect material, we have

$$g_{N,A}(0) = \frac{1}{2\pi} (\mathrm{I}_1 + \mathrm{II}_1),$$

where

$$\mathrm{I}_1 = \int_{\left\{\left|\frac{1}{\sqrt{\mathrm{N}}}\xi\right|\le\delta\right\}} \exp\left(-\mathrm{Nh}\left(\frac{\xi}{\sqrt{\mathrm{N}}}\right)\right)\mathrm{d}\xi, \tag{2.38}$$

$$0 < C_{11} \le |\mathrm{I}_1| \le \mathrm{C}_{12}, \quad \text{and} \quad |\mathrm{II}_1| \le \mathrm{C}_{13}\mathrm{N}\lambda_1^{\mathrm{N}-2}, \tag{2.39}$$

for some $0 < \lambda_1 < 1$ and positive constants C_{11}, C_{12}, C_{13} and $\left|h(\xi) - \frac{1}{2}\varsigma^2\xi^2\right| \le C|\xi|^3$ for $|\xi| \ll 1$ with ς^2 denotes the variance of μ^{σ_0}. According to Lemma 2.8, the constant λ_1 is given by

$$\lambda_1 = 1 - \frac{1}{2}\sqrt{C_{\sigma_0}}\left(1 - \exp\left(-\frac{\delta^2}{2\kappa_2}\right)\right),$$

with $0 < C_{\sigma_0} < 1$. Similarly,

$$g^P_{N,A}(0) = \frac{1}{2\pi}(\mathrm{I}_2 + \mathrm{II}_2), \tag{2.40}$$

where

$$\mathrm{I}_2 = \int_{\left\{\left|\frac{1}{\sqrt{\mathrm{N}}}\xi\right|\le\delta\right\}} \exp\left(-\sum_{\mathrm{j}=1}^{\mathrm{N}}\tilde{\mathrm{h}}_{\mathrm{j}}\left(\frac{\xi}{\sqrt{\mathrm{N}}}\right)\right)\mathrm{d}\xi, \tag{2.41}$$

$$0 < C_{21} \le |\mathrm{I}_2| \le \mathrm{C}_{22}, \quad \text{and} \quad |\mathrm{II}_2| \le \mathrm{C}_{23}\mathrm{N}\lambda_2^{\mathrm{N}-2}, \tag{2.42}$$

for some $0 < \lambda_2 < 1$ and positive constants C_{21}, C_{22}, C_{23} and

$$\left|\tilde{h}_1(\xi) - \frac{1}{2}\varsigma_{P,1}^2\xi^2\right| \le C|\xi|^3, \quad \text{for } |\xi| \text{ sufficiently small,}$$

$$\tilde{h_2} = \ldots = \tilde{h}_N, \quad \varsigma_{P,2} = \ldots = \varsigma_{P,N}, \left|\tilde{h}_j(\xi) - \frac{1}{2}\varsigma_{P,j}^2\xi^2\right| \le C|\xi|^3, \quad \text{for } |\xi| \text{ sufficiently small,}$$

where $\zeta_{P,1}^2$ and $\zeta_{P,2}^2$ are respectively the variances of $\nu^{\sigma_P^N}$ and $\mu^{\sigma_P^N}$.

The constant λ_2 is given by

$$\lambda_2 = \max\left\{1 - \frac{1}{2}\sqrt{C_{\sigma_P^N}}\left(1 - \exp\left(-\frac{\delta^2}{\kappa_2}\right)\right), 1 - \frac{1}{2}\sqrt{\tilde{C}_{\sigma_P^N}}\left(1 - \exp\left(-\frac{\delta^2}{\kappa_2 + \varsigma_2}\right)\right)\right\},$$

with $0 < C_{\sigma_P^N}, \tilde{C}_{\sigma_P^N} < 1$.

Hence we obtain

$$\frac{g^P_{N,A}(0)}{g_{N,A}(0)} - 1 = \frac{\mathrm{I}_2 + \mathrm{II}_2}{\mathrm{I}_1 + \mathrm{II}_1} - 1 = \frac{\mathrm{I}_2 - \mathrm{I}_1}{\mathrm{I}_1 + \mathrm{II}_1} + \frac{\mathrm{II}_2 - \mathrm{II}_1}{\mathrm{I}_1 + \mathrm{II}_1}. \tag{2.43}$$

It follows from (2.39) that $|I_1 + II_1| \le C$ for N sufficiently large, thus

$$\left|\frac{g^P_{N,A}(0)}{g_{N,A}(0)} - 1\right| \le |\mathrm{I}_2 - \mathrm{I}_1| + |\mathrm{II}_2 - \mathrm{II}_1|. \tag{2.44}$$

The second term decays exponentially fast since, from (2.39) and (2.42)

$$|\mathrm{II}_2 - \mathrm{II}_1| \le |\mathrm{II}_1| + |\mathrm{II}_2| \le \mathrm{CN}\lambda^{\mathrm{N}-2}, \tag{2.45}$$

with $\lambda = \max\{\lambda_1, \lambda_2\}$. It follows that $\lambda = 1 - O(\delta^2)$.

It remains to estimate $|\mathrm{I}_2 - \mathrm{I}_1|$. By changing variable $t := \frac{\xi}{\sqrt{N}}$, we have

$$\begin{aligned}
\mathrm{I}_1 - \mathrm{I}_2 &= \int_{\left\{\left|\frac{1}{\sqrt{N}}\xi\right|\leq\delta\right\}} \left[\exp\left(-Nh\Big(\frac{\xi}{\sqrt{N}}\Big)\right) - \exp\left(-\sum_{j=1}^{N}\tilde{h}_j\Big(\frac{\xi}{\sqrt{N}}\Big)\right)\right] \mathrm{d}\xi \\
&= \sqrt{N}\int_{-\delta}^{\delta}\left[\exp(-Nh(t)) - \exp\left(-\sum_{j=1}^{N}\tilde{h}_j(t)\right)\right]\mathrm{d}t \\
&= \sqrt{N}\int_{-\delta}^{\delta}\exp\Big(-Nh(t)\Big)\left(1-\exp\left(\sum_{j=1}^{N}(h(t)-\tilde{h}_j(t))\right)\right)\mathrm{d}t. \qquad (2.46)
\end{aligned}$$

Note that

$$\left|h(t) - \frac{1}{2N}\zeta^2 t^2\right| \leq C\frac{t^3}{N^{\frac{3}{2}}}, \quad \left|\tilde{h}_1(t) - \frac{1}{2N}\zeta_{P,1}^2 t^2\right| \leq C\frac{t^3}{N^{\frac{3}{2}}},$$

$$\tilde{h}_j(t) = \ldots = \tilde{h}_N(t), \quad \zeta_{P,j} = \zeta_{P,2} \quad \text{for} \quad j = 2,\ldots,N, \quad \text{and} \quad \left|\tilde{h}_j(t) - \frac{1}{2N}\zeta_{P,2}^2 t^2\right| \leq C\frac{t^3}{N^{\frac{3}{2}}},$$

where we recall that $\zeta^2, \zeta_{P,1}^2$ and $\zeta_{P,2}^2$ are, respectively, the variances of μ^{σ_0}, $\nu^{\sigma_P^N}$ and $\mu^{\sigma_P^N}$. It follows that, for $t < 1$,

$$\begin{aligned}
\Big|\exp(-Nh(t))\Big| &= \exp\left(-\frac{1}{2}\zeta^2 t^2\right)\left|\exp\left(-N\left(h(t) - \frac{1}{2N}\zeta^2 t^2\right)\right)\right| \\
&\leq \exp\left(-\frac{1}{2}\zeta^2 t^2\right)\exp\left(\frac{Ct^3}{N^{\frac{1}{2}}}\right) \\
&\leq \exp\left(\frac{Ct^2}{N^{\frac{1}{2}}}\right). \qquad (2.47)
\end{aligned}$$

Now we estimate

$$\begin{aligned}
\left|\sum_{j=1}^{N}(h(t)-\tilde{h}_j(t))\right| &= \left|\sum_{j=1}^{N}\left(h(t) - \frac{1}{2N}\zeta^2t^2 + \frac{1}{2N}\zeta^2t^2 - \frac{1}{2N}\zeta_{P,j}^2t^2 + \frac{1}{2N}\zeta_{P,j}^2t^2 - \tilde{h}_j(t)\right)\right| \\
&\leq \sum_{j=1}^{N}\left[\left|h(t) - \frac{1}{2N}\zeta^2t^2\right| + \left|\frac{1}{2N}\zeta^2t^2 - \frac{1}{2N}\zeta_{P,j}^2t^2\right| + \left|\frac{1}{2N}\zeta_{P,j}^2t^2 - \tilde{h}_j(t)\right|\right] \\
&\leq \frac{Ct^3}{N^{\frac{1}{2}}} + \frac{N-1}{2N}\left|\zeta^2 - \zeta_{P,2}^2\right|t^2 + \frac{1}{2N}\left|\zeta^2 - \zeta_{P,1}^2\right|t^2. \qquad (2.48)
\end{aligned}$$

From Lemma 2.5 and Lemma 2.9, we have

$$\left|\zeta^2 - \zeta_{P,2}^2\right| = \left|\Lambda(\sigma_0) - \Lambda(\sigma_P^N)\right| \leq C|\sigma_0 - \sigma_P^N| \leq \frac{C}{N}, \quad \text{and}$$

$$|\zeta^2 - \zeta_{P,1}^2| = |\Lambda_P(\sigma_P^N) - \Lambda(\sigma_0)| \leq |\Lambda_P(\sigma_P^N) - \Lambda_P(\sigma_0)| + |\Lambda_P(\sigma_0) - \Lambda(\sigma_0)| \leq \frac{C}{N} + C,$$

where $\Lambda_P(\sigma)$ is the variance of the measure $Z^{-1}\int\exp[-(\psi+P)(x)+\sigma x]\,\mathrm{d}x$ and the last inequality is obtained similarly as in Lemma 2.9.

