

MAXIMUM ENTROPY METHODS AS THE BRIDGE BETWEEN MICROSCOPIC AND MACROSCOPIC THEORY

JAMIE M. TAYLOR

*Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Road,
Oxford, OX2 6GG*

jamie.taylor@maths.ox.ac.uk

Received (Day Month Year)

Revised (Day Month Year)

Communicated by (xxxxxxxxxx)

This paper investigates a function of macroscopic variables known as the singular potential, building on previous work by Ball and Majumdar. The singular potential is a function of the admissible statistical averages of probability distributions on a state space, defined so that it corresponds to the maximum possible entropy given known observed statistical averages, although non-classical entropy-like objective functions will also be considered. First the set of admissible moments must be established, and under the conditions presented in this work the set is open, bounded and convex allowing a description in terms of supporting hyperplanes, which provides estimates on the development of singularities for related probability distributions. Under appropriate conditions it is shown that the singular potential is strictly convex, as differentiable as the microscopic entropy and blows up uniformly as the macroscopic variable tends to the boundary of the set of admissible moments. Applications of the singular potential are then discussed, and particular consideration will be given to certain free-energy functionals typical in mean-field theory, demonstrating an equivalence between certain microscopic and macroscopic free-energy functionals. This allows statements about L^1 -local minimisers of Onsager's free energy to be obtained which cannot be given by two-sided variations, and overcomes the need to ensure local minimisers are bounded away from zero and infinity before taking bounded variations. The analysis also permits the definition of a dual order parameter for which Onsager's free energy allows an explicit representation. Also the difficulties in approximating the singular potential by everywhere defined functions, in particular by polynomials, are addressed with examples demonstrating the failure of the Taylor approximation to preserve shape properties of the singular potential.

Keywords: Order parameter constraints; mean-field theory; maximum entropy methods; liquid crystals.

AMS Subject Classification: 46N10, 49N99, 82B26

1. Introduction

In many-body problems in physics it is often required to reduce the complexity of the problem by applying statistical methods. Consider a state space X with a corresponding measure μ , and a probability distribution ρ on X that describes the probability of a given body occupying the state $t \in X$. Two particular examples are firstly nematic liquid crystals, where axially symmetric molecules can be described

by their orientation, so that $X = \mathbb{S}^2$ ²⁴, and the Boltzmann equation with state space $X = \mathbb{R}^3 \times \mathbb{R}^3$ corresponding to the position and momentum of particles ¹⁶. Thermodynamic equilibria can then be described as minima of free energy functionals on $\mathcal{P}(X) = \{\rho \in L^1(X) : \int_X \rho d\mu = 1, \rho \geq 0 \text{ a.e}\}$, the set of probability distributions on X . A particularly common example is the mean-field free energy, based on the second virial expansion and due to Onsager ²⁵ and is typically given by

$$\mathcal{I}_{\mathcal{P}}(\rho) = T \int_X \rho(t) \ln \rho(t) d\mu(t) - \frac{1}{2} \int_X \int_X K(s, t) \rho(s) \rho(t) d\mu(s) d\mu(t). \quad (1.1)$$

where the function $K \in L^\infty(X \times X)$ is a symmetric positive kernel, $T > 0$ represents temperature and μ is a measure on X . This is analogous to the Helmholtz free energy, with the left-hand term representing an entropic contribution and the right-hand term representing chemical energy. This work will be concerned with the commonly considered case where the kernel K is of the form

$$K(s, t) = \sum_{i,j=1}^k c_{ij} a_i(t) a_j(s) \quad (1.2)$$

for some $k \in \mathbb{N}$, constants c_{ij} , and a set of linearly independent functions $a_i \in L^\infty(X)$ for $i = 1, \dots, k$. The functions a_i will often be denoted by a single vector valued function, $a = (a_i)_{i=1}^k \in L^\infty(X, \mathbb{R}^k)$. An alternative way to approach such problems is by considering only macroscopic variables typically called order parameters, defined as statistical averages corresponding to the mean field model by

$$b_i = \int_X \rho(t) a_i(t) d\mu(t). \quad (1.3)$$

The vector of moments will be denoted $b = (b_i)_{i=1}^k \in \mathbb{R}^k$. By considering a finite dimensional state space of order parameters and disregarding the full statistical nature of the problem, analysis becomes much simpler through tools such as the Landau expansion (see for example Ref. 29 for a broad review). Within the context of the Q-tensor model for nematic liquid crystals, Majumdar noted that by losing the statistical nature of the problem, physical constraints on b_i can be lost ²¹. For example from the Hölder inequality it must hold that $|b_i| \leq \|a_i\|_\infty$ for each i . Motivated by the notation used in nematic liquid crystals, define \mathcal{Q} to be the set of admissible order parameters,

$$\mathcal{Q} = \left\{ \left(\int_X a_i(t) \rho(t) dt \right)_{i=1}^k : \rho \in \mathcal{P}(X) \right\} \subset \mathbb{R}^k. \quad (1.4)$$

Majumdar investigated conditions under which equilibrium values of the order parameters are physical, in that they are elements of \mathcal{Q} , demonstrating that the Landau model can fail in this sense. Motivated by this problem Ball and Majumdar ⁵ defined a *singular potential* ψ_s on \mathcal{Q} within the Q-tensor theory of nematics that

builds on earlier work by Katriel *et. al.*¹⁸ The singular potential is a convex function, inspired by the entropic term in the mean-field free energy, that blows up as the order parameters approach the boundary of \mathcal{Q} . Whilst the work of Ball and Majumdar concentrated on ensuring physicality in static problems relating to nematic liquid crystals, it has further been applied to dynamic systems^{14,31}, nematic elastomers¹⁰, and the derivation of Q-tensor models¹⁵, demonstrating the versatility of the framework. The aim of this work is to extend the ideas of Ball and Majumdar to define a singular potential in a more general setting, as well as analyse its properties and develop applications.

Before the singular potential can be defined the set of admissible moments \mathcal{Q} must first be understood. This will form the bulk of Section 2. This is an example of the problem of moments², which in the general case is poorly understood. The classical problem considers $X = \mathbb{R}$, and functions $a_i(x) = x^i$, although others have considered the more general setting. For example, Lewis has provided a characterisation of the set of moments that roughly corresponds to the existence of solutions to certain optimisation problems²⁰. In Section 2 a more geometric description of the set \mathcal{Q} will be useful, in which the set is described in terms of supporting hyperplanes. Whilst this description of \mathcal{Q} is abstract and in general is unlikely to give an explicit expression for \mathcal{Q} it is nonetheless appropriate for the analysis in this work. In particular, this characterisation is used in providing growth bounds on the singular potential. Throughout both Sections 2 and 3 the assumption is made that the constraint functions $(a_i)_{i=1}^k$ and the constant function $a_0(t) = 1$ form a pseudo-Haar set of functions, a property first defined by Borwein and Lewis⁷, which provides an elegant theory without being overly restrictive to applications in mean-field theory.

The singular potential is defined and analysed in Section 3. The singular potential corresponds to the greatest possible entropy of a probability distribution, subject to observed moments. This is an application of the principle of maximum entropy, pioneered in the seminal work of Jaynes¹⁷, which has since been influential in a vast array of applications. Mathematically it can be phrased as a convex optimisation problem for the probability distribution subject to linear constraints. For the bulk of this work more general objective functions will be considered than the Shannon entropy, which is given by $\phi(x) = x \ln x$. The objective function will be required to share similar properties to the Shannon entropy (see Definition 3.1). One of these properties is that the objective function is not differentiable at 0, which combined with the non-negativity constraint $\rho \geq 0$ has a consequence of not permitting two-sided variations about the minimiser ρ^* , unless it is known beforehand that $\text{ess inf}_{t \in X} \rho^*(t) > 0$. Borwein and Lewis⁷ provide an ideal framework for this kind of problem, based on duality, that will be exploited in this work. The main results of this section are that, under appropriate conditions, the singular potential ψ_s is strictly convex, has the same differentiability as the objective function, and that $\psi_s(b)$ blows up to $+\infty$ uniformly as the distance from b to $\partial\mathcal{Q}$ approaches zero.

Finally in Section 4 some potential applications are discussed. The results

demonstrate an equivalence between certain minimisation problems in $\mathcal{P}(X)$ and \mathcal{Q} , and demonstrate the possibility of using the singular potential to rephrase harder questions relating to functional analysis in terms of simpler questions of several variable calculus. Particular mention is given to the mean-field approximation, where in the literature a rigorous treatment of the non-negativity constraint is often neglected, and the results in Section 4 provide an existence proof for minimisers as well as demonstrating that the solutions obtained using the first variation are correct. Furthermore the equivalence of local/global minimisers of the macroscopic and microscopic free energies has the implication that under the assumptions presented in this work, equilibria of mean-field free energies can be described using only macroscopic variables and the optimal entropy assumption. As noted by Decarreau *et. al*¹², the dual formulation of the optimisation problem, which rephrases the optimisation problem in terms of finitely many Lagrange multipliers, is the most desirable from the point of numerical analysis. The “moment-space” representation outlined in this work is loosely speaking equivalent, permits a simpler state space for analytical problems, and also allows the problem to be phrased directly in terms of order parameters. Expanding further on the work of Ball and Majumdar⁴, models with spatial inhomogeneities will be considered and it is shown that for a particular class of free-energy functionals, minimisers are strictly physical, in the sense that the minimiser is bounded away from $\partial\mathcal{Q}$. However the class of models in which this analysis works is rather limited so that it remains open if more general free energies have strictly physical minimisers.

The remainder of Section 4 discusses issues surrounding the approximation of the singular potential by globally defined functions, in particular by polynomials. The Landau theory of phase transitions²⁹ states that since the free energy should be analytic in the order parameters, one can consider a Taylor expansion of the free energy to low (typically fourth) order and perform an energy minimisation over the approximation. It is shown that even within some simple one dimensional examples, the fourth order Taylor approximation to the singular potential in general fails to reproduce various desirable properties of the singular potential such as convexity, the existence of a single critical point and even a lower bound. Whilst it is possible to approximate the singular potential by convex polynomials due to a Weierstrass-type result, constructive methods appear to be out reach. Similarly, due to the difficulty of establishing the convexity of fourth order polynomials in several variables, even providing conditions for the Taylor approximation to be convex appears to be unobtainable. In response to this, the Yosida-Moreau approximation²³ is suggested as a globally defined approximation that preserves meaningful properties of the singular potential. The drawback however is that the Yosida-Moreau approximation is similarly defined through an optimisation procedure that, loosely speaking, is no less, although fortunately no more, difficult to solve than the optimisation problem that defines the singular potential itself. To illustrate issues with approximation and provide a concrete example of the singular potential framework,

the McMillan model ²² for isotropic-nematic-smectic A phase transitions is used as a guiding example. The isotropic phase corresponds to a disordered liquid, the nematic phase corresponds to molecules having orientational, but no positional order, and the smectic A phase corresponds to molecules having orientational order, as well as positional order in one direction, parallel to the molecular orientation. This model is described by a mean field free energy, with state space $X = \mathbb{S}^2 \times [0, 1]$, and two order parameters (S, σ) defined as

$$\begin{aligned} S &= \frac{1}{2} \int_0^1 \int_{\mathbb{S}^2} \rho(p, x) (3 \cos(p \cdot e_1)^2 - 1) d\mathcal{H}^2(p) dx \\ \sigma &= \frac{1}{2} \int_0^1 \int_{\mathbb{S}^2} \rho(p, x) (3 \cos(p \cdot e_1)^2 - 1) \cos(2\pi x) d\mathcal{H}^2(p) dx. \end{aligned} \tag{1.5}$$

Here e_1 is a unit vector. If $S = \sigma = 0$ then the system is in the isotropic phase, if $S \neq 0$ and $\sigma = 0$ then the system is in the nematic phase, and finally if $S \neq 0$ and $\sigma \neq 0$, the system is in the smectic A phase.

2. The Moment Problem

The key condition required for the framework presented to be successful is the pseudo-Haar property of a finite set of functions.

Definition 2.1 (Pseudo-Haar functions, see Ref. 8). Let (X, μ) be a finite measure space. Let $k \in \mathbb{N}$ and $f_i : X \rightarrow \mathbb{R}$, for $i = 1, \dots, k$. The set $\{f_i : i = 1, \dots, k\}$ is called pseudo-Haar if for every $Y \subset X$ with $\mu(Y) > 0$, the set $\{f_i|_Y : i = 1, \dots, k\}$ is linearly independent.