Substituting these estimates into (2.48), we obtain that, for $t < 1$,

$$\Big|\sum_{j=1}^{N}(h(t)-\tilde{h}_j)(t)\Big| \le \frac{Ct^3}{N^{\frac{1}{2}}} + \frac{Ct^2}{N} + \frac{Ct^2}{N^2} \lesssim \frac{Ct^2}{N^{\frac{1}{2}}}.$$

Therefore by using the estimate $|\mathrm{e}^z - 1| \le |z|\mathrm{e}^{|z|}$, we obtain

$$\begin{aligned}\left|1-\exp\left(\sum_{j=1}^{N}(h(t)-\tilde{h}_j(t))\right)\right| &\le \Big|\sum_{j=1}^{N}(h(t)-\tilde{h}_j(t))\Big| \exp\left(\Big|\sum_{j=1}^{N}(h(t)-\tilde{h}_j(t))\Big|\right)\\ &\le \frac{Ct^2}{\sqrt{N}}\exp\left(\frac{Ct^2}{\sqrt{N}}\right). \qquad (2.49)\end{aligned}$$

Substituting the estimates (2.47)–(2.49) into (2.46), we obtain

$$\begin{aligned}|\mathrm{I}_1-\mathrm{I}_2| &\le \sqrt{N}\int_{-\delta}^{\delta}\exp\left(\frac{Ct^2}{N^{\frac{1}{2}}}\right)\frac{Ct^2}{\sqrt{N}}\exp\left(\frac{Ct^2}{\sqrt{N}}\right)\,\mathrm{d}t\\ &\le C\exp\left(\frac{C\delta^2}{N^{\frac{1}{2}}}\right)\int_{-\delta}^{\delta}t^2\,\mathrm{d}t = O(\delta^3).\end{aligned}$$

By choosing $\delta = N^{-\alpha}$ where $\frac{1}{3} < \alpha < \frac{1}{2}$ then

$$\begin{aligned}&|II_2 - II_1| \lesssim N\lambda^N \lesssim N(1-N^{-2\alpha})^N \lesssim N\left(\mathrm{e}^{-N^{-2\alpha}}\right)^N = N\mathrm{e}^{-N^{-2\alpha+1}} \lesssim N^{-1},\\ &|I_1 - I_2| \lesssim N^{-3\alpha} \lesssim N^{-1}.\end{aligned}$$

Substituting these estimates into (2.44), we obtain

$$\left|\frac{g^P_{N,A}(0)}{g_{N,A}(0)} - 1\right| \lesssim N^{-1},$$

implying that

$$\left|\log\frac{g^P_{N,A}(0)}{g_{N,A}(0)}\right| \lesssim N^{-1}.$$

This completes the proof of the proposition. □

2.2. Coarse-grained energy

In this section, we prove Theorem 1.4 for the case without external forces by deriving the formula for the coarse-grained energy and the representation of the thermodynamic limit $G_\infty(A)$.

We recall that the finite coarse-grained energy E_N^{cg} is defined as a minimization problem

$$E_N^{\mathrm{cg}}(A,y) := \inf_{\substack{u\in\mathbb{R}^N\\ u_1=y,u_N=NA}} \sum_{i=2}^{N}[W(u_i-u_{i-1})-W(A)]. \qquad (2.50)$$

Due to the separation of variables, which is a special property of the 1D model, the minimization is explicit (see Thm. 2.11 below and Thm. 3.1 for the case with external forces). This simplicity explains why the Cauchy-Born derivation from a continuum model leads to the correct coarse-grained energy in Theorem 1.4.

The main theorem of this section is the following.

Theorem 2.11.
(i) *The coarse-grained energy,* $E^{\mathrm{cg}}(A, y) = \lim_{N\to\infty} E^{\mathrm{cg}}_N(A, y)$*, exists and is given by*

$$E^{\mathrm{cg}}(A, y) = W'(A)(A - y). \tag{2.51}$$

In addition, for all $A, y \in \mathbb{R}$ *we have* $|E^{\mathrm{cg}}_N(A, y) - E^{\mathrm{cg}}(A, y)| \lesssim N^{-1}$.
(ii) *The defect formation free energy* $G_\infty(A)$ *can be represented in terms of* E^{cg} *as*

$$G_\infty(A) = -\log \frac{\int_{\mathbb{R}} \exp(-P(y) - \psi(y) - E^{\mathrm{cg}}(A, y))\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) - E^{\mathrm{cg}}(A, y))\,\mathrm{d}y}\cdot \tag{2.52}$$

Proof. We first prove (2.51). The minimizer of the minimization problem (2.50) satisfies the following Euler-Lagrange equation

$$-W'(u_{i+1} - u_i) + W'(u_i - u_{i-1}) = 0,$$

which implies that $W'(u_i - u_{i-1}) = \lambda$, *i.e.*, $u_N - u_{N-1} = \ldots = u_2 - u_1 (= (W')^{-1}(\lambda))$. This implies that

$$u_i - u_{i-1} = \frac{1}{N-1}\sum_{j=2}^{N}(u_j - u_{j-1}) = \frac{NA - y}{N-1} = A + \frac{A-y}{N-1}\cdot$$

Thus, we obtain

$$E^{\mathrm{cg}}_N(A, y) = (N-1)\left[W\Big(A + \frac{A-y}{N-1}\Big) - W(A)\right].$$

By applying the mean value theorem twice, there exist $0 \le \theta, \theta' \le 1$ such that

$$\begin{aligned}
E^{\mathrm{cg}}_N(A, y) - \mathrm{e}^{\mathrm{cg}}(A, y) &= (N-1)\left[W\left(A + \frac{A-y}{N-1}\right) - W(A)\right] - W'(A)(A-y)\\
&= (N-1)W'\left(A + \theta\frac{A-y}{N-1}\right)\frac{A-y}{N-1} - W'(A)(A-y)\\
&= \left[W'\left(A + \theta\frac{A-y}{N-1}\right) - W'(A)\right](A-y)\\
&= W''\left(A + \theta'\frac{A-y}{N-1}\right)\frac{(A-y)^2}{N-1}.
\end{aligned}$$

Let $x \in \mathbb{R}$ and let σ_x be the maximiser in the definition of $W(x)$. Then we have

$$x = \Psi(\sigma_x) \quad \text{and} \quad W(x) = \sigma_x x - \log\int_{\mathbb{R}} \exp[-\psi(y) + \sigma_x y]\,\mathrm{d}y.$$

It follows that

$$W'(x) = x\frac{\mathrm{d}\sigma_x}{\mathrm{d}x} + \sigma_x - \Psi(\sigma_x)\frac{\mathrm{d}\sigma_x}{\mathrm{d}x} = \sigma_x \quad \text{and} \quad W''(x) = \frac{\mathrm{d}\sigma_x}{\mathrm{d}x} = \frac{1}{\Psi'(\sigma_x)}\cdot \tag{2.53}$$

According to Lemma 2.4, we have

$$|W''(x)| \le C$$

for all $x \in \mathbb{R}$. It implies that $|W''(A + \theta'\frac{A-y}{N-1})| \le C$ and hence,

$$|E^{\mathrm{cg}}_N(A, y) - E^{\mathrm{cg}}(A, y)| \le \frac{C(A-y)^2}{N-1},$$

which gives (2.51).

The representation (2.52) is a direct consequence of (2.1) and (2.51). Indeed,

$$G_\infty(A) \overset{(2.1)}{=} -\log \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) + W'(A)y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) + W'(A)y]\,\mathrm{d}y}$$

$$= -\log \frac{\int_{\mathbb{R}} \exp[-(\psi+P)(y) - W'(A)(A-y)]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) - W'(A)(A-y)]\,\mathrm{d}y}$$

$$\overset{(2.51)}{=} -\log \frac{\int_{\mathbb{R}} \exp(-P(y) - \psi(y) - E^{\mathrm{cg}}(A,y))\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) - E^{\mathrm{cg}}(A,y))\,\mathrm{d}y}. \qquad \square$$

2.3. Propagation of error

In this section, we prove Theorem 1.5 for the case without external forces.

Proof of Theorem 1.5 for the case without external forces.

For shortening of the notation, we define $\tilde{\psi} := \psi + P$. We rewrite $G_N^{\mathrm{cg}}(A)$ as follows.

$$\begin{aligned}
G_N^{\mathrm{cg}}(A) &= -\log \frac{\int \exp[-\tilde{\psi}(y) - E_N^{\mathrm{cg}}(A,y)]\,\mathrm{d}y}{\int \exp[-\psi(y) - E_N^{\mathrm{cg}}(A,y)]\,\mathrm{d}y} \\
&= -\log \frac{\int \exp[-\tilde{\psi}(y) - E^{\mathrm{cg}}(A,y)]\,\mathrm{d}y}{\int \exp[-\psi(y) - E^{\mathrm{cg}}(A,y)]\,\mathrm{d}y} - \log \frac{\int \exp[-\tilde{\psi}(y) - E^{\mathrm{cg}}(A,y) - (E_N^{\mathrm{cg}}(A,y) - E^{\mathrm{cg}}(A,y))]\,\mathrm{d}y}{\int \exp\left[-\tilde{\psi}(y) - E^{\mathrm{cg}}(A,y)\right]\,\mathrm{d}y} \\
&\quad + \log \frac{\int \exp[-\psi(y) - E^{\mathrm{cg}}(A,y) - (E_N^{\mathrm{cg}}(A,y) - E^{\mathrm{cg}}(A,y))]\,\mathrm{d}y}{\int \exp[-\psi(y) - E^{\mathrm{cg}}(A,y)]\,\mathrm{d}y} \\
&= -\log \frac{\int \exp[-\tilde{\psi}(y) - E^{\mathrm{cg}}(A,y)]\,\mathrm{d}y}{\int \exp[-\psi(y) - E^{\mathrm{cg}}(A,y)]\,\mathrm{d}y} - \log \left\langle \exp[E_N^{\mathrm{cg}}(A,y) - E^{\mathrm{cg}}(A,y)] \right\rangle_{\zeta_1} \\
&\quad + \log \left\langle \exp[E_N^{\mathrm{cg}}(A,y) - E^{\mathrm{cg}}(A,y)] \right\rangle_{\zeta_2},
\end{aligned}$$

where ζ_1 and ζ_2 are two probability measures defined by

$$\zeta_1(y)\,\mathrm{d}y = \frac{\exp[-\tilde{\psi}(y) - E^{\mathrm{cg}}(y)]\,\mathrm{d}y}{\int \exp[-\tilde{\psi}(y) - E^{\mathrm{cg}}(y)]\,\mathrm{d}y} \quad \text{and} \quad \zeta_2(y)\,\mathrm{d}y = \frac{\exp[-\psi(y) - E^{\mathrm{cg}}(y)]\,\mathrm{d}y}{\int \exp[-\psi(y) - E^{\mathrm{cg}}(y)]\,\mathrm{d}y}.$$