Remark 2.1. If (X, μ) is a measure space with atoms, so that there exists some $t \in X$ with $\mu(\{t\}) > 0$, then any set of two or more functions on X cannot be pseudo-Haar, since restricted to the set $\{t\}$ one is simply a multiple of the other. In particular the theory presented here is not applicable to discrete state spaces, such as X a finite subset of \mathbb{N} .

Lewis ¹⁹ showed that for X a subset of \mathbb{R}^n , and $(f_i)_{i=1}^k$ analytic and linearly independent on a connected neighbourhood of X then the functions are pseudo-Haar with respect to the Lebesgue measure. This result can be slightly extended to manifolds, which has applications in liquid crystal theory, where in particular X is taken as \mathbb{S}^2 for axially symmetric molecules or $\text{SO}(3)$ for molecules with lower symmetry. By phrasing the problem in local coordinates, the result of Lewis can be used to extend the result to when the domain is a Riemannian manifold, and the measure taken will be the one induced by the metric, to be denoted by μ_g . The result can trivially be extended for any other measure μ' such that for all $A \subset X$, $\mu'(A) = 0 \Rightarrow \mu_g(A) = 0$. For a relatively self contained introduction to the theory of integration on manifolds, the reader is directed to Ref. 3. Before attempting the proof, some preliminaries will be taken covered.

Definition 2.2. Let be X an analytic manifold with atlas $((U_\alpha, \phi_\alpha))_{\alpha \in J}$, where J

is some index set. Say that $f : X \rightarrow \mathbb{R}$ is analytic if $f \circ \phi_\alpha^{-1} : \phi(U_\alpha) \rightarrow \mathbb{R}$ is analytic for all $\alpha \in J$.

Proposition 2.1 (Proposition 12.1.6 of Ref. 3). *Let (X, g) be an n -dimensional Riemannian manifold with induced measure μ_g . Then $\mu_g(A) = 0$ if and only if $\mathcal{L}^n(\phi(A \cap U)) = 0$ for all charts (U, ϕ) of X .*

Proposition 2.2 (From Ref. 19). *Let $\Omega \subset \mathbb{R}^n$ be open and connected, and let $f : \Omega \rightarrow \mathbb{R}$ be analytic. Then if there exists some set $A \subset \Omega$ with $\mathcal{L}^n(A) > 0$ so that $f|_A = 0$, then $f = 0$.*

Using these results it is now straightforward to obtain the required result.

Proposition 2.3. *Let (X, g) be a connected, analytic, Riemannian manifold. Denote by μ the measure on X induced by the metric g . Let $f_i : X \rightarrow \mathbb{R}$ be analytic functions on X for $i = 1, \dots, k$ for some $k \in \mathbb{N}$. Then the set $\{f_i : i = 1, \dots, k\}$ is linearly independent if and only if it is pseudo-Haar.*

Proof. If the functions are pseudo-Haar then linear independence is immediate. To show the converse, assume the result is false. This implies that there exists some $\xi \in \mathbb{R}^k$ with $\xi \neq 0$ and $A \subset X$ so that $\mu_g(A) > 0$ and $\sum_{i=1}^k \xi_i f_i(x) = 0$ for all $x \in A$.

For simplicity denote $F = \sum_{i=1}^k \xi_i f_i$. Then it is clear that F is analytic, and $F|_A = 0$.

Without loss of generality take all charts to be connected by splitting the charts into connected components where necessary. Using Prop. 2.1, there must exist some chart (U, ϕ) so that $\mathcal{L}^n(\phi(A \cap U)) > 0$. For brevity let $\tilde{U} = \phi(U)$ and $\tilde{A} = \phi(A \cap U)$. Also let $\tilde{F} = F \circ \phi^{-1} : \tilde{U} \rightarrow \mathbb{R}$. Since F is analytic, then \tilde{F} is analytic. Therefore since $\tilde{F}|_{\tilde{A}} = 0$ with $\mu_g(\tilde{A}) > 0$, and \tilde{F} analytic on the connected open set \tilde{U} , by Prop. 2.2 it must hold that $\tilde{F} = 0$ on \tilde{U} . Composing \tilde{F} with ϕ gives that $F = 0$ on U , which is open. Therefore since X is connected and analytic, F is zero on all of X . Recalling the definition of F , this implies $\sum_{j=1}^k \xi_j f_j = 0$, contradicting the linear independence assumption and completing the proof. \square

Definition 2.3 (The sets $\mathcal{P}(X)$ and \mathcal{Q}). Let (X, μ) be a finite measure space. Consider finitely many constraint functions $a_i \in L^\infty(X)$ for $i = 1, \dots, k$, such that the set $\{t \mapsto 1\} \cup \{a_i : i = 1, \dots, k\}$ is pseudo-Haar. For ease of notation let $a = (a_i)_{i=1}^k \in L^\infty(X, \mathbb{R}^k)$. Define the sets

$$\begin{aligned} \mathcal{P}(X) &= \left\{ \rho \in L^1(X) : \rho \geq 0 \text{ } \mu\text{-a.e., } \int_X \rho d\mu = 1 \right\}, \\ \mathcal{Q} &= \left\{ \int_X a \rho d\mu : \rho \in \mathcal{P}(X) \right\} \subset \mathbb{R}^k. \end{aligned} \tag{2.1}$$

Furthermore, if $b \in \mathcal{Q}$, then b will be called physical, and if $b = \int_X \rho a d\mu$, then b is generated by ρ .

Definition 2.4. Let $u \in \mathbb{S}^{k-1}$. Given the constraint functions $(a_i)_{i=1}^k$ with $a_i \in L^\infty(X)$ for $i = 1, \dots, k$, define $S_u \in \mathbb{R}$ by

$$S_u = \operatorname{ess\,sup}_{t \in X} u \cdot a(t). \quad (2.2)$$

Let $\epsilon > 0$. Define the set $E_\epsilon^u \subset X$ as

$$E_\epsilon^u = \{t \in X : S_u < u \cdot a(t) + \epsilon\}. \quad (2.3)$$

Note that since $\mu(X) < \infty$, $L^\infty(X) \subset L^2(X)$, so that up to an invertible linear transformation, it can be assumed that the functions $(a_i)_{i=1}^k$ are $L^2(X)$ orthonormal functions, so that $\int_X a_i a_j d\mu = \delta_{ij}$, and orthogonal to any constant so that $\int_X a_i d\mu = 0$. For simplicity of calculation, this will be assumed unless stated otherwise, although it has little effect on the results since the non-normalised framework is equivalent up to an affine map on \mathcal{Q} . This has the consequence that the uniform distribution, $\rho_U(t) = \frac{1}{\mu(X)}$, satisfies $\int_X \rho_U(t) a(t) d\mu(t) = 0$. This is relevant in Landau theory where the order parameters are required to be zero in the high temperature uniform state. In order to establish a necessary and sufficient condition for $b \in \mathcal{Q}$, some properties of \mathcal{Q} will be established. The method is to describe \mathcal{Q} in terms of supporting hyperplanes.

Proposition 2.4. *The set \mathcal{Q} is convex, bounded and open.*

Proof. Convexity is immediate from the convexity of $\mathcal{P}(X)$. To show boundedness, let $b \in \mathcal{Q}$, corresponding to some $\rho \in \mathcal{P}(X)$. Then

$$\begin{aligned} |b| &= \left| \int_X a(t) \rho(t) d\mu(t) \right| \\ &\leq \operatorname{ess\,sup}_{t \in X} |a(t)| \end{aligned} \quad (2.4)$$

so that \mathcal{Q} is bounded. To see that \mathcal{Q} is open, consider some $b \in \mathcal{Q}$ and let $\xi \in \mathbb{R}^k$. Since the functions $t \mapsto 1$ and a_i are pseudo-Haar, by Ref. 7 there exists some $\rho \in \mathcal{P}(X)$ that generates b and $\epsilon > 0$ such that $\rho(t) \geq \epsilon > 0$ almost everywhere. Define the function $\chi_\xi \in L^\infty(X)$ by

$$\chi_\xi(t) = \sum_{i=1}^k a_i(t) \xi_i. \quad (2.5)$$

Then $\int_X \chi_\xi d\mu = 0$ and $\int_X a(t) \chi_\xi(t) d\mu(t) = \xi$ due to the orthogonality of $(a_i)_{i=1}^k$ and the constant function. Note that $|\chi_\xi(t)| = |\xi \cdot a(t)| \leq |\xi| \operatorname{ess\,sup}_{s \in X} |a(s)|$. Hence if

$$|\xi| < \frac{\epsilon}{\operatorname{ess\,sup}_{s \in X} |a(s)|}, \quad (2.6)$$

then $\rho_\xi = \rho + \chi_\xi$ is non-negative almost everywhere. Combining the previous results, this gives that $\rho_\xi \in \mathcal{P}(X)$, and $\int_X a(t) \rho_\xi(t) d\mu(t) = b + \xi$, so that \mathcal{Q} is open. \square

Using these properties it is now possible to characterise \mathcal{Q} in terms of its supporting hyperplanes, and this will provide the necessary and sufficient condition for $b \in \mathcal{Q}$.

Theorem 2.1. *Let $b \in \mathbb{R}^k$. Then $b \in \mathcal{Q}$ if and only if*

$$b \cdot u < \operatorname{ess\,sup}_{t \in X} a(t) \cdot u = S_u \quad (2.7)$$

for all $u \in \mathbb{S}^{k-1}$.

Proof. By Ref. 26 it is known that a closed convex set $K \subset \mathbb{R}^k$ can be written as the intersection of all closed half spaces containing K , so that $x \in K$ is equivalent to

$$x \cdot u \leq \sup_{z \in K} z \cdot u \quad (2.8)$$

for all $u \in \mathbb{S}^{k-1}$. If $\Omega \subset \mathbb{R}^k$ is a bounded open convex set, then by considering $K = \bar{\Omega}$ it is straightforward to deduce that $x \in \Omega$ is equivalent to

$$x \cdot u < \sup_{z \in K} z \cdot u = \sup_{z \in \Omega} z \cdot u \quad (2.9)$$

for all $u \in \mathbb{S}^{k-1}$, since a non-zero linear function cannot attain its maximum over a set in the interior. By taking $\Omega = \mathcal{Q}$, the supremum can be computed as

$$\begin{aligned} \sup_{z \in \mathcal{Q}} z \cdot u &= \sup_{\rho \in \mathcal{P}(X)} \int_X u \cdot a(t) \rho(t) d\mu(t) \\ &= \operatorname{ess\,sup}_{t \in X} u \cdot a(t) \\ &= S_u, \end{aligned} \quad (2.10)$$

by considering ρ uniform on the set $E_\epsilon^u = \{t \in X : \operatorname{ess\,sup} u \cdot a(s) < u \cdot a(t) + \epsilon\}$ and zero elsewhere, then taking $\epsilon \rightarrow 0$, which gives the necessary and sufficient condition. \square

Using the necessary and sufficient condition given previously it is possible to formalise the intuition that if b is close to the boundary of \mathcal{Q} , any probability distribution ρ generating b must be concentrated on some subset of X with small measure. Firstly, it will be shown that if $b \in \mathcal{Q}$ is close to the boundary of \mathcal{Q} , then there is a certain set, depending on b , on which any $\rho \in \mathcal{P}(X)$ with $\int_X a \rho d\mu = b$ must have most of its mass. Then secondly, it will be shown that as b approaches the boundary of \mathcal{Q} , this set becomes arbitrarily small in measure. These results will also be important in providing a growth bound for the singular potential in the next section.

Lemma 2.1. *Let $b \in \mathcal{Q}$, $b_0 \in \partial\mathcal{Q}$ with $|b_0 - b| = \min_{\tilde{b} \in \partial\mathcal{Q}} |\tilde{b} - b|$. Define $u \in \mathbb{S}^{k-1}$ to be the unit vector in the direction of $b_0 - b$. Then u is normal to a supporting hyperplane*

of \mathcal{Q} at b_0 , so that $b_0 \cdot u = S_u$, and $d(b, \partial\mathcal{Q}) = S_u - b \cdot u$. Furthermore, there is only one supporting hyperplane of \mathcal{Q} at b_0 , with outward-pointing unit normal vector given by u .

Proof. First it will be shown that u is normal to a supporting hyperplane at b_0 . Assume otherwise for the sake of contradiction. Then there exists some $u' \in \mathbb{S}^2$, normal to a supporting hyperplane at b_0 , which must satisfy $u \neq u'$. Let $x \in \mathbb{R}^k$ be the projection of b onto the hyperplane given by the set of all $\tilde{b} \in \mathbb{R}^k$ with $\tilde{b} \cdot u' = S_{u'}$. In particular, $x \notin \mathcal{Q}$. Furthermore, $|x - b| < |b_0 - b|$. By considering the straight line from b to x , there must be some $y \in \partial\mathcal{Q}$ such that $|y - b| \leq |x - b| < |b_0 - b|$, contradicting the assumption that $|b_0 - b| = \min_{\tilde{b} \in \partial\mathcal{Q}} |\tilde{b} - b|$. Then it holds that

$$\begin{aligned} S_u - b \cdot u &= (b_0 - b) \cdot u \\ &= \frac{1}{|b_0 - b|} (b_0 - b) \cdot (b_0 - b) \\ &= |b_0 - b| = d(b, \partial\mathcal{Q}). \end{aligned} \tag{2.11}$$

To show that this is the only such hyperplane, assume that $u_1 \in \mathbb{S}^{k-1}$, $u_1 \neq u$ and $u_1 \cdot b_0 = S_{u_1}$. By the same argument as before, by projecting b onto the supporting hyperplane with normal u_1 , it contradicts that b_0 was a closest point on the boundary. \square

Proposition 2.5. *Let $b \in \mathcal{Q}$, $b_0 \in \partial\mathcal{Q}$ with*

$$|b - b_0| = \min_{\tilde{b} \in \partial\mathcal{Q}} |b - \tilde{b}|, \tag{2.12}$$

define $u = \frac{1}{|b_0 - b|} (b_0 - b)$ and $\epsilon^2 = |b_0 - b|$. Let S_u , E_ϵ^u be as in Def. 2.4. Then $S_u - b \cdot u = \epsilon^2$ and

$$\int_{E_\epsilon^u} \rho(t) d\mu(t) \geq 1 - \epsilon \tag{2.13}$$

for all $\rho \in \mathcal{P}(X)$ generating b .

Proof. To see that $S_u - b \cdot u = \epsilon^2$ is a result of the previous lemma, and observing that $S_u - b \cdot u = (b_0 - b) \cdot u$. Letting $E = E_\epsilon^u$ for brevity, from the assumptions it can be seen that for any $\rho \in \mathcal{P}(\mathbb{S}^2)$ that generates b ,

$$\begin{aligned} 0 &= u \cdot b - S_u + \epsilon^2 \\ &= \int_X (u \cdot a(t) - S_u + \epsilon^2) \rho(t) d\mu(t) \\ &= \int_E (u \cdot a(t) - S_u + \epsilon^2) \rho(t) d\mu(t) + \int_{X \setminus E} (u \cdot a(t) - S_u + \epsilon^2) \rho(t) d\mu(t) \\ &\leq \epsilon^2 \int_E \rho d\mu + (\epsilon^2 - \epsilon) \int_{X \setminus E} \rho d\mu, \end{aligned} \tag{2.14}$$

using that for $t \in X$, $u \cdot a(t) - S_u \leq 0$, and for $t \in X \setminus E$ that $S_u - u \cdot a(t) \geq \epsilon$. If $I = \int_E \rho d\mu$ for brevity, then this chain of inequalities gives that

$$0 \leq \epsilon^2 I + (\epsilon^2 - \epsilon)(1 - I) \quad (2.15)$$

so that dividing through by ϵ and rearranging gives the desired result. \square \square

Proposition 2.6. *For $u \in \mathbb{S}^{k-1}$ and $\epsilon > 0$, let $E_\epsilon^u \subset X$ be as in Def. 2.4. Then $\lim_{\epsilon \rightarrow 0} \mu(E_\epsilon^u) = 0$, where the convergence is uniform in u .*

Proof. That $\lim_{\epsilon \rightarrow 0} \mu(E_\epsilon^u) = 0$ is a consequence of the pseudo-Haar property of the constraint functions. Since $E_\epsilon^u = (S_u - a \cdot u)^{-1}((-\infty, \epsilon))$, it holds that

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \mu(E_\epsilon^u) &= \mu \left(\bigcap_{\epsilon > 0} E_\epsilon^u \right) \\ &= \mu(\{t \in X : S_u - u \cdot a(t) = 0\}) = 0 \end{aligned} \quad (2.16)$$

by the pseudo-Haar property.

To see that the convergence is uniform, assume otherwise so that there exists some $\gamma > 0$, and $(u_j)_{j \in \mathbb{N}}$, $(\alpha_j)_{j \in \mathbb{N}}$ so that the sets $E_j = E_{\alpha_j}^{u_j}$ satisfy $\mu(E_j) \geq \gamma$ for all j , and $\lim_{j \rightarrow \infty} \alpha_j = 0$. For brevity, denote the functions $f_j(t) = S_{u_j} - u_j \cdot a(t) - \alpha_j$. Up to a subsequence (not relabelled) it can be assumed that $u_j \rightarrow u^*$, and since the convergence of $u_j \cdot a(t) \rightarrow u^* \cdot a(t)$ is uniform, there is a corresponding $S^* = S_{u^*} = \lim_{j \rightarrow \infty} S_{u_j}$. Let $\epsilon > 0$. Let $f = \lim_{j \rightarrow \infty} f_j$. Then for sufficiently large j since $f_j \rightarrow f$ uniformly, it must hold that if $t \in E_j$ then $f(t) > -\epsilon$. This means that $E_j \subset E_\epsilon^{u^*}$ for sufficiently large j . Therefore for large j ,

$$\begin{aligned} \mu(E_j) &\leq \mu(E_\epsilon^{u^*}) \\ \Rightarrow \limsup_{j \rightarrow \infty} \mu(E_j) &\leq \mu(E_\epsilon^{u^*}). \end{aligned} \quad (2.17)$$

Since ϵ was arbitrary and $\lim_{\epsilon \rightarrow 0} \mu(E_\epsilon^{u^*}) = 0$, this leads to the conclusion that $0 < \gamma \leq \limsup_{j \rightarrow \infty} \mu(E_j) \leq 0$, a contradiction. \square \square

3. The Singular Potential

In order to define the singular potential, there are certain constraints on the objective function ϕ that are necessary to ensure that the minimisation problem is well posed and the resulting singular potential has desirable properties. Motivated by the Shannon entropy, $\phi(x) = x \ln x$ as the prototypical example, any function satisfying these constraints will be called *entropy-like*.

Definition 3.1. Let $\phi : \mathbb{R} \rightarrow (-\infty, \infty]$. Define $\text{dom}(\phi) = \{x \in \mathbb{R} : \phi(x) < +\infty\}$. Then ϕ is entropy-like if and only if $\text{int dom}(\phi) = (0, \infty)$, $\phi|_{(0, \infty)}$ is strictly convex, continuously differentiable, $\lim_{x \rightarrow \infty} \frac{\phi(x)}{x} = +\infty$ and $\lim_{x \rightarrow 0^+} \phi'(x) = -\infty$.

In particular any entropy-like function is of Legendre type, so that ϕ' is a continuous bijection between $(0, \infty)$ and \mathbb{R} ²⁶. This definition does not place any restriction on the limiting value of ϕ at zero. It may hold that $\lim_{x \rightarrow 0^+} \phi(x)$ is bounded, such as in the case of the Shannon entropy, or the limit may be infinite, such as for the example $\phi(x) = \frac{1}{x} + x^2$.

Definition 3.2. Let ϕ be an entropy-like function. Define the singular potential $\psi_s : \mathcal{Q} \rightarrow \mathbb{R}$ by

$$\psi_s(b) = \min_{\rho \in \mathcal{A}_b} \int_X \phi(\rho) d\mu \tag{3.1}$$

where the set \mathcal{A}_b is defined as

$$\mathcal{A}_b = \left\{ \rho \in \mathcal{P}(X) : \int_X \rho(t)a(t) d\mu(t) = b \right\}. \tag{3.2}$$

Using the definition of the singular potential, the next results will be concerned with establishing some key properties of the function.

Proposition 3.1. *Let $b \in \mathcal{Q}$. Then the singular potential, $\psi_s(b)$ is well defined in the sense that the minimisation problem admits solutions. Furthermore, the minimiser is unique, and ψ_s is a strictly convex function.*

Proof. The minimisation problem is a convex minimisation problem subject to linear constraints, with constraint functions that are pseudo-Haar. Together with the growth conditions imposed on ϕ , this ensures the existence of a unique minimising ρ , with ρ bounded away from 0 and $+\infty$, so that $\psi_s(b) < +\infty$ ⁷. To see convexity, first since \mathcal{Q} is convex there is no problem in taking convex combinations of elements of \mathcal{Q} . Let $b_1, b_2 \in \mathcal{Q}$, $\epsilon \in [0, 1]$. If ρ_1, ρ_2 solve the minimisation problems corresponding to b_1, b_2 respectively, then $\epsilon\rho_1 + (1 - \epsilon)\rho_2 \in \mathcal{A}_{\epsilon b_1 + (1 - \epsilon)b_2}$. Hence

$$\begin{aligned} \psi_s(\epsilon b_1 + (1 - \epsilon)b_2) &= \min_{\rho \in \mathcal{A}_{\epsilon b_1 + (1 - \epsilon)b_2}} \int_X \phi(\rho(t)) d\mu(t) \\ &\leq \int_X \phi(\epsilon\rho_1(t) + (1 - \epsilon)\rho_2(t)) d\mu(t) \\ &\leq \int_X \epsilon\phi(\rho_1(t)) + (1 - \epsilon)\phi(\rho_2(t)) d\mu(t) \\ &= \epsilon\psi_s(b_1) + (1 - \epsilon)\psi_s(b_2). \end{aligned} \tag{3.3}$$

Furthermore, note that the strict convexity of ϕ gives that for $b_1 \neq b_2$ and $\epsilon \in (0, 1)$, $\rho_1 \neq \rho_2$ on a set of positive measure, giving that this inequality is strict. \square \square

Theorem 3.1. *There exists some $c > 0$ such that for all $b \in \mathcal{Q}$, $b_0 \in \partial\mathcal{Q}$ with $\min_{\tilde{b} \in \partial\mathcal{Q}} |\tilde{b} - b| = |b_0 - b| = \epsilon^2$ and $\epsilon < c$, the singular potential satisfies the inequality*

$$\psi_s(b) \geq \mu(E_\epsilon^u)\phi\left(\frac{1 - \epsilon}{\mu(E_\epsilon^u)}\right) + \mu(X \setminus E_\epsilon^u)\phi\left(\frac{\epsilon}{\mu(X \setminus E_\epsilon^u)}\right), \tag{3.4}$$

where E_ϵ^u is as defined in Def. 2.4 and $u = \frac{1}{|b_0 - b|}(b_0 - b)$.