We next show that the logarithmic terms are of order $O(N^{-1})$. The argument will be similar to the paragraph following (2.30) in the proof of Proposition 2.7. Applying the estimate $|e^t - 1| \le |t|e^{|t|}$ and using the estimate in Theorem 2.11, we get

$$\begin{aligned}
|\exp[E_N^{\mathrm{cg}}(A,y) - E^{\mathrm{cg}}(A,y)] - 1| &\le |E_N^{\mathrm{cg}}(A,y) - E^{\mathrm{cg}}(A,y)| \exp[|E_N^{\mathrm{cg}}(A,y) - E^{\mathrm{cg}}(A,y)|] \\
&\le \frac{C}{N}(A-y)^2 \exp\left(\frac{C}{N}(A-y)^2\right) \\
&\le \frac{C}{N}(A-y)^2 \exp\left(\frac{\kappa_1 + \varsigma_1}{2}(A-y)^2\right), \quad \text{for } N \text{ sufficiently large.}
\end{aligned}$$

Therefore,

$$\begin{aligned}\Big|\Big\langle \exp[E_N^{\rm cg}(A,y)-E^{\rm cg}(A,y)]\Big\rangle_{\zeta_1}-1\Big| &= \Big|\Big\langle \exp[E_N^{\rm cg}(A,y)-E^{\rm cg}(A,y)]-1\Big\rangle_{\zeta_1}\Big| \\ &\le \Big\langle \Big|\exp[E_N^{\rm cg}(A,y)-E^{\rm cg}(A,y)]-1\Big|\Big\rangle_{\zeta_1} \\ &\le \frac{C}{N}\Big\langle (A-y)^2\exp\Big(\frac{\kappa_1+\varsigma_1}{2}(A-y)^2\Big)\Big\rangle_{\zeta_1}.\end{aligned}$$

Thanks to Assumption 1.1, the last average term will be finite. Therefore,

$$\Big|\Big\langle \exp[E_N^{\rm cg}(A,y)-E^{\rm cg}(A,y)]\Big\rangle_{\zeta_1}-1\Big| \le \frac{C}{N},$$

which implies that

$$\Big|\log\Big\langle \exp[E_N^{\rm cg}(A,y)-E^{\rm cg}(A,y)]\Big\rangle_{\zeta_1}\Big| \le \frac{C}{N}.$$

Similarly we also have

$$\Big|\log\Big\langle \exp[E_N^{\rm cg}(A,y)-E^{\rm cg}(A,y)]\Big\rangle_{\zeta_2}\Big| \le \frac{C}{N}.$$

Therefore, we obtain that

$$\left|-\log\frac{\int \exp[-\tilde\psi(y)-E_N^{\rm cg}(y)]\,{\rm d}y}{\int \exp[-\psi(y)-E_N^{\rm cg}(y)]\,{\rm d}y}+\log\frac{\int \exp[-\tilde\psi(y)-E^{\rm cg}(y)]\,{\rm d}y}{\int \exp[-\psi(y)-E^{\rm cg}(y)]\,{\rm d}y}\right| \le \frac{C}{N}.$$

This completes the proof. □

3. External forces case

In this section, we consider the case where the external forces are present. Recall that in this case, the perfect free energy is unchanged as the external forces are used to model the decay rate of the defect away from the core. The perfect material energy is given by

$$\begin{cases} F_N(A) = -\beta^{-1}\log\int_{\mathbb{R}^{N-1}}\exp\Big[-\beta\sum_{i=1}^N\psi(u_i-u_{i-1})\Big]\,{\rm d}u_1\ldots{\rm d}u_{N-1} \\ u_0=0, u_N=NA. \end{cases} \tag{3.1}$$

The deformed free energy is influenced by the external forces

$$\begin{cases} F_N^P(A) = -\beta^{-1}\log\int_{\mathbb{R}^{N-1}}\exp\Big[-\beta\sum_{i=1}^N\psi_i(u_i-u_{i-1})-\beta P(u_1)\Big]\,{\rm d}u_1\ldots{\rm d}u_{N-1} \\ u_0=0, u_N=NA, \end{cases} \tag{3.2}$$

where $\psi_i(y)=\psi(y)+h_iy$. The defect-formation free energy is defined as the free energy difference, $G_\infty(A) := \lim_{N\to\infty}G_N(A)$, where

$$G_N(A) = F_N^P(A) - F_N(A). \tag{3.3}$$

Finally, the finite-domain coarse-grained energy is given by

$$E_N^{\rm cg}(A,y) := \inf_{\substack{u\in\mathbb{R}^N \\ u_1=y, u_N=NA}}\sum_{i=2}^N\Big[W(u_i-u_{i-1})-W(A)+h_i(u_i-u_{i-1})\Big]. \tag{3.4}$$

Recall also that the external forces $\{h_i\}_{i=1}^n$ satisfy Assumption 1.2 and $H=\sum_{i=2}^\infty h_i$.

3.1. Coarse-grained energy

We now establish the formula for the coarse-grained energy, thus proving Theorem 1.4 for the case with external forces.

Theorem 3.1. *The coarse-grained energy,* $E^{\mathrm{cg}}(A,y) := \lim_{N\to\infty} E_N^{\mathrm{cg}}(A,y)$, *is given by*

$$E^{\mathrm{cg}}(A,y) = (A-y)W'(A) + AH + \inf_{\substack{v\in\mathbb{R}^{\mathbb{N}}\\ v_1=0}} J_\infty(A;v), \tag{3.5}$$

where

$$J_\infty(A;v) = \sum_{i=2}^{\infty}[W(A+v_i') - W(A) - W'(A)v_i' + h_i v_i']. \tag{3.6}$$

In addition, for all $A, y \in \mathbb{R}$, *we have the estimate*

$$|E_N^{\mathrm{cg}}(A,y) - E^{\mathrm{cg}}(A,y)| \lesssim N^{-1} + A\,\Big|\sum_{i=N+1}^{\infty} h_i\Big| + \sum_{i=N+1}^{\infty}|h_i|^2. \tag{3.7}$$

Proof. By changing variables $v_i' = u_i' - A$ and substituting to (3.4), we obtain

$$E_N^{\mathrm{cg}}(A,y) = \inf_{\substack{v\in\mathbb{R}^N\\ v_1=y-A, v_N=0}} I_N(A;v), \tag{3.8}$$

where

$$\begin{aligned} I_N(A;v) &= \sum_{i=2}^{N}[W(A+v_i') - W(A) + h_i(v_i'+A)] \\ &= \sum_{i=2}^{N}[W(A+v_i') - W(A) - W'(A)v_i' + h_i v_i'] + A\sum_{i=2}^{N} h_i + (A-y)W'(A) \\ &= J_N(A;v) + A\sum_{i=2}^{N} h_i + (A-y)W'(A), \end{aligned}$$

with

$$J_N(A;v) = \sum_{i=2}^{N}[W(A+v_i') - W(A) - W'(A)v_i' + h_i v_i']. \tag{3.9}$$

Therefore

$$E_N^{\mathrm{cg}}(A,y) = A\sum_{i=2}^{N} h_i + (A-y)W'(A) + \inf_{\substack{v\in\mathbb{R}^N\\ v_1=y-A, v_N=0}} J_N(A;v). \tag{3.10}$$

We now show that

$$\lim_{N\to\infty} \inf_{\substack{v\in\mathbb{R}^N\\ v_1=y-A, v_N=0}} J_N(A;v) = \inf_{v_1=y-A} J_\infty(A;v), \tag{3.11}$$

where

$$J_\infty(A;v) = \sum_{i=2}^{\infty}[W(A+v_i') - W(A) - W'(A)v_i' + h_i v_i'].$$

In fact, since $J_\infty(A;v)$ depends only on v_i', we have that

$$\inf_{\substack{v\in\mathbb{R}^{\mathbb{N}}\\ v_1=y-A}} J_\infty(A;v) = \inf_{\substack{v\in\mathbb{R}^{\mathbb{N}}\\ v_1=0}} J_\infty(A;v).$$

To shorten the notation, we define $\Theta_i(A,z) = W(A+z) - W(A) - W'(A)z + h_i z$ so that

$$J_N(A;v) = \sum_{i=2}^{N} \Theta(A, v_i'), \quad \text{and} \quad J_\infty(A;v) = \sum_{i=2}^{\infty} \Theta_i(A, v_i').$$

A minimizer of $J_\infty(A;)$ satisfies the following Euler-Langrange equation for $i = 2,\ldots,N$

$$\Theta_i'(A; v_i') = 0,$$

together with the boundary condition $v_1 = y - A$. In particular, since $\Theta_i'(A,z) = W'(A+z) - W'(A) + h_i = W''(A+\theta z)z + h_i$ for some $\theta \in \mathbb{R}$, it follows that

$$|v_i'| = \frac{|h_i|}{|W''(A+\theta v_i')|} \leq \frac{|h_i|}{\kappa_1}.$$

We define an admissible sequence $\tilde{v}_i$ as follows

$$\tilde{v}_1 = y - A, \quad \tilde{v}_N = 0, \quad \tilde{v}_i' = v_i' + C_N,$$

for some C_N. Since $\{v_i'\} \in l^1$, we have $\sum_{i=2}^N v_i' \to a$ for some $a \in \mathbb{R}$. By summing up the above equalities, it follows that

$$|C_N| \lesssim \frac{|y-A| + |a|}{N}.$$

Since v_i' minimizes Θ_i we have

$$0 \leq \Theta_i(\tilde{v}_i') - \Theta_i(v_i) \lesssim C_N^2 \lesssim N^{-2}.$$

As a consequence, we obtain

$$\begin{aligned}\inf_{\substack{w\in\mathbb{R}^N\\ w_1=y-A, w_N=0}} J_N(A;w) &\leq J_N(A;\tilde{v}) = J_N(A;v) + \sum_{i=2}^{N} [\Theta_i(\tilde{v}_i') - \Theta_i(v_i')] \\ &\leq J_\infty(A;v) + CN^{-1} + \sum_{i=N+1}^{\infty} \Theta_i(v_i') \\ &\leq J_\infty(A;v) + CN^{-1} + C\sum_{i=N+1}^{\infty} |h_i|^2. \end{aligned} \tag{3.12}$$

Note that in the estimation above we have used the fact that $|\Theta_i(v_i')| \leq C(|h_i|^2 + |v_i'|^2)| \leq C|h_i|^2$.