Proof. Let $b \in \mathcal{Q}$, with corresponding b_0, ϵ, u . Let $E = E_\epsilon^u$ for brevity. From Prop. 2.5 it is known that $\int_E \rho d\mu > 1 - \epsilon$ for all ρ with $\int_X \rho a d\mu = b$. In particular by taking $\rho \in \mathcal{A}_b$ to be the unique minimiser given in Prop. 3.1 so that $\psi_s(b) = \int_X \phi(\rho) d\mu$, then

$$\begin{aligned} \psi_s(b) &= \int_E \phi(\rho) d\mu + \int_{X \setminus E} \phi(\rho) d\mu \\ &\geq \mu(E) \phi\left(\frac{1}{\mu(E)} \int_E \rho d\mu\right) + \mu(X \setminus E) \phi\left(\frac{1}{\mu(X \setminus E)} \int_{X \setminus E} \rho d\mu\right) \end{aligned} \quad (3.5)$$

by applying Jensen's inequality. Since $\frac{1}{\mu(E)} \int_E \rho d\mu \geq \frac{1 - \epsilon}{\mu(E)}$, using the uniform convergence of $\mu(E) \rightarrow 0$ in u from Prop. 2.6 and that ϕ is an increasing function for sufficiently large argument, it must hold that there exists some $c_1 > 0$ so that if $\epsilon < c_1$ then

$$\phi\left(\frac{1}{\mu(E)} \int_E \rho d\mu\right) \geq \phi\left(\frac{1 - \epsilon}{\mu(E)}\right). \quad (3.6)$$

Similarly for $\epsilon < c_2$ since ϕ is decreasing for sufficiently small argument, it must hold that

$$\phi\left(\frac{1}{\mu(X \setminus E)} \int_{X \setminus E} \rho d\mu\right) \geq \phi\left(\frac{\epsilon}{\mu(X \setminus E)}\right). \quad (3.7)$$

Combining these inequalities gives the lower bound. \square

Corollary 3.1. *For $j \in \mathbb{N}$ let $b_j \in \mathcal{Q}$ with $\lim_{j \rightarrow \infty} \inf_{\tilde{b} \in \partial \mathcal{Q}} |\tilde{b} - b_j| = 0$. Then $\lim_{j \rightarrow \infty} \psi_s(b_j) = +\infty$. Furthermore the convergence is uniform in $\inf_{\tilde{b} \in \partial \mathcal{Q}} |b - b_j|$, so that*

$$\lim_{\epsilon \rightarrow 0} (\inf \{\psi_s(b) : d(b, \partial \mathcal{Q}) \leq \epsilon\}) = +\infty. \quad (3.8)$$

Proof. Let $b \in \mathcal{Q}$, and $E_\epsilon^u = E$ for brevity. Using the lower bound for ψ_s , and that ϕ is bounded below, for some real constant C , independent of b ,

$$\psi_s(b) \geq \mu(E) \phi\left(\frac{1 - \epsilon}{\mu(E)}\right) + C = (1 - \epsilon) \frac{\phi(\xi)}{\xi} + C, \quad (3.9)$$

where $\xi = \frac{1 - \epsilon}{\mu(E)}$. Using that $\mu(E) \rightarrow 0$ uniformly as $b \rightarrow \partial \mathcal{Q}$ and the superlinear growth of ϕ gives the required result. \square

The precise form of the minimiser can be obtained by considering the dual optimisation problem. For a more thorough account of duality the reader is referred to Ref. 26, although for completeness a brief heuristic argument will be provided here.

Let X be some Banach space, $F : X \rightarrow \mathbb{R}$ be convex, and $T : X \rightarrow \mathbb{R}^k$ be continuous and linear. Let $b \in \mathbb{R}^k$ and consider the minimisation problem

$$\begin{aligned} & \min_{x \in X} F(x) \\ & \text{subject to } Tx = b. \end{aligned} \tag{3.10}$$

Trivially, if x does not satisfy $Tx = b$, then $\max_{\lambda \in \mathbb{R}^k} (-\lambda \cdot (Tx - b)) = +\infty$, and if $Tx = b$, then $(-\lambda \cdot (Tx - b)) = 0$ for all $\lambda \in \mathbb{R}^k$. Therefore the minimisation problem can be written as

$$\min_{x \in X} \max_{\lambda \in \mathbb{R}^k} F(x) - \lambda \cdot (Tx - b). \tag{3.11}$$

Let X^* be the dual space to X with duality pairing $\langle \cdot, \cdot \rangle$, $T^* : \mathbb{R}^k \rightarrow X^*$ be the adjoint of T , and $F^* : X^* \rightarrow \mathbb{R}$ be the convex conjugate of F defined by $F^*(x^*) = \sup_{x \in X} \langle x^*, x \rangle - F(x)$. Assuming that the minimum and maximum in (3.11) can be exchanged, this minimisation problem can be written as

$$\begin{aligned} \max_{\lambda \in \mathbb{R}^k} \min_{x \in X} F(x) - \lambda \cdot (Tx - b) &= \max_{\lambda \in \mathbb{R}^k} \lambda \cdot b - \left(\max_{x \in X} \langle T^* \lambda, x \rangle - F(x) \right) \\ &= \max_{\lambda \in \mathbb{R}^k} \lambda \cdot b - F^*(T^* \lambda). \end{aligned} \tag{3.12}$$

If the interchange of minimum and maximum is permitted, then the dual problem

$$\max_{\lambda \in \mathbb{R}^k} \lambda \cdot b - F^*(T^* \lambda) \tag{3.13}$$

can provide information about the original so-called primal problem. Furthermore, since the minimisation is over a finite dimensional set without constraints, it may be easier to approach. This is what is done in the work of Borwein and Lewis⁷, and their results will be exploited for the sake of this work. In particular the minimiser of the primal problem can be described in terms of maximiser of the dual problem.

Proposition 3.2. *Let $\phi : \mathbb{R} \rightarrow (-\infty, \infty]$ be an entropy-like function. Then for $b \in \mathcal{Q}$, the unique minimiser of*

$$\min_{\rho \in \mathcal{A}_b} \int_X \phi(\rho) d\mu \tag{3.14}$$

is achieved by

$$\rho_b(t) = (\phi')^{-1}(\alpha + \lambda \cdot a(t)), \tag{3.15}$$

where $\alpha \in \mathbb{R}$, $\lambda \in \mathbb{R}^k$ are the unique maximisers of the dual optimisation problem,

$$\max_{\tilde{\alpha}, \tilde{\lambda}} \tilde{\alpha} + \tilde{\lambda} \cdot b - \int_X \phi^* \left(\tilde{\alpha} + \tilde{\lambda} \cdot a(t) \right) d\mu. \tag{3.16}$$

Proof. This is a straightforward application of Ref. 7, \square

The next stage is to establish smoothness properties of the singular potential. In order to show differentiability of ψ_s , the key point is to observe that the map from the Lagrange multipliers to the moments b is explicitly given and invertible. As such the differentiability of ψ_s is inherited from the explicit map from the Lagrange multipliers to \mathcal{Q} .

Proposition 3.3. *Let $\phi : \mathbb{R} \rightarrow (-\infty, \infty]$ be entropy-like, C^m on $(0, \infty)$ for $m \geq 2$, with ϕ'' positive everywhere. Then the map $b \mapsto (\alpha, \lambda)$ is C^{m-1} , where α, λ solve the dual optimisation problem.*

Proof. For $(\alpha, \lambda) \in \mathbb{R} \times \mathbb{R}^k$, consider the map

$$h : (\alpha, \lambda) \mapsto \int_X (1, a(t)) (\phi')^{-1} (\alpha + \lambda \cdot a(t)) d\mu(t) = (\beta, b). \quad (3.17)$$

Since ϕ'' is positive, the inverse function theorem can be applied to $(\phi')^{-1}$, and the differentiability of ϕ gives that $(\phi')^{-1}$ is C^{m-1} on \mathbb{R} . In particular, this implies that $h \in C^{m-1}(\mathbb{R}^{k+1}, \mathbb{R}^{k+1})$. The Jacobian matrix of h is given by

$$\frac{\partial h(\alpha, \lambda)}{\partial(\alpha, \lambda)} = \int_X (1, a(t)) \otimes (1, a(t)) \left(\phi'' \circ (\phi')^{-1} (\alpha + \lambda \cdot a(t)) \right)^{-1} d\mu(t). \quad (3.18)$$

Here \circ denotes function composition. If this Jacobian matrix failed to be invertible, there would exist some $\xi \in \mathbb{R}^{k+1}$ with $\xi \neq 0$ so that

$$0 = \int_X ((1, a(t)) \cdot \xi)^2 \left(\phi'' \circ (\phi')^{-1} (\alpha + \lambda \cdot a(t)) \right)^{-1} d\mu(t). \quad (3.19)$$

Since ϕ'' is positive everywhere, this implies that $(\xi \cdot (1, a(t)))^2 = 0$ almost everywhere in X ; however by the pseudo-Haar property this cannot hold, a contradiction. By applying the inverse function theorem to h , this gives that the map

$$(\beta, b) \mapsto (\alpha, \lambda) \quad (3.20)$$

is C^{m-1} for $(\beta, b) \in h(\mathbb{R}^{k+1})$. In particular, $h^{-1}(1, b)$, which maps $b \in \mathcal{Q}$ to (α, λ) is also a C^{m-1} function. Global invertibility is not problematic, since the existence and uniqueness of solutions for dual problem ensures the existence and unique pair (α, λ) for each $b \in \mathcal{Q}$, and non-existence otherwise. \square

Proposition 3.4. *Let ϕ be a C^2 entropy-like function with ϕ'' positive. Then for $b \in \mathcal{Q}$, $\frac{\partial \psi_s}{\partial b}(b) = \lambda$, where λ solves the dual optimisation problem corresponding to b .*

Proof. From Ref. 7 it is possible to write the singular potential in terms of the Lagrangian dual optimisation problem, so that for dual optimal α, λ , the singular potential can be written as

$$\psi_s(b) = \alpha + \lambda \cdot b - \int_X \phi^*(\alpha + \lambda \cdot a(t)) d\mu(t) \quad (3.21)$$

Using from Prop. 3.3 that the map $b \mapsto (\alpha, \lambda)$ is at least C^1 , differentiating through gives that

$$\begin{aligned} \frac{\partial \psi_s}{\partial b} &= \frac{\partial \alpha}{\partial b} + \frac{\partial \lambda}{\partial b} \cdot b + \lambda - \int_X (\phi')^{-1} (\alpha + \lambda \cdot a(t)) \left(\frac{\partial \alpha}{\partial b} + \frac{\partial \lambda}{\partial b} \cdot a(t) \right) d\mu(t) \\ &= \lambda + \frac{\partial \alpha}{\partial b} \left(1 - \int_X \rho_b(t) d\mu(t) \right) + \frac{\partial \lambda}{\partial b} \cdot \left(b - \int_X \rho_b(t) a(t) d\mu(t) \right) \\ &= \lambda. \end{aligned} \quad (3.22)$$

□

□

Corollary 3.2. *If ϕ is entropy-like, C^m for $m \geq 2$ on $(0, \infty)$ and ϕ'' is positive, then $\psi_s \in C^m(\mathcal{Q}, \mathbb{R})$.*

Proof. Combining the results of Prop. 3.3 and 3.4, it holds that

$$\frac{\partial \psi_s}{\partial b} = \lambda \in C^{m-1}(\mathcal{Q}; \mathbb{R}^k), \quad (3.23)$$

so that $\psi_s \in C^m(\mathcal{Q}, \mathbb{R})$. □

□

Proposition 3.5. *Assume ϕ is entropy-like, C^m for $m \geq 2$ on the interior of its domain and that ϕ'' is positive. Let $b_0 \in \mathcal{Q}$, and denote $\rho = \rho_{b_0}$, $\sigma = \frac{1}{\phi''(\rho)}$, $Z = \int_X \sigma d\mu$. and $\tilde{\sigma} = \frac{1}{2}\sigma$. Then*

$$\frac{\partial \lambda}{\partial b}(b_0) = \frac{1}{Z} \left(\int_X a \otimes a \tilde{\sigma} d\mu - \int_X a \tilde{\sigma} d\mu \otimes \int_X a \tilde{\sigma} d\mu \right)^{-1}. \quad (3.24)$$

In particular, if ϕ is the Shannon entropy $\phi(x) = x \ln x$, then

$$\frac{\partial \lambda}{\partial b}(b_0) = \left(\int_X a \otimes a \rho d\mu - b_0 \otimes b_0 \right)^{-1}. \quad (3.25)$$

Proof. Under the standing assumptions α, λ as functions of b are at least C^1 . Following a similar argument to Prop. 3.3,

$$\begin{aligned} 1 &= \int_X (\phi')^{-1} (\alpha + \lambda \cdot a) d\mu \\ \Rightarrow 0 &= \int_X \left(\frac{\partial \alpha}{\partial b_i} + \frac{\partial \lambda}{\partial b_i} \cdot a \right) \sigma d\mu \\ &= Z \frac{\partial \alpha}{\partial b_i} + \int_X a \sigma d\mu \cdot \frac{\partial \lambda}{\partial b_i}. \end{aligned} \quad (3.26)$$

Similarly, considering the relation between α, λ and b gives

$$\begin{aligned}
b_i &= \int_X a_i (\phi')^{-1} (\alpha + \lambda \cdot a) d\mu \\
\Rightarrow \delta_{ij} &= \int_X \left(a_i \frac{\partial \alpha}{\partial b_j} + a_i \frac{\partial \lambda}{\partial b_j} \cdot a \right) \sigma d\mu \\
&= -\frac{1}{Z} \int_X a_i \sigma d\mu \int_X a \sigma d\mu \cdot \frac{\partial \lambda}{\partial b_j} + \int_X a_i a \sigma d\mu \cdot \frac{\partial \lambda}{\partial b_j} \\
\Rightarrow I &= Z \left(\int_X a \otimes a \tilde{\sigma} d\mu - \int_X a \tilde{\sigma} d\mu \otimes \int_X a \tilde{\sigma} d\mu \right) \frac{\partial \lambda}{\partial b}.
\end{aligned} \tag{3.27}$$

Since ϕ'' , and therefore σ is positive, and $\tilde{\sigma} = \frac{1}{\int_X \sigma d\mu} \sigma$, this gives that $\tilde{\sigma} \in \mathcal{P}(X)$, and furthermore due to the bounds on ρ , $\tilde{\sigma}$ is bounded away from zero. Therefore by use of the pseudo-Haar condition analogously to in Prop. 3.3, the matrix

$$Z \left(\int_X a \otimes a \tilde{\sigma} d\mu - \int_X a \tilde{\sigma} d\mu \otimes \int_X a \tilde{\sigma} d\mu \right) \tag{3.28}$$

is invertible, so that taking its inverse gives the result. To see the case when ϕ is the Shannon entropy follows immediately since $\phi''(x) = \frac{1}{x}$, which gives $\sigma = \rho$, so $Z = 1$ and $\tilde{\sigma} = \rho$. \square

Proposition 3.6. *Let $\phi : \mathbb{R} \rightarrow (-\infty, \infty]$ be an entropy-like function, C^2 with ϕ'' positive on $(0, \infty)$. Then the map $F : \mathcal{Q} \rightarrow \mathcal{P}(X)$ given by $F(b) = \rho_b$ is continuous with respect to the L^∞ topology on $\mathcal{P}(X)$.*

Proof. From the argument in Prop. 3.3 it is sufficient to show that the map $G : (\alpha, \lambda) \mapsto (\phi')^{-1} (\alpha + \lambda \cdot a)$ is continuous as a function from \mathbb{R}^{k+1} to L^∞ . That $G(\alpha, \lambda) \in L^\infty$ is a consequence of $(\phi')^{-1}$ being continuous and the inequality $|\alpha + \lambda \cdot a(t)| < |\alpha| + |\lambda| \|a\|_\infty$. To show continuity, fix α_0, λ_0 . Define $K = |\alpha_0| + |\lambda_0| \|a\|_\infty$. Consider α, λ with $|\alpha - \alpha_0| + |\lambda - \lambda_0| \|a\|_\infty < \epsilon$ for $\epsilon > 0$. Then

$$\begin{aligned}
|\alpha + \lambda \cdot a(t)| &= |\alpha - \alpha_0 + (\lambda - \lambda_0) \cdot a(t) + \alpha_0 + \lambda_0 \cdot a(t)| \\
&\leq \epsilon + K.
\end{aligned} \tag{3.29}$$

Since $(\phi')^{-1}$ is C^1 on \mathbb{R} , it is Lipschitz on $[-(K + \epsilon), K + \epsilon]$ with Lipschitz constant L , say. Hence

$$\begin{aligned}
\left| (\phi')^{-1} (\alpha + \lambda \cdot a(t)) - (\phi')^{-1} (\alpha_0 + \lambda_0 \cdot a(t)) \right| &\leq L |\alpha - \alpha_0 + (\lambda - \lambda_0) \cdot a(t)| \\
&\leq L (|\alpha - \alpha_0| + |\lambda - \lambda_0| \|a\|_\infty) \\
&\leq L\epsilon.
\end{aligned} \tag{3.30}$$

\square

\square

4. Applications

For the remainder of this section, ϕ will be assumed to be entropy-like and C^2 with positive second derivative on $(0, \infty)$. In particular from Corollary 3.2 this implies ψ_s is C^2 . Given $b \in \mathcal{Q}$, ρ_b will be as in Proposition 3.2, and given $\rho \in \mathcal{P}(X)$, define $b_\rho = \int_X a\rho d\mu$. Within this section all the examples of constraint functions can be identified with linearly independent sets of analytic functions on connected subsets of \mathbb{R}^n , with corresponding measure absolutely continuous with respect to the Lebesgue measure, so that by ¹⁹, the pseudo-Haar condition is satisfied.

4.1. Mean Field and Nonlinear Constraint Models

The aim of this section is to establish some elementary results on minimisation problems for certain functionals on $\mathcal{P}(X)$ by reducing the problem to a finite dimensional problem via the machinery of the singular potential. The key point is to establish equivalence between global minima, local minima and critical points of functionals on $\mathcal{P}(X)$ and global minima, local minima and critical points (respectively) of functions on \mathcal{Q} . Once this has been shown it is possible to establish standard results such as existence of minimisers by using simple tools of real analysis rather than having to resort to more complicated arguments based on functional analysis. Critical points of the functionals on $\mathcal{P}(X)$ will be expressed as implicit relations on the dual variables. This loosely provides a framework for dealing with such minimisation problems, whereby the model is derived as a problem in $\mathcal{P}(X)$ by some physical argument, then the singular potential allows analytical questions to be asked in the simpler language of real analysis as a problem in \mathcal{Q} , and then finally the dual variables provide a framework that is attractive from the point of view of numerical analysis.

Let $\emptyset \neq A \subset \mathcal{Q}$ be relatively closed in \mathcal{Q} . Let $f : \mathcal{Q} \rightarrow \mathbb{R}$ be continuous on A and bounded from below, with $f|_{\mathcal{Q} \setminus A} = +\infty$. Define the functionals $\mathcal{I}_{\mathcal{P}} : \mathcal{P}(X) \rightarrow \mathbb{R}$ and $\mathcal{I}_{\mathcal{Q}} : \mathcal{Q} \rightarrow \mathbb{R}$ by

$$\begin{aligned} \mathcal{I}_{\mathcal{P}}(\rho) &= \int_X \phi(\rho) d\mu + f(b_\rho), \\ \mathcal{I}_{\mathcal{Q}}(b) &= \psi_s(b) + f(b) = \mathcal{I}_{\mathcal{P}}(\rho_b). \end{aligned} \tag{4.1}$$

Consider the minimisation problems (P1) and (P2) given by

$$\begin{aligned} \text{(P1)} \quad & \min_{\rho \in \mathcal{P}(X)} \mathcal{I}_{\mathcal{P}}(\rho) \\ \text{(P2)} \quad & \min_{b \in \mathcal{Q}} \mathcal{I}_{\mathcal{Q}}(b). \end{aligned} \tag{4.2}$$

The aim of this section is to show equivalence of global minima, local minima and critical points for (P1) and (P2) and characterise such points. For applications (P1) will be obtained through a physical argument, and (P2) loosely speaking is a simpler, macroscopic, equivalent model. The motivation for the problem (P1) comes from two different areas. Firstly one can see maximum entropy methods subject to

nonlinear constraints as an example. Let $g : \mathcal{Q} \rightarrow \mathbb{R}$ be a continuous function. Consider the minimisation problem

$$\begin{aligned} & \min_{\rho \in \mathcal{P}(X)} \int_X \phi(\rho) d\mu \\ & \text{subject to } g(b_\rho) = 0 \end{aligned} \quad (4.3)$$

Define $f(b) = 0$ if $g(b) = 0$ and $f(b) = +\infty$ if $g(b) \neq 0$ so that $A = \{b \in \mathcal{Q} : g(b) = 0\}$, which by the continuity of g is relatively closed. Then the minimisation problem

$$\min_{\rho \in \mathcal{P}(X)} \int_X \phi(\rho) d\mu + f(b_\rho)$$

is equivalent to the nonlinear constraint model. Examples from the literature of such models are seen in statistical models of isotropic elasticity³⁰, where $X = \mathbb{S}^2$, $a_i(p) = p_i$ (in Cartesian coordinates) for $p \in \mathbb{S}^2$ represents the orientation of molecules in a polymer chain and b corresponds to the end to end vector spanned by the entire chain. The constraint in this case is of the form $|b| = r$ for $r \in [0, 1)$. Maximum entropy methods with to nonlinear constraints were also considered by Decarreau *et al.*¹² as a method for dealing with the phase problem in crystallography. In diffraction experiments it is much simpler to observe the intensity of a wave than its phase, and mathematically this corresponds to having knowledge of the modulus of a Fourier coefficient representing the wave whilst its argument is unknown. For example, Decarreau *et al.* consider minimisation problems such as

$$\left\{ \begin{array}{l} \min_{\rho \in \mathcal{P}(\Omega)} \int_\Omega \phi(\rho(x)) dx \\ \left(\int_\Omega \cos(n \cdot x) \rho(x) dx \right)^2 + \left(\int_\Omega \sin(n \cdot x) \rho(x) dx \right)^2 = m_n^2 \quad \forall n \in N \end{array} \right\},$$

where $\Omega = [0, 2\pi]$, N is some finite subset of \mathbb{N} and m_n are given real numbers.

A second example of problems of the form (P1) are given by the mean field approximation. Given some state space X and constraint functions $(a_i)_{i=1}^k$, the free energy is typically of the form

$$T \int_X \phi(\rho) d\mu - \frac{1}{2} K b_\rho \cdot b_\rho - H \cdot b_\rho \quad (4.4)$$

where $T > 0$ represents temperature, K is a positive definite $k \times k$ matrix that represents some kind of interaction potential, and $H \in \mathbb{R}^k$ is representative of some kind of external influence such as magnetic/electric fields. In practice the Shannon entropy $\phi(x) = x \ln x$ is almost exclusively used, although for this analysis this is not necessary.

In the literature these minimisation problems often are not dealt with rigorously, and in particular the issue of non-differentiability of the objective function ϕ at 0 is often neglected, so that the given solution is obtained by considering only points where the first variation of the free energy is zero. The following results will aim to show that under rather non-restrictive assumptions the minimisation problems are well posed and the solutions behave as one would expect with a non-rigorous

analysis. Through use of the singular potential the minimisation problem can be reduced to an analytically simpler but equivalent problem (P2) in finite dimensions, for which one can easily obtain the desired results.

Proposition 4.1. *There exists global minimisers b^* and ρ^* of $\mathcal{I}_{\mathcal{Q}}$ and $\mathcal{I}_{\mathcal{P}}$ respectively, and all global minimisers of $\mathcal{I}_{\mathcal{Q}}$ are in one-to-one correspondence with global minimisers of $\mathcal{I}_{\mathcal{P}}$ via the map $b \mapsto \rho_b$. In particular, global minimisers of $\mathcal{I}_{\mathcal{P}}$ are bounded away from zero and infinity.*

Proof. From the assumptions on ϕ, f , it is immediate that these functionals are bounded below. Furthermore, their infima coincide, since

$$\begin{aligned} \inf_{\rho \in \mathcal{P}(X)} \mathcal{I}_{\mathcal{P}}(\rho) &= \inf_{b \in \mathcal{Q}} \left(\inf_{\rho \in \mathcal{A}_b} \mathcal{I}_{\mathcal{P}}(\rho) \right) \\ &= \inf_{b \in \mathcal{Q}} \mathcal{I}_{\mathcal{Q}}(b). \end{aligned} \quad (4.5)$$

Since \mathcal{Q} is a precompact set, $\mathcal{I}_{\mathcal{Q}}$ is continuous on the relatively closed set A , and $\lim_{b \rightarrow \partial \mathcal{Q}} \mathcal{I}_{\mathcal{Q}}(b) = +\infty$, it must hold that a minimum exists for $\mathcal{I}_{\mathcal{Q}}$, with corresponding minimiser b^* . Furthermore, since $\mathcal{I}_{\mathcal{P}}(\rho_{b^*}) = \mathcal{I}_{\mathcal{Q}}(b^*)$, the minimum of $\mathcal{I}_{\mathcal{P}}$ must be attained at ρ_{b^*} . Finally, if ρ^* is a global minimiser for $\mathcal{I}_{\mathcal{P}}$, then it must hold that $\rho^* = \rho_{b^*}$ for $b^* = b_{\rho^*}$, since otherwise $\mathcal{I}_{\mathcal{P}}(\rho^*) > \mathcal{I}_{\mathcal{Q}}(b^*) = \mathcal{I}_{\mathcal{P}}(\rho_{b^*})$. \square \square