On the other hand, using again the fact that v_i' minimizes Θ_i for each $i = 2,\ldots,N$, we have

$$\inf_{\substack{w\in\mathbb{R}^N\\ w_1=y-A, w_N=0}} J_N(A;w) \geq J_N(A;v) = J_\infty(A;v) - \sum_{i=N+1}^{\infty} \Theta(v_i') \geq J_\infty(A;v) - C\sum_{i=N}^{\infty} |h_i|^2. \tag{3.13}$$

From (3.12) and (3.13), we obtain

$$\left| \inf_{\substack{v\in\mathbb{R}^N\\ v_1=y-A, v_N=0}} J_N(A;v) - \inf_{\substack{v\in\mathbb{R}^N\\ v_1=y-A}} J_\infty(A;v) \right| \lesssim N^{-1} + \sum_{i=N+1}^{\infty} |h_i|^2, \tag{3.14}$$

from which (3.11) follows. Finally, from (3.10) and (3.14), we get

$$|E_N^{\mathrm{cg}}(A,y) - \lim_{N\to\infty} E_N^{\mathrm{cg}}(A,y)| \lesssim N^{-1} + A\left|\sum_{i=N+1}^{\infty} h_i\right| + \sum_{i=N+1}^{\infty} |h_i|^2,$$

which is (3.7) (and hence (3.5)) as claimed. This finishes the proof of Theorem 3.1. □

3.2. Thermodynamic limit

The main result of this section is the following theorem on the representation of the defect formation free energy.

Theorem 3.2. *The thermodynamic limit is given by*

$$G_\infty(A) = -\log \frac{\int_{\mathbb{R}} \exp[-(\psi_1 + P)(y) - E^{\mathrm{cg}}(A,y)]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) - E^{\mathrm{cg}}_{\mathbf{h}=0}(A,y)]\,\mathrm{d}y}\cdot \tag{3.15}$$

where $E^{\mathrm{cg}}(A,y)$ is defined in (3.5).

Proof of Theorem 3.2. The proof is analogous to that of Theorem 2.1 which consists of three main steps.

Step (1). Express the defect-formation free energy in terms of the energy difference and a ratio of the densities of random variables based on Lemma 2.2.
Step (2). Establish the limit of the energy difference.
Step (3). Show that the ratio of the densities of random variables are of order $O(1/N)$.

We now only sketch out the main computations in Step (1) and Step (2). Applying Lemma 2.2 for the case $\tilde{\psi}_1 = \psi_1 + P, \tilde{\psi}_2 = \psi_i$, for $i = 2, \ldots, N$ to obtain

$$\begin{aligned} W_N^P(A) &= \sup_{\sigma\in A}\left\{\sigma A - \frac{1}{N}\log\int_{\mathbb{R}} \exp[-(\psi_1(y) + P(y) + \sigma y]\,\mathrm{d}y - \frac{1}{N}\log\int_{\mathbb{R}}\sum_{i=2}^{N}\exp[-\psi_i(y) + \sigma y]\,\mathrm{d}y\right\}, \\ &= \sigma_N A - \frac{1}{N}\log\int_{\mathbb{R}} \exp[-(\psi_1(y) + P(y) + \sigma_N y]\,\mathrm{d}y - \frac{1}{N}\log\int_{\mathbb{R}}\sum_{i=2}^{N}\exp[-\psi_i(y) + \sigma_N y]\,\mathrm{d}y. \end{aligned}$$

The optimal value σ_N solves

$$\begin{aligned} A &= \frac{1}{N}\frac{\int_{\mathbb{R}} y\exp[-(\psi_1(y) + P(y)) + \sigma y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-(\psi_1(y) + P(y)) + \sigma y]\,\mathrm{d}y} + \frac{1}{N}\sum_{i=2}^{N}\frac{\int_{\mathbb{R}} y\exp[-\psi_i(y) + \sigma y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi_i(y) + \sigma y]\,\mathrm{d}y} \\ &= \frac{1}{N}\Psi_P(\sigma - h_1) + \frac{1}{N}\sum_{i=2}^{N}\Psi(\sigma - h_i), \end{aligned} \tag{3.16}$$

where Ψ is defined in (2.17) and Ψ_P is given by

$$\Psi_P(\sigma) = \frac{\int_{\mathbb{R}} y\exp[-(\psi_1(y) + P(y)) + \sigma y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-(\psi_1(y) + P(y)) + \sigma y]\,\mathrm{d}y}\cdot \tag{3.17}$$

Since $W(A)$ is unchanged, it is the same as in (2.23)–(2.24), so that

$$\begin{aligned} N[W_N^P(A) - W(A)] &= N(\sigma_N - \sigma_0)A - \log \frac{\int_{\mathbb{R}} \exp[-(\psi_1(y) + P(y)) + \sigma_N y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) + \sigma_0 y]\,\mathrm{d}y} \\ &\quad - \sum_{i=2}^N \log \frac{\int_{\mathbb{R}} \exp[-\psi_i(y) + \sigma_N y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) + \sigma_0 y]\,\mathrm{d}y} \\ &= N(\sigma_N - \sigma_0)A - \log \frac{\int_{\mathbb{R}} \exp[-(\psi_1(y) + P(y)) + \sigma_N y]\,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) + \sigma_0 y]\,\mathrm{d}y} \\ &\quad - \sum_{i=2}^N [W^*(-h_i + \sigma_N) - W^*(\sigma_0)], \end{aligned} \tag{3.18}$$

where $\sigma_0 = W'(A)$. We will need the following lemma whose proof is postponed after the proof of Theorem 3.2.

Lemma 3.3. *It holds that*

$$|\sigma_N - \sigma_0| \le \frac{C}{N}. \tag{3.19}$$

To proceed, we will compare this free energy difference with the finite-domain coarse-grained energy. Recalling that the latter is defined by (see (3.4)),

$$\begin{aligned} E_N^{\mathrm{cg}}(y) &:= \inf_{\substack{u:\{1,N\}\to\mathbb{R} \\ u(1)=y,u(N)=NA}} \sum_{i=2}^N \Big[W(u_i - u_{i-1}) - W(A) + h_i(u_i - u_{i-1})\Big] \\ &= A\sum_{i=2}^N h_i + \inf_{\substack{u:\{1,N\}\to\mathbb{R} \\ u(1)=y,u(N)=NA}} \sum_{i=2}^N \Big[W(u_i - u_{i-1}) + h_i(u_i - u_{i-1}) - (h_i A + W(A))\Big]. \end{aligned} \tag{3.20}$$

The Euler-Lagrange equation for a minimizer of E_N^{cg} is

$$-W'(u_{i+1} - u_i) + W'(u_i - u_{i-1}) - (h_{i+1} - h_i) = 0,$$

which implies that

$$W'(u_i - u_{i-1}) = -h_i + \lambda$$

for $i = 2, \ldots, N$ and for some $\lambda \in \mathbb{R}$. We note that $(W')^{-1}(z) = (W^*)'(z)$, where W^* is the Legendre transformation of W. It follows from the definition of W that

$$W^*(x) = \log \int \exp[-\psi(z) + xz]\,\mathrm{d}z,$$

and so

$$(W^*)'(x) = \frac{\int x \exp[-\psi(z) + xz]\,\mathrm{d}z}{\int \exp[-\psi(z) + xz]\,\mathrm{d}z} = \Psi(x).$$

Therefore, we obtain that

$$u_i - u_{i-1} = (W')^{-1}(-h_i + \lambda) = (W^*)'(-h_i + \lambda) = \Psi(-h_i + \lambda).$$

Summing up these equalities from $i = 2$ to N and using the boundary condition on u, we obtain the following equation for $\lambda = \lambda_N$

$$NA - y = \sum_{i=2}^N \Psi(-h_i + \lambda_N). \tag{3.21}$$

Next, we use the following relations of the Legendre transform

$$W(x) = W'(x)x - W^*(W'(x)), \qquad W'((W^*)'(x)) = x$$

to obtain $W(A) = W'(A)A - W^*(W'(A))$ and

$$\begin{aligned} W(u_i - u_{i-1}) &= W((W^*)'(-h_i + \lambda_N)) \\ &= W'((W^*)'(-h_i + \lambda_N))(W^*)'(-h_i + \lambda_N) - W^*(W'((W^*)'(-h_i + \lambda_N))) \\ &= (-h_i + \lambda_N)(W^*)'(-h_i + \lambda_N) - W^*(-h_i + \lambda_N). \end{aligned}$$

Therefore, the sum inside the inf in (3.20) can be re-written as (recalling that $u_N = NA, u_1 = y$)

$$\begin{aligned} &\sum_{i=2}^{N} \Big[W(u_i - u_{i-1}) + h_i(u_i - u_{i-1}) - h_i A - W(A) \Big] \\ &\quad = \sum_{i=2}^{N} \Big[(-h_i + \lambda_N)(W^*)'(-h_i + \lambda_N) - W^*(-h_i + \lambda_N) + h_i (W^*)'(-h_i + \lambda_N) \\ &\quad - h_i A - W'(A)A + W^*(W'(A)) \Big] \\ &\quad = \lambda_N \sum_{i=2}^{N} (W^*)'(-h_i + \lambda_N) - \sum_{i=2}^{N} \Big[W^*(-h_i + \lambda_N) - W^*(W'(A)) + h_i A + W'(A)A \Big] \\ &\quad = \lambda_N \sum_{i=2}^{N} (u_i - u_{i-1}) - \sum_{i=2}^{N} \Big[W^*(-h_i + \lambda_N) - W^*(W'(A)) + h_i A + W'(A)A \Big] \\ &\quad = \lambda_N (NA - y) - A \sum_{i=1}^{N} h_i - (N-1) W'(A)A - \sum_{i=2}^{N} \Big[W^*(-h_i + \lambda_N) - W^*(W'(A)) \Big]. \end{aligned}$$

Substituting this expression back into (3.20), we get

$$\begin{aligned} E_N^{\mathrm{cg}}(y) &= \lambda_N (NA - y) - (N-1) W'(A)A - \sum_{i=2}^{N} [W^*(-h_i + \lambda_N) - W^*(W'(A))] \\ &= \lambda_N (A - y) + (N-1)(\lambda_N - W'(A))A - \sum_{i=2}^{N} [W^*(-h_i + \lambda_N) - W^*(W'(A))] \cdot \end{aligned} \tag{3.22}$$