Proposition 4.2. *There is a one-to-one correspondence between local minimisers of $\mathcal{I}_{\mathcal{Q}}$ and L^1 -local minimisers of $\mathcal{I}_{\mathcal{P}}$, given by the map $b \mapsto \rho_b$, so that $\rho^* \in \mathcal{P}(X)$ is an L^1 -local minimiser if and only if $b_{\rho^*} \in \mathcal{Q}$ is a local minimiser, and vice versa. In particular, all local minimisers of $\mathcal{I}_{\mathcal{P}}$ are bounded away from zero and infinity. The equivalence also holds for strict local minimisers.*

Proof. The proof for strict local minimisers is identical to that for non-strict minimisers, so the proof for strict local minimisers will be omitted. Let $b^* \in \mathcal{Q}$ be a local minimiser, so that $\mathcal{I}_{\mathcal{Q}}(b^*) \leq \mathcal{I}_{\mathcal{Q}}(b)$ for all $b \in \mathcal{Q}$ with $|b - b^*| < \epsilon$. Define $\rho^* = \rho_{b^*}$. Let $\rho \in \mathcal{P}(X)$ with $\|\rho - \rho^*\|_1 < \delta$. Then if $b = b_{\rho}$, $|b - b^*| < \delta \|a\|_{\infty}$. Finally,

$$\begin{aligned} \mathcal{I}_{\mathcal{P}}(\rho) - \mathcal{I}_{\mathcal{P}}(\rho^*) &= \int_X \phi(\rho) d\mu + f(b) - \mathcal{I}_{\mathcal{Q}}(b^*) \\ &\geq \psi_s(b) + f(b) - \mathcal{I}_{\mathcal{Q}}(b^*) \\ &= \mathcal{I}_{\mathcal{Q}}(b) - \mathcal{I}_{\mathcal{Q}}(b^*). \end{aligned} \quad (4.6)$$

Therefore if $\delta < \frac{\epsilon}{\|a\|_{\infty}}$, it holds that $|b - b^*| < \epsilon$ so that consequently $\mathcal{I}_{\mathcal{P}}(\rho) \geq \mathcal{I}_{\mathcal{P}(X)}(\rho^*)$.

To show the converse statement, note that if ρ^* is an L^1 -local minimiser of $\mathcal{I}_{\mathcal{P}}$, then $\rho^* = \rho_{b^*}$. This can be seen by considering $\rho = (1 - \gamma)\rho^* + \gamma\rho_{b^*}$ for $\gamma > 0$ small, and using the strict convexity of ϕ , which gives that $\mathcal{I}_{\mathcal{P}(X)}(\rho) < \mathcal{I}_{\mathcal{P}(X)}(\rho^*)$.

Now let ρ^* be an L^1 -local minimiser, so that for all $\rho \in \mathcal{P}(X)$ with $\|\rho - \rho^*\|_1 < \epsilon$, $\mathcal{I}_{\mathcal{P}}(\rho) \geq \mathcal{I}_{\mathcal{P}}(\rho^*)$. Now consider $b \in \mathcal{Q}$. Then

$$\mathcal{I}_{\mathcal{Q}}(b) - \mathcal{I}_{\mathcal{Q}}(b^*) = \mathcal{I}_{\mathcal{P}}(\rho_b) - \mathcal{I}_{\mathcal{P}}(\rho^*). \quad (4.7)$$

All that remains to show is that there exists some $\delta > 0$ so that for all $\tilde{b} \in \mathcal{Q}$ with $|\tilde{b} - b^*| < \delta$, then $\|\rho_{\tilde{b}} - \rho^*\|_1 < \epsilon$. This holds, since from Prop. 3.6 it is known that the map $b \mapsto \rho_b$ is continuous in L^∞ , and hence is continuous in L^1 . \square

For the following result, a critical point of a function $F : \mathcal{P}(X) \rightarrow \mathbb{R}$ is defined to be any ρ in $\mathcal{P}(X)$ that is bounded away from zero, such that for all $\xi \in L^\infty(X)$ with $\int_X \xi d\mu = 0$,

$$\left. \frac{d}{d\tau} F(\rho + \tau\xi) \right|_{\tau=0} = 0. \quad (4.8)$$

Proposition 4.3. *Let $\text{int}(A) \neq \emptyset$ and let f be C^1 on $\text{int}(A)$. Then $b^* \in \text{int}(A)$ is a critical point of $\mathcal{I}_{\mathcal{Q}}$ if and only if $\rho^* = \rho_{b^*}$ is a critical point of $\mathcal{I}_{\mathcal{P}}$ in $\{\rho \in \mathcal{P}(X) : b_\rho \in \text{int}(A)\}$.*

Proof. Since b^* is a critical point of the C^1 function $\mathcal{I}_{\mathcal{Q}}$, it holds that

$$\begin{aligned} 0 &= \nabla \mathcal{I}_{\mathcal{Q}}(b^*) \\ &= \lambda(b^*) + \nabla f(b^*). \end{aligned} \quad (4.9)$$

This implies that $\phi'(\rho^*) = \alpha - \nabla f(b^*) \cdot a$. Consider any $\xi \in L^\infty(X)$ with $\int_X \xi d\mu = 0$. This gives

$$\begin{aligned} \left. \frac{d}{d\tau} \mathcal{I}_{\mathcal{P}}(\rho^* + \tau\xi) \right|_{\tau=0} &= \int_X (\phi'(\rho^*) - \nabla f(b^*) \cdot a) \xi d\mu \\ &= \alpha \int_X \xi d\mu \\ &= 0. \end{aligned} \quad (4.10)$$

The converse follows by the same argument. \square

Remark 4.1. Whilst Proposition 4.2 shows that there is a one-to-one correspondence between local minima of $\mathcal{I}_{\mathcal{P}}$ and $\mathcal{I}_{\mathcal{Q}}$, an analogous result does not hold for local maxima. Let $b \in \mathcal{Q}$ be a local maximum for $\mathcal{I}_{\mathcal{Q}}$, and let $\xi \in L^\infty(X) \setminus \{0\}$ be such that $\int_X \xi d\mu = 0$ and $\int_X a\xi d\mu = 0$. Then $\rho_b + \tau\xi \in \mathcal{A}_b$ for $\tau \in \mathbb{R}$ sufficiently small to ensure the non-negativity constraint is satisfied, and hence

$$\begin{aligned} \mathcal{I}_{\mathcal{P}}(\rho_b + \tau\xi) &= \int_{\mathbb{S}^2} \phi(\rho_b + \tau\xi) + f(b) \\ &> \psi_s(b) + f(b) \\ &= \mathcal{I}_{\mathcal{Q}}(b) \\ &= \mathcal{I}_{\mathcal{P}}(\rho_b). \end{aligned} \quad (4.11)$$

By taking τ sufficiently small shows that ρ_b is not a local maximum, although if $f \in C^1$, ρ_b will be a critical point by Prop. 4.3.

Corollary 4.1. *Consider the minimisation problem (P1). If $A = \mathcal{Q}$ and $f \in C^1(\overline{\mathcal{Q}}; \mathbb{R})$ then there exists a global minimum, and all L^1 -local minima and critical points satisfy*

$$\rho^*(t) = (\phi')^{-1}(\alpha - \nabla f(b_\rho) \cdot a(t)) \quad (4.12)$$

for some constant $\alpha \in \mathbb{R}$.

Proof. This is a straightforward application of Propositions 4.1, 4.2 and 4.3. \square

Corollary 4.2. *Consider the minimisation problem (P1). Let $g \in C^1(\mathcal{Q}, \mathbb{R})$ and $A = \{b \in \mathcal{Q} : g(b) = 0\} \neq \emptyset$. Then there exists a global minimum, and all L^1 -local minima and critical points satisfy*

$$\rho^*(t) = (\phi')^{-1}(\alpha + \eta \nabla g(b) \cdot a(t)) \quad (4.13)$$

for all $t \in X$ and some Lagrange multiplier $\eta \in \mathbb{R}$.

Proof. Again this is a straightforward application of Propositions 4.1, 4.2 and 4.3. \square

Although the equivalent problem (P2) is simpler by virtue of being finite-dimensional, the non-explicit representation of the energy limits its usefulness. However, an order parameter that is in some sense dual to $b \in \mathcal{Q}$ can be considered, and reduces the dual optimisation scheme needed at each point in the domain to a simple 1-dimensional problem. First, a lemma is required.

Lemma 4.1. *Let $\lambda \in \mathbb{R}^k$. Then there exists a unique $\alpha_\lambda \in \mathbb{R}$ such that $\int_X (\phi')^{-1}(\alpha_\lambda + \lambda \cdot a(t)) d\mu(t) = 1$, which can be found by solving*

$$\max_{\alpha \in \mathbb{R}} \alpha - \int_X \phi^*(\alpha + \lambda \cdot a(t)) d\mu(t). \quad (4.14)$$

Furthermore, if ϕ is C^m on its effective domain for $m \geq 2$ then α is a C^{m-1} function of λ .

Proof. Existence of such an α_λ can be seen by recalling that $(\phi')^{-1}$ is an increasing bijection between \mathbb{R} and $(0, \infty)$ and the map $\alpha \mapsto \int_X (\phi')^{-1}(\alpha + \lambda \cdot a(t)) d\mu(t)$ is a continuous function of α by applying the intermediate value theorem. Uniqueness follows from the strict monotonicity of $(\phi')^{-1}$. The regularity follows from the implicit function theorem, using an argument analogous to Proposition 3.3. The existence of a maximiser for the problem in Equation (4.14) follows from the existence of a unique critical point and the strict concavity of the functional. \square

Remark 4.2. In the case of the Shannon entropy, the additive property of the exponential greatly simplifies the previous lemma, since α_λ can explicitly be given as

$$\alpha_\lambda = 1 - \ln \left(\int_X \exp(\lambda \cdot a(t)) d\mu(t) \right). \quad (4.15)$$

We now define $\lambda \in \mathbb{R}^k$ to be the dual order parameter. It corresponds to the classical order parameter through the relationship

$$b = \int_X (\phi')^{-1}(\alpha_\lambda + \lambda \cdot a(t)) a(t) d\mu(t). \quad (4.16)$$

If ϕ is C^m for $m \geq 2$, then it follows that the relationship between the dual and classical order parameters is a C^{m-1} bijection between \mathbb{R}^k and \mathcal{Q} . Furthermore, if $b^\lambda \in \mathcal{Q}$ is the classical order parameter corresponding to a dual order parameter λ , then it is immediate that

$$\mathcal{I}_{\mathcal{Q}}(b^\lambda) = \alpha_\lambda + \lambda \cdot b^\lambda + f(b^\lambda) = \mathcal{I}_{\mathbb{R}^k}(\lambda) \quad (4.17)$$

The map $\lambda \mapsto b^\lambda$ is a continuous open map, which implies that local minimisers of $\mathcal{I}_{\mathbb{R}^k}$ and $\mathcal{I}_{\mathcal{Q}}$ are in one-to-one correspondence also. The advantage of solving $\mathcal{I}_{\mathbb{R}^k}$ is that the domain is no longer a constrained set, and at each point one must only solve a 1-dimensional optimisation problem rather than a $(k+1)$ -dimensional one. In the case of the Shannon entropy, the explicit expression of α_λ as mentioned in Remark 4.2 requires no optimisation problem to be solved in order to evaluate the energy. In the case of the Onsager energy, the following proposition explicitly states these conclusions.