It follows from (3.18) and (3.22) that

$$\begin{aligned} N[W_N(A) - W(A)] - E_N^{\mathrm{cg}}(A) &= (\sigma_N - \sigma_0)A - \log \frac{\int_{\mathbb{R}} \exp[-(\psi_1(y) + P(y)) + \sigma_N y] \,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) + \sigma_0 y] \,\mathrm{d}y} \\ &\quad + \sum_{i=2}^{N} \left([\sigma_N A - W^*(-h_i + \sigma_N)] - [\lambda_N A - W^*(-h_i + \lambda_N)] \right) \\ &= (\sigma_N - \sigma_0)A - \log \frac{\int_{\mathbb{R}} \exp[-(\psi_1(y) + P(y)) + \sigma_N y] \,\mathrm{d}y}{\int_{\mathbb{R}} \exp[-\psi(y) + \sigma_0 y] \,\mathrm{d}y} \\ &\quad + b_N(\sigma_N) - b_N(\lambda_N), \end{aligned} \tag{3.23}$$

where

$$b_N(x) := \sum_{i=2}^{N}[xA - W^*(-h_i + x)].$$

Then we have

$$b_N'(x) = (N-1)A - \sum_{i=2}^{N}(W^*)'(-h_i + x),$$

$$b_N''(x) = -\sum_{i=2}^{N}(W^*)''(-h_i + x) = -\sum_{i=2}^{N}\Psi'(-h_i + x) \le 0,$$

where we have used (2.19) to obtain the last inequality. Therefore $b_N'(x)$ is a non-increasing function. Furthermore, from (3.21) and (3.16), we have

$$b_N'(\lambda_N) = (N-1)A - \sum_{i=2}^{N}(W^*)'(-h_i + \lambda_N) = y - A,$$

$$b_N'(\sigma_N) = (N-1)A - \sum_{i=2}^{N}(W^*)'(-h_i + \sigma_N) = \Psi_P(\sigma_N - h_1) - A.$$

Since $\frac{\mathrm{d}}{\mathrm{d}\sigma}\Psi_P(\sigma) \le \frac{1}{\kappa_1+\varsigma_1}$, we have

$$\begin{aligned}|\Psi_P(\sigma_N - h_1)| &\le |\Psi_P(0)| + \frac{1}{\kappa_1+\varsigma_1}|\sigma_N - h_1| \le |\Psi_P(0)| + \frac{1}{\kappa_1+\varsigma_1}(|\sigma_0 - h_1| + |\sigma_N - \sigma_0|)\\ &\le \left(|\Psi_P(0)| + \frac{1}{\kappa_1+\varsigma_1}(|\sigma_0 - h_1| + C)\right).\end{aligned}$$

Therefore both $b_N'(\lambda_N)$ and $b_N'(\sigma_N)$ are uniformly bounded. It follows that

$$\begin{aligned}|b_N(\sigma_N) - b_N(\lambda_N)| &= |\sigma_N - \lambda_N||b_N'(\theta_N)|\\ &\le |\sigma_N - \lambda_N| \max\{|b_N'(\sigma_N), b_N'(\lambda_N)|\}\\ &\le C|\sigma_N - \lambda_N|\\ &\le C[|\sigma_N - W'(A)| + |\lambda_N - W'(A)|]\\ &\le C(N-1)^{-1}.\end{aligned}$$

Substituting this estimate into (3.23), we obtain

$$\left|N[W_N(A) - W(A)] - \left(E_N^{\mathrm{cg}}(A) + (\sigma_N - \sigma_0)A - \log\frac{\int_{\mathbb{R}}\exp[-(\psi_1(y) + P(y)) + \sigma_N y]\,\mathrm{d}y}{\int_{\mathbb{R}}\exp[-\psi(y) + \sigma_0 y]\,\mathrm{d}y}\right)\right| \le \frac{C}{N}. \tag{3.24}$$

An analogous argument as in the proof of Proposition 2.7 we obtain

$$\left|\log\frac{\int_{\mathbb{R}}\exp[-(\psi_1(y) + P(y)) + \sigma_N y]\,\mathrm{d}y}{\int_{\mathbb{R}}\exp[-\psi(y) + \sigma_0 y]\,\mathrm{d}y} - \log\frac{\int_{\mathbb{R}}\exp[-(\psi_1(y) + P(y)) + \sigma_0 y]\,\mathrm{d}y}{\int_{\mathbb{R}}\exp[-\psi(y) + \sigma_0 y]\,\mathrm{d}y}\right| \le \frac{C}{N}. \tag{3.25}$$

The assertion (3.15) of Theorem 3.2 is then followed from (3.24), Theorem 3.1, Lemma 3.3 and (3.25). □

We now prove Lemma 3.3.

Proof of Lemma 3.3. Define $L(\sigma) := \frac{1}{N}\Psi_P(\sigma_N) + \frac{1}{N}\sum_{i=2}^N \Psi(\sigma_N - h_i)$. Then we have

$$A = \Psi(\sigma_0) = L(\sigma_N).$$

Hence,

$$L(\sigma_N) - L(\sigma_0) = \Psi(\sigma_0) - L(\sigma_0) = \frac{1}{N}(\Psi(\sigma_0) - \Psi_P(\sigma_0 - h_1)) + \frac{1}{N}\sum_{i=2}^N (\Psi(\sigma_0) - \Psi(\sigma_0 - h_i)).$$

By the mean value theorem, there exists θ such that

$$L(\sigma_N) - L(\sigma_0) = L'(\theta)(\sigma_N - \sigma_0). \tag{3.26}$$

We have

$$|L'(\theta)||\sigma_N - \sigma_0| = |L(\sigma_N) - L(\sigma_0)| \leq \frac{1}{N}\left[|\Psi(\sigma_0) - \Psi_P(\sigma_0 - h_1)| + \sum_{i=2}^N |\Psi(\sigma_0) - \Psi(\sigma_0 - h_i)|\right]$$
$$\leq \frac{1}{N}\left[|\Psi(\sigma_0) - \Psi_P(\sigma_0 - h_1)| + \frac{1}{\kappa_1}\sum_{i=2}^N |h_i|\right]\cdot$$

Since $0 < |L'(\theta)| \leq C$, it implies that

$$|\sigma_N - \sigma_0| \leq \frac{1}{N|L'(\theta)|}\left[|\Psi(\sigma_0) - \Psi_P(\sigma_0 - h_1)| + \frac{1}{\kappa_1}\sum_{i=2}^N |h_i|\right] \leq \frac{C}{N}\cdot$$ □

4. Defect-formation free energy *versus* Gibbs conditioning principle (GCP)

There is a close relationship between the thermodynamic limit studied in this paper with the Gibbs conditioning principle (GCP) in probability theory and statistical mechanics [7, 9]. To make a comparison we first review the result of [7]. Consider an exponential family of probability measures $\{P_\lambda : \lambda \in \Lambda = (\alpha, \beta)\}$ on the fixed interval $I = (a, b)$

$$P_\lambda(\mathrm{d}x) = e^{\lambda x} h(x)\,\mathrm{d}x / c(\lambda) \quad c(\lambda) = \int_I e^{\lambda x} h(x)\,\mathrm{d}x.$$

Let $X_1, \ldots, X_n$ be i.i.d. random variables with common distribution P_λ. For $k < n$, let Q_{nsk} be the law of $X_1, \ldots, X_k$ given $S_n = X_1 + \ldots + X_n = s_n$ which can be computed explicitly by

$$Q_{nsk}(x_1, \ldots, x_k)(dx_1, \ldots, dx_k) = h(x_1)\ldots h(x_k) h^{*(n-k)}(s - x_1 - \ldots - x_k)/h^{*n}(s),$$

where $h^{*j} := h \underbrace{* \ldots *}_{j} h$.

Theorem 4.1. [7], (Thm. 1.6 (a)). *Under certain conditions (smoothness, boundedness of fourth moments, growth condition and a maximal condition) and if $k = o(n)$, then uniformly in s,*

$$\|Q_{nsk} - P_{\lambda^*}^{\otimes k}\| = \gamma\frac{k}{n} + o\left(\frac{k}{n}\right),$$

where $\|\|$ is the total variation distance, $\gamma = (2/\pi e)^{1/2}$ and λ^* solves

$$m_\lambda = \frac{\mathrm{d}}{\mathrm{d}\lambda}\log c(\lambda) = \frac{s}{n}\cdot$$

We note that the GCP can also be found from the theory of large-deviation principle, see *e.g.*, [10] for generalities on large deviation theory and [5, 9] for its applications to the GCP. We discuss here a simplified case that is relevant to this paper. Suppose as above that $X_1, \dots, X_n$ are i.i.d. random variables with common distribution $\pi(\mathrm{d}x)$ (not necessary parameterized by λ as above) and $X_1 + \dots X_n = nA$. The empirical measure associated to these variables is defined by

$$\pi_n(\mathrm{d}x) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}(\mathrm{d}x).$$

The condition $X_1 + \dots X_n = nA$ can be written as $\langle x, \pi_n\rangle := \int_{\mathbb{R}} x\pi_n(dx) = A$. For any two probability measures μ, ν on $\mathbb{R}$, let $H(\mu||\nu)$ be the relative entropy of of μ with respect to ν, *i.e.*,

$$H(\mu||\nu) = \begin{cases} \int_{\mathbb{R}} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right) \mathrm{d}\mu, & \text{if} \quad \mathrm{d}\mu \ll \mathrm{d}\nu, \\ +\infty, & \text{otherwises.} \end{cases}$$

Now the GCP states that the conditional distribution of a finite k number of variables becomes i.i.d. with new condition μ_* obtained by minimizing the relative entropy

$$\inf_{\mu:\langle x,\mu\rangle = A} H(\mu||\pi). \tag{4.1}$$

The solution of this minimization problem satisfies the following Euler-Langrange equation

$$\mu_*(\mathrm{d}x) = \frac{\mathrm{e}^{\sigma_* x}\pi(\mathrm{d}x)}{\int_{\mathbb{R}} \mathrm{e}^{\sigma_* x}\pi(\mathrm{d}x)},$$

where σ_* satisfies

$$A = \int_{\mathbb{R}} x\mu_*(\mathrm{d}x) = \frac{\int_{\mathbb{R}} xe^{\sigma_* x}\pi(\mathrm{d}x)}{\int_{\mathbb{R}} e^{\sigma_* x}\pi(\mathrm{d}x)}.$$

In (4.1) the fact that the functional to be minimized is the relative entropy comes from Sanov's theorem that states that the empirical measure π_N satisfies a large-deviation principle as $N \to \infty$ with a rate function given by the relative entropy $H(\cdot||\pi)$, see *e.g.*, [10].

Now we compare the computations of the defect-formation free energy with the GCP. In the case where the external force is absent, the defect-formation free energy for the specific example that we consider can be regarded as an application of Theorem 4.1. Indeed, using a change of variables $y_i = u_i - u_{i-1}$, we obtain

$$\begin{aligned} G_N^P(A) :&= F_N^P(A) - F_N(A) \\ &= -\log\int_{\mathbb{R}} \exp(-P(y_1)) \frac{\exp(-\psi(y_1)) \int_{\mathbb{R}^{N-2}} \exp(-\sum_{i=2}^{N-1}\psi(y_i) - \psi(NA - y_1 - \sum_{i=2}^{N-1} y_i))\mathrm{d}y_2 \dots \mathrm{d}y_{N-1}}{\int_{\mathbb{R}^{N-1}} \exp(-\sum_{i=1}^{N-1}\psi(y_i) - \psi(NA - \sum_{i=1}^{N-1} y_i))\mathrm{d}y_1 \dots \mathrm{d}y_{N-1}} \mathrm{d}y_1. \end{aligned} \tag{4.2}$$

Let $Y_i, i = 1, \dots, N$ be i.i.d. random variables with common distribution $P_\lambda(y) \propto e^{\lambda y}q(y)$ where $q(y) = \exp(-\psi(y))$. The conditional density of Y_1 given $S := Y_1 + \dots + Y_N = s_N$ is given by

$$\begin{aligned} Q_{Ns_N1} &= \frac{q(y_1)q^{*(N-1)}(s_N - y_1)}{q^{*N}(s_N)} \\ &= \frac{\exp(-\psi(y_1)) \int_{\mathbb{R}^{N-2}} \exp(-\sum_{i=2}^{N-1}\psi(y_i) - \psi(s_N - y_1 - \sum_{i=2}^{N-1} y_i))\,\mathrm{d}y_2 \dots \mathrm{d}y_{N-1}}{\int_{\mathbb{R}^{N-1}} \exp(-\sum_{i=1}^{N-1}\psi(y_i) - \psi(s_N - \sum_{i=1}^{N-1} y_i))\,\mathrm{d}y_1 \dots \mathrm{d}y_{N-1}}. \end{aligned} \tag{4.3}$$

Taking $s_N = NA$ in (4.3) and comparing with (4.2) we obtain that

$$G_N^P(A) = -\log \int_{\mathbb{R}} \exp(-P(y_1)) Q_{N(NA)1}(\mathrm{d}y_1) = -\log \mathbb{E}\Big(\exp(-P(Y_1)) | Y_1 + \ldots + Y_N = NA \Big).$$

According to Theorem 4.1,

$$\|Q_{N(NA)1} - \mu\| = \frac{\gamma}{N} + o\Big(\frac{1}{N}\Big),$$

where μ is given by

$$\mu(y_1) \propto \mathrm{e}^{\lambda^* y_1 - \psi(y_1)}$$

where λ^* solves

$$A = \frac{\int_{\mathbb{R}} y_1 \exp(\lambda y_1 - \psi(y_1))\,\mathrm{d}y_1}{\int_{\mathbb{R}} \exp(\lambda y_1 - \psi(y_1))\,\mathrm{d}y_1},$$

which implies that $\lambda^* = W'(A)$ where

$$W(A) = \sup_{\sigma} \{\sigma A - \log \int_{\mathbb{R}} \exp(\sigma y - \psi(y))\,\mathrm{d}y\}.$$

As a consequence, $|G_N^P(A) - G_\infty(A)| \lesssim 1/N$ uniformly in A where

$$G_\infty(A) = -\log \frac{\int_{\mathbb{R}} \exp(-(\psi + P)(y) + W'(A)y)\,\mathrm{d}y}{\int_{\mathbb{R}} \exp(-\psi(y) + W'(A)y)\,\mathrm{d}y}$$

as expected.

However, when there is external force, Theorem 4.1 can not be applied per se since the variables Y_i are not identically distributed. Instead, Y_i has distribution $\propto \exp(-\psi_i(y))$ with $\psi_i(y) = \psi(y) + h_i y$. The method in [14] (and hence our Lem. 2.2) can be applied to non-identical distributions. The proof of the local Cramér theorem in [14] is an extension of Proof of Theorem 1.6 from [7], to non-identical distributions.

5. Harmonic potentials

In this section, we provide explicit computations for the quadratic case,

$$\psi(y) = \alpha |y|^2, \qquad P(y) = \beta |y|^2, \quad \text{for some} \quad \alpha, \beta > 0. \tag{5.1}$$

5.1. Harmonic potentials without forcing

We recall that

$$F_N(A) = -\log \int_{\mathbb{R}^{N-1}} \exp\left[-\alpha \sum_{i=1}^{N-1} y_i^2 - \alpha \left(NA - \sum_{i=1}^{N-1} y_i \right)^2 \right] \mathrm{d}y_1 \ldots \mathrm{d}y_{N-1}.$$

and

$$F_N^P(A) = -\log \int_{\mathbb{R}^{N-1}} \exp\left[-(\alpha + \beta)\, y_1^2 - \alpha \sum_{i=2}^{N-1} y_i^2 - \alpha \left(NA - \sum_{i=1}^{N-1} y_i \right)^2 \right] \mathrm{d}y_1 \ldots \mathrm{d}y_{N-1}.$$

The main result of the present section is the following.

Theorem 5.1. *The defect-formation free energy is given by*

$$\begin{aligned} G_N(A) &:= F_N^P(A) - F_N(A) \\ &= \frac{1}{2}\log\frac{\alpha+\beta}{\alpha} + \frac{\alpha\beta A^2}{\alpha+\beta} - \frac{N\alpha\beta^2 A^2}{(N(\alpha+\beta)-\beta)^2} + \frac{\alpha\beta A^2}{\alpha+\beta}\left(\frac{2\beta}{N(\alpha+\beta)-\beta} + \frac{\beta^2}{(N(\alpha+\beta)-\beta)^2}\right) \\ &\quad + \frac{1}{2}\log\left(1 - \frac{\beta}{N(\alpha+\beta)}\right)\cdot \end{aligned}$$

The thermodynamic limit is given by

$$G_\infty(A) := \lim_{N\to\infty} G_N(A) = \frac{\alpha\beta\, A^2}{\alpha+\beta} + \frac{1}{2}\log\frac{\alpha+\beta}{\alpha}\cdot$$

Moreover, the following error estimate holds for all $A \in \mathbb{R}$ and $N \geq 2$ and for some positive constant C

$$|G_N(A) - G_\infty(A)| \leq \frac{C}{N}.$$

Proof. The proof consists of lengthy and elementary computations. Hence, we omit it here and refer to the preprint version [6] for detailed computations. ☐

5.2. Harmonic potentials with external forces

Now we consider the quadratic case with external forces. Recall that the perfect energy is

$$\begin{cases} F_N(A) = -\beta^{-1}\log\int_{\mathbb{R}^{N-1}}\exp\Big[-\beta\sum_{i=1}^N \psi(u_i - u_{i-1})\Big]\,\mathrm{d}u_1\ldots\mathrm{d}u_{N-1} \\ u_0 = 0, u_N = NA. \end{cases} \tag{5.2}$$

and the deformed energy is

$$\begin{cases} F_N^P(A) = -\beta^{-1}\log\int_{\mathbb{R}^{N-1}}\exp\Big[-\beta\sum_{i=1}^N \psi_i(u_i - u_{i-1}) - \beta P(u_1)\Big]\,\mathrm{d}u_1\ldots\mathrm{d}u_{N-1} \\ u_0 = 0, u_N = NA, \end{cases} \tag{5.3}$$

where $\psi_i(y) = \psi(y) + h_i y = \alpha y^2 + h_i y$, where $\{h_i\}$ represent the external forces.

The finite coarse-grained energy is given by

$$E_N^{\mathrm{cg}}(y) = \sum_{i=2}^N \left[\frac{1}{4\alpha}(-h_i+\lambda)^2 + \frac{1}{2\alpha}h_i(-h_i+\lambda) - \alpha A^2\right].$$

In view of Assumption 1.2 we define

$$H := \sum_{i=2}^\infty h_i \quad \text{and} \quad \bar{H} = \sum_{i=2}^\infty h_i^2.$$

The main result of this section is the following.

Theorem 5.2.
(1) *The thermodynamic limit has the following explicit formula*

$$G_\infty(A) = \frac{1}{2}\log\frac{\alpha+\beta}{\alpha} + \frac{\alpha\beta A^2}{\alpha+\beta} + \frac{\alpha A h_1}{\alpha+\beta} - \frac{h_1^2}{4(\alpha+\beta)} + AH - \frac{1}{4\alpha}\bar{H}.$$

(2) *The coarse-grained energy is given by*

$$E^{\text{cg}}(y) = \lim_{N\to\infty} E_N^{\text{cg}}(y) = 2\alpha A(A-y) + AH - \frac{1}{4\alpha}\bar{H}.$$

(3) *The thermodynamic limit can be represented as*

$$G_\infty = -\log \frac{\int \exp[-(\psi(y) + P(y) + h_1 y) - E^{\text{cg}}(y)]\,\mathrm{d}y}{\int \exp[-\psi(y) - E^{\text{cg}}_{\mathbf{h}=0}(y)]\,\mathrm{d}y}.$$

Proof. The proof is elementary and lengthy. Hence, we omit it here and refer to the preprint version [6] for detailed computations. □

5.3. Harmonic coarse-graining

In this section, we provide a direct method to coarse-graining for the harmonic case. We consider as before the potential energy

$$V(u) = \sum_{i=1}^{N} \psi(u_i - u_{i-1}), \tag{5.4}$$

and the perturbed energy

$$V(u) + P(u_1) = \sum_{i=1}^{N} \psi(u_i - u_{i-1}) + P(u_1), \tag{5.5}$$

where we consider the harmonic case $\psi(r) = K_1 r^2$ and $P(r) = K_2 r^2$. We are interested in the free energy difference

$$F_N(x,P) - F_N(x,0) = -\log \int_{\mathbb{R}^{N-1}} \exp(-V(u) - P(u)) + \log \int_{\mathbb{R}^{N-1}} \exp(-V(u)). \tag{5.6}$$

As seen above, this can be analytically computed. However, we consider coarse-graining the potential energy and using the free energy difference of the coarse-grained model to approximate the free energy difference for the full model. We show that the free energy difference for the coarsened model is identical to that of the full model.