Proposition 4.4. *Let $\lambda \in \mathbb{R}^k$. Define*

$$\begin{aligned} Z_\lambda &= \int_X \exp(\lambda \cdot a(t)) d\mu(t), \\ b^\lambda &= \frac{1}{Z_\lambda} \int_X \exp(\lambda \cdot a(t)) a(t) d\mu(t). \end{aligned} \quad (4.18)$$

Then all global minimisers, L^1 -local minimisers and critical points (respectively) of Onsager's free energy

$$\mathcal{I}_{\mathcal{P}(X)}(\rho) = T \int_X \rho(t) \ln \rho(t) d\mu(t) - \frac{1}{2} K b_\rho \cdot b_\rho \quad (4.19)$$

are in one-to-one correspondence with global minimisers, local minimisers and critical points (respectively) of the function

$$T(\lambda \cdot b^\lambda - \ln(Z_\lambda)) - \frac{1}{2} K b^\lambda \cdot b^\lambda \quad (4.20)$$

by the relation

$$\rho = \frac{1}{Z_\lambda} \exp(\lambda \cdot a(t)). \quad (4.21)$$

Proof. This follows from the preceding discussion. \square

Finally using the machinery of the singular potential it is possible to obtain some simple estimates on the stability of the isotropic phase $\rho(t) = \frac{1}{\mu(X)}$ in mean field free energy models. In the following ϕ will be taken as the Shannon entropy $\phi(x) = x \ln x$, and the constraint functions a_i are taken to be orthonormal in $L^2(X)$. Define the mean field free energy on $\mathcal{P}(X)$ for $T > 0$ and a positive definite $k \times k$ matrix K as

$$\mathcal{I}_{\mathcal{P}}(\rho) = T \int_X \rho \ln \rho d\mu - \frac{1}{2} K b \cdot b \quad (4.22)$$

Proposition 4.5. *If $T > \|a\|_{\infty}^2 \lambda_{\max}(K)$ then the isotropic state is globally stable and there are no other critical points of $\mathcal{I}_{\mathcal{P}}$. If $T > \frac{1}{\mu(X)} \lambda_{\max}(K)$ then the isotropic state is at least locally stable. If $T < \frac{1}{\mu(X)} \lambda_{\max}(K)$ then the isotropic state is unstable.*

Proof. Consider the corresponding free energy functional $\mathcal{I}_{\mathcal{Q}} : \mathcal{Q} \rightarrow \mathbb{R}$ defined by

$$\mathcal{I}_{\mathcal{Q}}(b) = T \psi_s(b) - \frac{1}{2} K b \cdot b$$

Note first that the isotropic state $b = 0$ is always a critical point for $\mathcal{I}_{\mathcal{Q}}$. The Hessian matrix of $\mathcal{I}_{\mathcal{Q}}$ at b is given by

$$H(b) = \frac{\partial^2 \mathcal{I}_{\mathcal{Q}}}{\partial b^2}(b) = T \left(\int_X a \otimes a \rho_b - b \otimes b \right)^{-1} - K. \quad (4.23)$$

The Hölder inequality and that $b \otimes b$ is positive semi-definite gives

$$\int_X a \otimes a \rho_b - b \otimes b \leq \int_X a \otimes a \rho_b \leq \|a\|_{\infty}^2 I \quad (4.24)$$

so that $H(b) \geq \frac{T}{\|a\|_{\infty}^2} I - K$. This gives that for $T > \|a\|_{\infty}^2 \lambda_{\max}(K)$ that $H(b)$ is positive definite for all $b \in \mathcal{Q}$ so that $\mathcal{I}_{\mathcal{Q}}$ is a strictly convex function on \mathcal{Q} , so that there exists at most one critical point. This implies that the isotropic state must be the global minimum and only critical point. Applying Proposition 4.1 gives the corresponding result for $\mathcal{I}_{\mathcal{P}}$. Due to the normalisation of the functions $(a_i)_{i=1}^k$ it holds that $H(0) = T \mu(X) I - K$. Consequently if $T > \frac{1}{\mu(X)} \lambda_{\max}(K)$ then $H(0)$ is positive definite so that $b = 0$ is a local minimum. Conversely, if $T < \frac{1}{\mu(X)} \lambda_{\max}(K)$ then $H(0)$ has a negative eigenvalue so that $b = 0$ is unstable. Applying Proposition 4.2 gives the desired result for the functional $\mathcal{I}_{\mathcal{P}}$. \square

Remark 4.3. If L^{∞} -local minimisers are *a priori* known to be bounded away from zero and $+\infty$, then L^{∞} variations can obtain a critical point condition for L^{∞} -local minimisers analogous to Corollaries 4.1 and 4.2. Similarly, taking the second variation can provide estimates analogous to Proposition 4.5 to assess the stability

under L^∞ perturbations. The advantages of the methods presented in this work are twofold; firstly that all statements are in reference to L^1 -local minimisers, a stronger condition which could not be obtained by two-sided variations since $\mathcal{P}(X)$ has empty relative interior in $L^1(X)$, and secondly the technicality that local minimisers must be shown to be bounded away from zero and $+\infty$ is removed.

Remark 4.4. More general mean-field like free energies can be defined by

$$\mathcal{I}_{\mathcal{P}}(\rho) = \int_X \rho \ln \rho \, d\mu - \int_X \int_X K(t, s) \rho(t) \rho(s) \, d\mu(t) \, d\mu(s) \quad (4.25)$$

with $K \in L^\infty(X \times X)$ a symmetric kernel. If, for example, X is a connected, analytic, bounded, Riemannian manifold then the eigenvectors of the Laplace-Beltrami operator on X , denoted $(a_i)_{i=1}^\infty$, form a countable, dense, orthogonal basis for $L^2(X)$ with a_i analytic for $i \in \mathbb{N}$ ⁶. By Prop. 2.3 they must therefore form a pseudo-Haar set. This allows the decomposition of K as

$$K(t, s) = \sum_{i,j=1}^{\infty} c_{ij} a_i(t) a_j(s), \quad (4.26)$$

so that by truncating this series to finitely many terms the kernel can be approximated, and the approximating model fits into the framework presented in this paper. In many models related to liquid crystals this decomposition would be superfluous, as the kernel is assumed to have a decomposition of the form in (4.26) with c_{ij} non-zero for only finitely many i, j . For example, if $V_\lambda(X)$ denotes the eigenspace of the Laplace-Beltrami operator on X with corresponding eigenvalue λ , then the constraint functions corresponding to the Maier-Saupe potential (dipolar potential, respectively) for nematic liquid crystals are in $V_{-6}(\mathbb{S}^2)$ ($V_{-2}(\mathbb{S}^2)$, respectively)¹³, the McMillan model (see Subsubsection 4.3.2) has two constraint functions, one in $V_{-6}(\mathbb{S}^2 \times [0, 1])$ and another in $V_{2\pi}(\mathbb{S}^2 \times [0, 1])$, and the Straley model²⁸ uses elements of $V_{-6}(\text{SO}(3))$.

Remark 4.5. Propositions 4.1, 4.2 and 4.3 presented together imply that globally stable, locally stable and unstable equilibrium points of the functionals on $\mathcal{P}(X)$ can be found by considering a macroscopic functional defined only on \mathcal{Q} , combined with the maximum entropy assumption.

4.2. Models with Spatial Inhomogeneities

If, as in the Q-tensor theory of liquid crystals, one wishes to consider models with inhomogeneities in a domain $\Omega \subset \mathbb{R}^n$, then one could appeal to the calculus of variations and consider functions $b : \Omega \rightarrow \overline{\mathcal{Q}}$ (note that the closure of \mathcal{Q} is taken for compactness properties), and investigate minimisers of the functional

$$\mathcal{I}(b) = \int_{\Omega} W(\nabla b(x), b(x)) + F(b(x)) \, dx \quad (4.27)$$

in some appropriate function space, where $W : \mathbb{R}^{k \times n} \times \mathcal{Q} \rightarrow \mathbb{R}$ represents the energy of distortions, and $F : \overline{\mathcal{Q}} \rightarrow \mathbb{R}$ represents the free energy of a homogeneous system. However, if F were to be bounded on $\partial\mathcal{Q}$, as is the case for a Landau expansion, then it is possible that minimisers could satisfy $b(x) \in \partial\mathcal{Q}$ for all x in some set of positive measure, providing unphysical solutions. Furthermore this is undesirable from a mathematical perspective since the minimiser would not satisfy the Euler-Lagrange equation. However, if one considers a functional \mathcal{I}_s , given by

$$\mathcal{I}_s(b) = \int_{\Omega} W(\nabla b(x), b(x)) + \psi_s(b(x)) + f(b(x)) dx \quad (4.28)$$

for some f, W bounded away from $-\infty$, then provided finite energy configurations exist, it is immediate that minimisers satisfy $b(x) \in \mathcal{Q}$ for almost every $x \in \Omega$. Furthermore, in particular simple models it is possible to show strict physicality, in the sense that there exists some compact set $K \subset \mathcal{Q}$ for which any global minimiser b satisfies $b \in K$ almost everywhere, which has the consequence that global minimisers satisfy the Euler-Lagrange equation allowing PDE methods to be applied to the problem. This is shown by a projection method based similar to that in Ref. 4.

Lemma 4.2. *Let $g : \mathcal{Q} \rightarrow \mathbb{R}$ be Lipschitz continuous with Lipschitz constant L . Given $M \in \mathbb{R}$ define the set $K_M = \{b \in \mathcal{Q} : \psi_s(b) \leq M\}$, and the nearest point projection $P_M : \mathcal{Q} \rightarrow K_M$. Furthermore given $b \in \mathcal{Q}$ denote $b_M = P_M b$. Then there exists $M_0 \in \mathbb{R}$ so that for all $M > M_0$ and $b \in \mathcal{Q}$,*

$$\psi_s(b) + g(b) \geq \psi_s(b_M) + g(b_M). \quad (4.29)$$

Proof. First note that the projection is well defined, since the convexity and continuity of ψ_s ensure that K_M is a closed convex set. Suppose that the result is false. Then there must exist sequences $M^k \in \mathbb{R}$, $b_k \in \mathcal{Q}$ so that $\lim_{k \rightarrow \infty} M^k = +\infty$, and defining $c_k = P_{M^k} b_k$, for each k the inequality

$$\psi_s(c_k) + g(c_k) > \psi_s(b_k) + g(b_k)$$

is satisfied. By taking a subsequence if necessary (not relabelled), assume that $(M^k)_{k \in \mathbb{N}}$ is an increasing sequence. Using the fact that g is Lipschitz and elementary inequalities for convex functions, it holds that

$$\begin{aligned} L|c_k - b_k| &\geq g(c_k) - g(b_k) \\ &> \psi_s(b_k) - \psi_s(c_k) \\ &\geq \nabla\psi_s(c_k) \cdot (b_k - c_k). \end{aligned} \quad (4.30)$$

From the construction of K_M , ∂K_M is a level set of ψ_s , so that $\nabla\psi_s(c_k)$ is normal to the surface ∂K_{M^k} at c_k . Also from the properties of projections it holds that $b_k - c_k$ is normal to ∂K_{M^k} at c_k , with the same sign as $\nabla\psi_s(c_k)$ in the sense that $\nabla\psi_s(c_k) = s(b_k - c_k)$ for some $s > 0$. This implies $\nabla\psi_s(c_k) \cdot (b_k - c_k) = |\nabla\psi_s(c_k)| |b_k - c_k|$.

Combining this with the previous chain of inequalities gives

$$\begin{aligned} L|c_k - b_k| &\geq |\nabla\psi_s(c_k)||c_k - b_k| \\ &\Rightarrow L \geq |\nabla\psi_s(c_k)| \end{aligned} \tag{4.31}$$

for all k . However since \mathcal{Q} is bounded, a subsequence (not relabelled) can be taken so that $c_k \rightarrow c^*$, and since $c_k \in \partial K_{M^k}$ for all k , this gives that $c^* \in \partial\mathcal{Q}$. Since $\lim_{b \rightarrow \partial\mathcal{Q}} \psi_s(b) = +\infty$, and ψ_s is convex, this also implies that $\lim_{b \rightarrow \partial\mathcal{Q}} |\nabla\psi_s(b)| = +\infty$. Taking the limit as $k \rightarrow \infty$ of the inequality $L \geq |\nabla\psi_s(c_k)|$ gives a contradiction, completing the proof. \square

Theorem 4.1. *Let $\Omega \subset \mathbb{R}^n$. Consider the minimisation problem*

$$\min_{b \in \mathcal{A}} \int_{\Omega} |\nabla b(x)|^p + \psi_s(b(x)) + g(b(x)) dx, \tag{4.32}$$

where $\mathcal{A} = \{b \in W^{1,p}(\Omega; \overline{\mathcal{Q}}) : b|_{\partial\Omega} = b_0\} \neq \emptyset$, $g : \overline{\mathcal{Q}} \rightarrow \mathbb{R}$ is Lipschitz continuous with Lipschitz constant L , $p > 1$ and there exists some compact $K \subset \mathcal{Q}$ such that $b_0(x) \in K$ for \mathcal{H}^{n-1} -almost every $x \in \partial\Omega$. Then there exists a global minimiser, and all global minimisers are strictly physical in the sense that $\psi_s(b) \in L^\infty(\Omega, \mathbb{R})$. In particular, any global minimiser $b^* : \Omega \rightarrow \overline{\mathcal{Q}}$ satisfies the Euler-Lagrange equation for the energy functional if $g \in C^1(\mathcal{Q})$.