Since our interactions are first-neighbor only and the defect potential is restricted to the first bond, we leave the first bond fully resolved and use a uniform coarsening elsewhere. That is, associated to the displacement $w \in \mathbb{R}^M$, we have the piecewise linear interpolation operator $I_h : \mathbb{R}^M \to \mathbb{R}^N$ where $(I_h w)_{p(j-1)+1} = u_j$. In particular, $N = p(M-1)+1$. The coarse-grained potential energy is then

$$V_{\text{cg}}(w) = \psi(w_1 - w_0) + \sum_{i=2}^{M} p\psi(p^{-1}(w_j - w_{j-1})) = K_1 w_1^2 + \sum_{i=2}^{M} K_1 p^{-1}(w_j - w_{j-1})^2. \tag{5.7}$$

The technique given here for computing the free energy will differ from that in the previous sections. Here, we successively complete squares on the energy, starting from w_{M-1}, and we define a recurrence for the coefficients

c_i, d_i, and f_i that are introduced in the expansion.

$$
\begin{aligned}
V_{\mathrm{cg}}(w) &= K_1 w_1^2 + \sum_{i=2}^{N} K_1 p^{-1}(w_j - w_{j-1})^2 \\
&= K_1 p^{-1} N^2 x^2 - 2K_1 p^{-1} N x w_{M-1} + \sum_{i=2}^{M-1} K_1 p^{-1}\left[2w_j - 2w_j w_{j-1}\right] + K_1(1+p^{-1})w_1^2 \\
&= K_1 p^{-1}\Bigg[N^2x^2 + 2(w_{N-1} - \frac{1}{2}(Nx + w_{N-2}))^2 - \frac{1}{2}(Nx + w_{N-2})^2 \\
&\quad + \sum_{i=2}^{M-2}(2w_j - 2w_j w_{j-1})\Bigg] + K_1(1+p^{-1})w_1^2 \\
&= K_1 p^{-1}\Bigg[f_i N^2 x^2 + \sum_{i=m}^{M-1} c_i (w_i - c_i^{-1}(w_{i-1} + d_i N x))^2 - c_m^{-1}(w_{m-1} + d_m N x)^2 \\
&\quad + \sum_{i=2}^{m-1}(2w_j - 2w_j w_{j-1})\Bigg] + K_1(1+p^{-1})w_1^2,
\end{aligned}
$$

where the coefficients satisfy the following recurrences:

$$
\begin{aligned}
c_{i-1} &= 2 - c_i^{-1} \qquad c_{M-1} = 2 \\
d_{i-1} &= \frac{d_i}{c_i} \qquad d_{M-1} = 1 \\
f_{i-1} &= f_i - \frac{d_i^2}{c_i} \qquad f_{M-1} = 1.
\end{aligned}
$$

We then find for $i = 2, \ldots, M-1$,

$$
\begin{aligned}
c_i &= \frac{M-i+1}{M-i} \\
d_i &= \frac{1}{M-i} \\
f_i &= \frac{1}{M-i}.
\end{aligned}
$$

So, for the coarse-grained energy, we compute:

$$
\begin{aligned}
V_{\mathrm{cg}}(w) = K_1 p^{-1}\Bigg[&\sum_{i=2}^{M-1} c_i (w_i - c_i^{-1}(w_{i-1} + d_i N x))^2 \\
&+ \left(\frac{M}{M-1} + p - 1\right)\left(w_1 - \frac{d_1 N x}{c_1 + p - 1}\right)^2 + \frac{N^2 x^2 p}{M + (p-1)(M-1)}\Bigg],
\end{aligned}
$$

where the lowest order terms do not satisfy the recursion because of the factor of p, but they are computed manually. Using the same recursion, we can also transform the energy with the defect, taking care to modify the lowest-order term.

$$
\begin{aligned}
V_{\mathrm{cg}}(w) + P(w_1) = K_1 p^{-1} &\left[\sum_{i=2}^{M-1} c_i (w_i - c_i^{-1}(w_{i-1} + d_i N x))^2 \right] \\
&+ \left(\frac{K_1}{p} \left(\frac{M}{M-1} + p - 1 \right) + K_2 \right) \left(w_1 - \frac{K_1 d_1 N x}{K_1(c_1 + p - 1) + K_2 p} \right)^2 \\
&+ \frac{K_1 N^2 x^2 (K_1 + K_2)}{K_1(M + (p-1)(M-1)) + K_2 p (M-1)}.
\end{aligned}
$$

When we take free energy differences, we can directly integrate starting from w_{M-1} downwards, and the only differences in the two energies are in the lowest-order terms. Also, we note that $M + (p-1)(M-1) = N$, and $p(M-1) = N-1$, so that the p will fall out. We have

$$
F_M^{\mathrm{cg}}(x, P) - F_M^{\mathrm{cg}}(x, 0) = \frac{K_1 N^2 x^2 (K_1 + K_2)}{K_1 N + K_2(N-1)} - \frac{N^2 x^2}{N} + \frac{1}{2} \log \left[\frac{K_1 N + K_2 (N-1)}{K_1 N} \right].
$$

We note that this is exactly the result arrived at in Section 5.1, and that there is no p or M dependence here. That is, any uniform coarse-graining of the chain that leaves the first bond refined exactly computes the free energy difference.

6. Numerical free energy

We present numerical experiments to illustrate the results of the paper using standard free energy computation techniques as in [13]. We compare the finite chain energy G_N, coarse grained energy G_N^{cg}, and G_∞ computed using numerical quadrature of the limit expression. We see the theoretically expected N^{-1} rate of convergence, where the asymptotic rate is observed to be valid even for small N. An application of the triangle inequality shows that $G_N - G_N^{\mathrm{cg}}$ decays at least as fast as N^{-1}, and we observe that N^{-1} is the asymptotic decay rate in our experiments.

6.1. Free energy perturbation

A standard approach for computing free energy differences is called the free energy perturbation technique which rewrites the free energy difference as an ensemble average of the energy perturbation with respect to the invariant measure of the unperturbed system. To compute the free energy difference between V and V^P, we write

$$
\begin{aligned}
G_N = F_N^P - F_N &= -\log \frac{\int_\Gamma \exp(-V^P(z)) \mathrm{d}z}{\int_\Gamma \exp(-V(z)) \mathrm{d}z} \\
&= -\log \frac{\int_\Gamma \exp(-(V^P(z) - V(z))) \exp(-V(z)) \mathrm{d}z}{\int_\Gamma \exp(-V(z)) \mathrm{d}z} \\
&= -\log \langle \exp(-P(u)) \rangle_{\mu_0}.
\end{aligned}
$$

Therefore, one samples $\exp(-P(u))$ with respect to the invariant measure given by V.

6.2. Staging

Direct sampling to compute the free energy perturbation can be very slow to converge when $V^P - V$ is large, particularly when the minima of V and V^P are separated. Many samples are chosen near the global minimum of V, which may not significantly contribute to the value of the integral. Instead, one can employ staging, where the free energy difference is broken into a telescopic sum. That is, we write $V_\lambda = V + \lambda P$, and $F_\lambda = -\beta^{-1} \log \int_\Gamma V_\lambda(z)\,\mathrm{d}z$. Then the free energy difference can be written

$$F_N^P - F_N = \sum_{i=1}^{N_{\text{stages}}} F_{\lambda_i} - F_{\lambda_{i-1}},$$

so that one must sample $\exp(-\beta(\lambda_i - \lambda_{i-1})P)$ with respect to the invariant measure corresponding to $V_{\lambda_{i-1}}$. Since the energies V_{λ_i} and $V_{\lambda_{i-1}}$ are closer than V and V^P, this speeds convergence and reduces the overall computed variance.

6.3. Metropolis adjusted langevin algorithm

In the following, we apply the Metropolis Adjusted Langevin Algorithm (MALA), which proceeds as a series of overdamped Langevin steps followed by an accept/reject step:

$$q^* = q^n - h\nabla V(q^n) + \sqrt{h}G \quad \text{where} \quad G \sim \mathcal{N}(0, Id).$$

Then we accept the new step and set $q^{n+1} = q^*$ with probability

$$r(q^n, q^*) = \min\left(1, \frac{T(q^*, \mathrm{d}q^n)\mu(\mathrm{d}q^*)}{T(q^n, \mathrm{d}q^*)\mu(\mathrm{d}q^n)}\right),$$

where

$$T(q, dq') = \left(\frac{1}{4\pi h}\right)^{d/2} \exp\left(\frac{-|q' - q + h\nabla V|^2}{4h}\right).$$

Otherwise, we set $q^{n+1} = q^n$. The accept/reject step assures that we are sampling the invariant measure μdq for any stepsize h. The choice of h is driven by two competing interests: larger h speeds up convergence from the initial condition to the invariant measure, whereas smaller h means that a step is more likely to be accepted.

6.4. Unforced nonlinear chain

We consider the nonlinear energy

$$\psi(r) = \frac{1}{2}(r-1)^4 + \frac{1}{2}r^2, \tag{6.1}$$

which was the test case used in [1]. Note that while this does not satisfy the upper bound of the growth condition in (1.7), we do observe the expected rates of convergence. We take a harmonic defect perturbation $P(y) = y^2$ and choose $A = 2$.

The free difference G_N is sampled using the MALA algorithm with 100 staging steps and 100 independent replicas to compute confidence intervals. In addition, the coarse-grained approximation G_N^{cg} is also computed. Due to the 1D nature of the problem, the minimizer for the CG energy is given by an affine function, so that the computations involved are low-dimensional integrals. First the energy density

$$W(A) = \sup_{\sigma\in\mathbb{R}} \left\{\sigma A - \log \int_{\mathbb{R}} \exp(-\psi(y) + \sigma y)\,\mathrm{d}y\right\}$$

is computed by quadrature, giving coarse-grained energy

$$E_N^{\text{cg}}(y) = (N-1)\left[W\Big(A + \frac{A-y}{N-1}\Big) - W(A)\right].$$

Then we may compute

$$G_N^{\text{cg}} = -\log \frac{\int \exp(-P(y) - \psi(y) - E_N^{\text{cg}}(y)\,\mathrm{d}y}{\int \exp(-\psi(y) - E_N^{\text{cg}}(y))\,\mathrm{d}y}$$

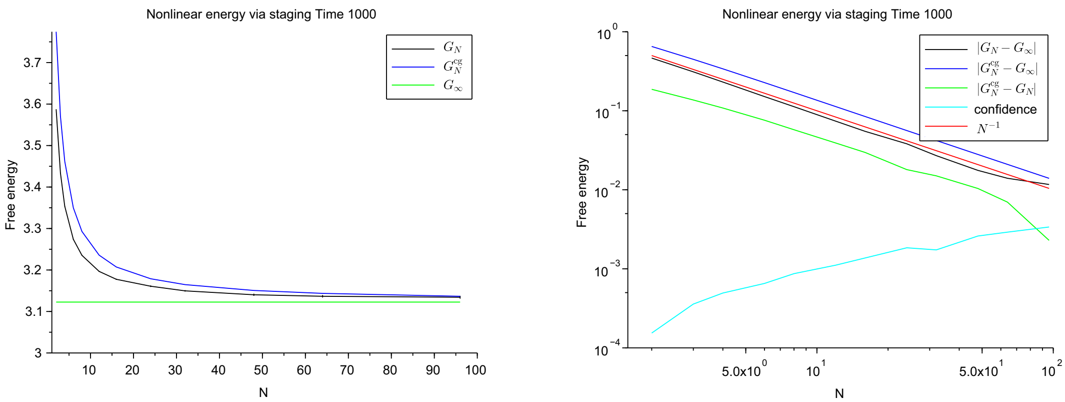

FIGURE 1. For the nonlinear potential, the free energy difference is sampled using staging, and the result is compared to the coarse-grained approximation. The limiting free energy is computed *via* numerical quadrature and plotted in green. On the right, we show the rate of convergence to the limiting energy G_∞, where both approximations show $O(N^{-1})$ convergence. The difference between G_N and G_N^{cg} is also $O(N^{-1})$.

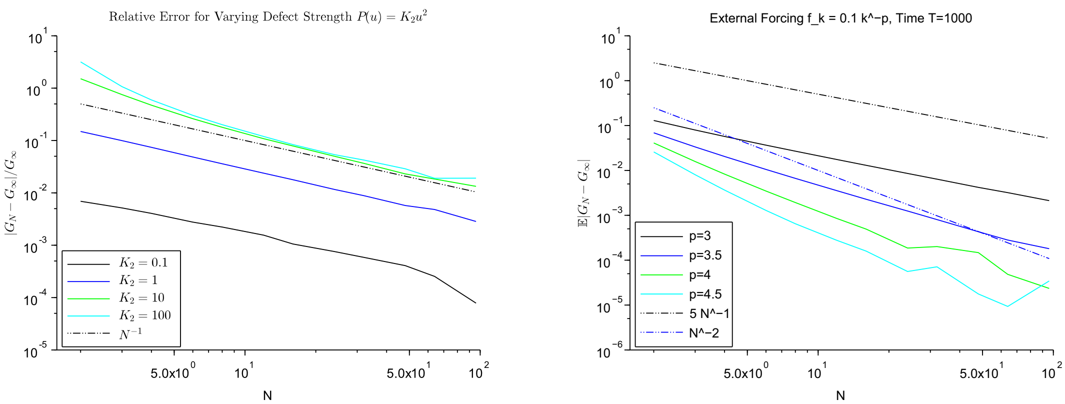

FIGURE 2. Convergence is studied under varying defect behavior. On the left, the strength of the defect potential $P(u) = K_2 u^2$ is varied and the relative error in the free energy is plotted. In each case we observe the expected N^{-1} error, and the error is much lower for a weaker external potential. On the right, a nonlinear chain is sampled where the defect is modeled by decaying forces $f_j = 0.1 j^{-p}$, and we show the rate of convergence to the limiting energy G_∞, where the approximations seem to have p-dependent rates of convergence. Note that for exponents $p = 4$ and $p = 4.5$, the computed energy quickly approaches the limiting energy up to statistical noise.

using standard quadrature techniques. In Figure 1, the sampled free energy difference G_N is compared to G_N^{cg} as well as G_∞. The $O(N^{-1})$ convergence is seen throughout the chosen range of N. We observe through the numerics that $|G_N - G_N^{cg}|$ is also $O(N^{-1})$.

6.5. Varying defect properties

We consider the convergence properties as the defect is varied. We first examine the nonlinear chain as above, with varying strengths for the defect potential, choosing $P(u) = K_2 u^2$ where K_2 will vary over $K_2 = 0.1, 1, 10, 100$. As a second example, we compute the free energy difference with external forces but no defect potential, $P(u) = 0$. The external forces are only present in the defective chain, and the forces impose effective decay rates for the defect providing an analog for the slow decay in the elastic field that surrounds defects in higher dimensional problems. The non-defective chain has nonlinear interaction potentials (6.1), and the defective chain has external forces $f_j = 0.1j^{-p}$ on each degree of freedom u_j, or $h_i = -\sum_{j=i}^{N-1} f_j$. The free energy G_N chain is sampled using MALA with 100 stages, and the limiting expression for G_∞ (3.15) is computed numerically, where it is noted that the minimization problem in the limit separates into single variable problems. As the forces decay sufficiently fast, a Taylor series approximation is used for all but the first four terms in $E^{cg}(A)$.

In Figure 2, the differences $G_N - G_\infty$ are plotted for various rates of decay in the external forces $f_j = 0.1j^{-p}, \quad p = 3, 3.5, 4, 4.5$. The observed rates of convergence depend on the decay rate and are observed to be faster than $\mathrm{O}(N^{-1})$.

7. Conclusion

We have provided a rigorous analysis of the defect-formation free energy (1.6) for a one-dimensional, nearest neighbour chain with nonlinear local defect and external forces. The limiting energy is written in terms of a coarse-grained energy that is based on the Cauchy-Born strain energy density. The form of the coarse-grained energy was chosen because its variational structure is amenable to analysis and approximation by methods in variational mechanics.

The analysis required many restrictions on the model. The nonlinear perturbation P could be extended to a finite region rather than the first bond without additional difficulty. Including interactions beyond nearest neighbour in V would entail extension of the arguments here, for example the bonds are no longer independently distributed in Lemma 2.2, compare the work done for the free energy density in [1]. Moving beyond one spatial dimension for the chain requires significant additional work; however, the inclusion of external forces was motivated in part by the higher dimensional cases as a way to model slowly-decaying stress field around a defect present in dimensions higher than one.

Acknowledgements. MHD and CO were supported by ERC Starting Grant 335120. MD was partially supported by the DARPA EQUiPS program under the grant W911NF1520122

References

[1] X. Blanc, C.L. Bris, F. Legoll and C. Patz, Finite-temperature coarse-graining of one-dimensional models: Mathematical analysis and computational approaches. *J. Nonlin. Sci.* **20** (2010) 241–275.

[2] X. Blanc and F. Legoll, A numerical strategy for coarse-graining two-dimensional atomistic models at finite temperature: The membrane case. *Comput. Mat. Sci.* **66** (2013) 84–95.

[3] P. Caputo, Uniform Poincaré inequalities for unbounded conservative spin systems: the non-interacting case. *Stochastic Process. Appl.* **106** (2003) 223–244.

[4] W.D. Callister and D.G. Rethwisch. Materials Science and Engineering: An Introduction. Wiley (2010).

[5] I. Csiszár, Sanov property, generalized I-projection and a conditional limit theorem. *Ann. Probab.* **3** (1975) 146–158.

[6] M. Dobson and M. H. Duong and C. Ortner, On assessing the accuracy of defect free energy computations. To published in: *ESAIM: M2AN* DOI: 10.1051/m2an/2017052 (2018).

[7] P. Diaconis and D.A. Freedman, Conditional limit for exponential families and finite versions of de Finetti's theorem. *J. Theor. Probab.* **1** (1988) 381–410.

[8] L.M. Dupuy, E.B. Tadmor, R.E. Miller and R. Phillips, Finite-temperature quasicontinuum: Molecular dynamics without all the atoms. *Phys. Rev. Lett.* **95** (2005) 060202.

[9] A. Dembo and O. Zeitouni, Refinements of the Gibbs conditioning principle. *Probability Theory and Related Fields* **104** (1996) 1–14.

[10] A. Dembo and O. Zeitouni. Large deviation techniques and applications. Springer, New York. (1998).

[11] N. Grunewald, F. Otto, C. Villani and M.G. Westdickenberg, A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (2009) 302–351.

[12] F.W. Herbert, A. Krishnamoorthy, W. Ma, K.J. Van Vliet and B. Yildiz, Dynamics of point defect formation, clustering and pit initiation on the pyrite surface. *Electr. Acta* **127** (2014) 416–426.

[13] T Lelièvre, M Rousset and G Stoltz. Free Energy Computations. Imperial College Press (2012).

[14] G. Menz, LSI for Kawasaki dynamics with weak interaction. *Commun. Math. Phys.* **307** (2011) 817–860.

[15] J. Marian, G. Venturini, B.L. Hansen, J. Knap, M. Ortiz and G.H. Campbell, Finite-temperature extension of the quasicontinuum method using langevin dynamics: entropy losses and analysis of errors. *Modell. Simul. Mater. Sci. Eng.* **18** (2010) 015003.

[16] A. Putnis. An Introduction to Mineral Sciences. Cambridge University Press. Cambridge Books Online (1992)

[17] E.G. Seebauer and M.C. Kratzer, Fundamentals of defect ionization and transport. In Charged Semiconductor Defects, Engineering Materials and Processes. Springer London (2009).

[18] A.V. Shapeev and M. Luskin, Accuracy of computation of crystalline defects at finite temperature. Preprint arXiv:1409.5739 (2014).

[19] E.B. Tadmor, F.F. Legoll, W.K. Kim, L.M. Dupuy and R.E. Miller, Finite-Temperature Quasi-Continuum. *ASME. Appl. Mech. Rev.* **65** (2013) 010803–010803-27.

[20] A. Walsh, A.A. Sokol and C.R.A. Catlow, Free energy of defect formation: Thermodynamics of anion frenkel pairs in indium oxide. *Phys. Rev. B* **83** (2011) 224105.