Proof. The existence of minimisers follows from a standard direct method argument¹¹. To see that the global minimiser must be strictly physical, assume for the sake of contradiction that there exists a global minimiser $b \in \mathcal{A}$ such that $\psi_s(b)$ is unbounded on Ω . Let $M > M_0$ (as defined in Lemma 4.2) and b_M be the projection of b onto the set K_M as defined in Lemma 4.2. Take M sufficiently large so that $K \subset K_M$, so that b_M satisfies the same boundary conditions as b . Then $\psi_s(b_M) + g(b_M) \leq \psi_s(b) + g(b)$ almost everywhere in Ω by Lemma 4.2. Furthermore, since the term on the left is bounded, and by assumption the term on the right is unbounded this equality is strict on a set of positive measure. Finally since $|\nabla Pb| \leq |\nabla b|$ for any function in $W^{1,p}$ and projection P onto a convex set (see Lemma 10 in Ref. 4), it holds that

$$\int_{\Omega} |\nabla b_M(x)|^p + \psi_s(b_M(x)) + g(b_M(x)) dx < \int_{\Omega} |\nabla b(x)|^p + \psi_s(b(x)) + g(b(x)) dx, \tag{4.33}$$

contradicting that b is a global minimiser. Since the minimiser is bounded away from $\partial\mathcal{Q}$ it is possible to take smooth variations so that the minimiser satisfies the Euler-Lagrange equation. \square

4.3. Approximation of the Singular Potential by Everywhere Defined Functions

In some situations it may be preferable to approximate the singular potential by a globally defined function. In particular, as mentioned in Subsection 4.2, in models

with inhomogeneities the blow up property can pose issues with minimisers satisfying the Euler-Lagrange equation. One tool kit for providing such an approximation is to use the Landau theory²⁹, which argues that the free energy of a system must be analytic in the order parameters, therefore it is possible to replace the free energy with a polynomial in the order parameters that respects the symmetry of the system. Typically, this polynomial will be taken to fourth order. However, as will be shown in this section, the Landau expansion is not necessarily compatible with free energies of the form given in Subsection 4.1, since the Taylor approximation does not necessarily preserve shape properties possessed by the singular potential. This section will only be concerned with the Shannon entropy, due to the relative simplicity of its Taylor approximation as well as its physical relevance.

4.3.1. One Dimensional Examples and Counterexamples

The fourth order Taylor approximation about zero of the singular potential is derived in Appendix Appendix A. For simplicity consider only a single constraint function, and state space $X = [-1, 1]$. Define $m_i = \frac{1}{2} \int_X a(x)^i dx$. The the fourth order Taylor approximation $\psi_s^4 : \mathbb{R} \rightarrow \mathbb{R}$ is given by

$$\begin{aligned} \psi_s^4(b) &= \psi_s(0) + \sum_{j=2}^4 \frac{d^j \psi_s}{db^j}(0) \frac{b^j}{j!} \\ &= \frac{1}{2m_2} b^2 - \frac{m_3}{6m_2^2} b^3 + \frac{3m_3^2 - m_2m_4 + 3m_2^3}{24m_2^5} b^4 + \psi_s(0). \end{aligned}$$

If the Taylor approximation is to represent a kind of macroscopic version of entropy, and approximate the singular potential, then at the very least it should be convex, possess a single local minimum, and be coercive so that the approximation blows up to $+\infty$ as $b \rightarrow \pm\infty$. Coercivity perhaps the most important property, since if it fails then energy minimisation is not possible. The second condition prevents the appearance of “phantom” minimisers for the entropy which should not exist. Finally convexity is a useful property in minimisation problems that one would want to inherit from the singular potential. In the one dimensional case presented, these three conditions have a chain of implication, so that convexity implies the existence of a single critical point, and a single critical point implies coercivity. Note that these implications rely on the fact that ψ_s^4 is convex in some neighbourhood of the origin, and that it is not linear. Similarly since ψ_s^4 is a polynomial that is not affine, convexity implies strict convexity.

Fortunately in the one dimensional case, verifying if these properties hold for a given example is relatively straightforward. In order to be coercive, the coefficient of b^4 must be positive, which gives the condition

$$d_1 = 3m_3^2 - m_2m_4 - 3m_2^3 > 0.$$

If $b \neq 0$ is a critical point of ψ_s^4 , this implies the existence of a non-zero real solution

to

$$\frac{d\psi_s^4}{db}(b) = b \left(\frac{1}{m_2} - \frac{1}{2m_2^3}m_3b + \frac{1}{6m_2^5} (3m_2^3 - m_2m_4 + 3m_3^2) b^2 \right) = 0.$$

By considering the discriminant of the quadratic, this means that ψ_s^4 possesses non-trivial critical points if and only if

$$d_2 = 72m_2^3 - 24m_2m_4 + 63m_3^2 > 0.$$

Finally for convexity, the second derivative of ψ_s^4 is a quadratic in b given by

$$\frac{d^2\psi_s^4}{db^2}(b) = \frac{1}{m_2} - \frac{m_3}{m_2^3}b + \frac{1}{2m_2^5} (3m_2^3 - m_2m_4 + 3m_3^2) b^2,$$

so by the same argument ψ_s^4 is convex if and only if

$$d_3 = 6m_2^3 - 2m_2m_4 + 5m_3^2 > 0.$$

Within the chain of implication that

$$\text{Convexity} \Rightarrow \text{A single critical point} \Rightarrow \text{Coercivity},$$

none of the implications are equivalent for the singular potential. Using the domain $X = [-1, 1]$, there are four examples that produces all four possibilities.

| | | | | |
|-----------------------|-----------------|---|--------------------------------|---------------------|
| $a(x) =$ | x | $x^4 - 7x^3 - 2x^2 + 3x + \frac{7}{15}$ | $7x^3 - x^2 + x + \frac{1}{3}$ | x^3 |
| $m_2 =$ | $\frac{1}{3}$ | $\frac{904}{525}$ | $\frac{92}{9}$ | $\frac{1}{7}$ |
| $m_3 =$ | 0 | $\frac{-6796096}{3378375}$ | $\frac{-12112}{945}$ | 0 |
| $m_4 =$ | $\frac{1}{5}$ | $\frac{4297061248}{287161875}$ | $\frac{6938192}{19305}$ | $\frac{1}{13}$ |
| $d_1 =$ | $\frac{2}{45}$ | $\frac{765059068785152}{452732233078125}$ | $\frac{998000512}{42567525}$ | $\frac{-10}{4459}$ |
| Coercive | Yes | Yes | Yes | No |
| $d_2 =$ | $\frac{16}{15}$ | $\frac{624228970176512}{150910744359375}$ | $\frac{-618770176}{675675}$ | $\frac{-240}{4459}$ |
| Single Critical Point | Yes | Yes | No | No |
| $d_3 =$ | $\frac{4}{45}$ | $\frac{-905889167393792}{1358196699234375}$ | $\frac{-2998034944}{25540515}$ | $\frac{-20}{4459}$ |
| Convex | Yes | No | No | No |

Fig. 1: Table of examples and discriminant values

As these simple examples show, the Taylor expansion approach to providing a polynomial approximation to the free energy will not necessarily work in general, and this is perhaps unsurprising given that the Taylor approximation is generally

only a local approximation of a function. However, the following Weierstrass-type result shows that convex polynomial approximation is possible.

Proposition 4.6 (From Ref. 27). *Let $M \subset \mathbb{R}^k$ be convex and compact, and $f : M \rightarrow \mathbb{R}$ be convex and continuous. Then for every $\epsilon > 0$ there exists some polynomial p_ϵ which is convex on \mathbb{R}^k such that*

$$\sup_{x \in M} |f(x) - p_\epsilon(x)| < \epsilon.$$

The main limitations of this result however are that the approximation can only be performed on compact subsets of \mathcal{Q} , and also, much like the classical Weierstrass result, the proof is non-constructive. It should also be noted that this result requires the function to be only continuous, rather than analytic, so that the Landau theory's assumption of an analytic free energy is unnecessary. One might hope that it is possible to establish a necessary and sufficient condition for the fourth order Taylor approximation to be convex, but in general the problem of establishing if a given polynomial is convex is an NP-hard problem¹. Rather than pursuing a polynomial approximation, the Yosida-Moreau regularisation may be more appropriate, due to its shape preservation properties.

Definition 4.1. For $J > 0$, define the Yosida-Moreau regularisation of ψ_s , denoted ψ^J by

$$\psi^J(b) = \min_{\tilde{b} \in \mathcal{Q}} \psi_s(b) + \frac{J}{2} |\tilde{b} - b|^2. \quad (4.34)$$

The following proposition outlines several of the key properties which suggest that the Yosida-Moreau approximation is an appropriate approximation for the singular potential. The results presented are for the general entropy-like objective functions and pseudo-Haar constraint functions.

Proposition 4.7.

- (1) *The Yosida-Moreau approximation preserves minima, in the sense that for all $J > 0$, $\min \psi^J = \psi^J(0) = \psi_s(0) = \min \psi_s$. For all $b \in \mathbb{R}^k$, $J > 0$, the minimisation problem defining ψ^J , as defined in Equation (4.34), admits a unique solution denoted $G_J(b)$.*
- (2) *For every $b \in \mathbb{R}^k$, $\psi^J(b) \nearrow \psi_s(b)$, where the limit is infinite for $b \notin \mathcal{Q}$.*
- (3) *For all $J > 0$, ψ^J is a convex, differentiable function. Furthermore, the gradient is Lipschitz with Lipschitz constant J . The derivative of ψ^J can be given in terms of $G_J(b)$ by*

$$\frac{\partial \psi^J}{\partial b}(b) = J(b - G_J(b)) = \lambda((G_J(b))). \quad (4.35)$$

- (4) *The Yosida-Moreau approximation can be evaluated numerically by the dual problem,*

$$\psi^J(b) = \max_{\alpha \in \mathbb{R}, \lambda \in \mathbb{R}^k} \alpha + \lambda \cdot b - \int_X \phi^*(\alpha + \lambda \cdot a) d\mu - \frac{1}{2J} |\lambda|^2. \quad (4.36)$$

Proof. The preservation of minima in statement 1 is immediate from the definition, by testing $\tilde{b} = 0$. The existence and uniqueness of a minimiser follow from the coercivity and strict convexity of the objective function. Statement 2 can be found in Ref. 23. Statement 3 can be found in Ref. 9. Statement 4 is given in Ref. 12. $\square\square$

Remark 4.6. The previous proposition has several consequences. Firstly, statements 1 and 2 say that ψ^J approximates ψ_s , as well as preserving desirable shape properties. Statement 3 gives that ψ^J is sufficiently regular for first order methods to be used. Lastly, statement 4 shows that the dual optimisation problem defining to ψ^J is, at face value, no harder to approach than the optimisation problem defining ψ_s .

4.3.2. *The McMillan Model*

As an illustrative and physically meaningful example consider the McMillan model for Isotropic-Nematic-Smectic-A phase transitions²². This is a mean field theory as described previously where the state space is $X = \mathbb{S}^2 \times [0, 1]$, and two constraint functions are given by

$$\begin{aligned} a_1(p, x) &= \frac{1}{2} (3(p \cdot e_1)^2 - 1), \\ a_2(p, x) &= \frac{1}{2} (3(p \cdot e_1)^2 - 1) \cos(2\pi x). \end{aligned}$$

Here e_1 is a unit vector, physically corresponding to the orientation of the material. Due to the rotational symmetry of the constraint functions, it is possible to consider only state variables $(\theta, x) \in [0, \pi] \times [0, 1]$, where $\cos(\theta) = p \cdot e_1$. This approach views X as equivalent to a subset of \mathbb{R}^2 with measure $d\mu(\theta, x) = 2\pi \sin(\theta) d\theta dx$. The constraint functions are analytic and linearly independent, and since the measure μ has the same null sets as the Lebesgue measure, the pseudo-Haar condition is satisfied. These constraints give two order parameters, denoted

$$\begin{aligned} S &= 2\pi \int_0^1 \int_{-\pi}^{\pi} a_1(\theta, x) \rho(\theta, x) d\theta dx \\ \sigma &= 2\pi \int_0^1 \int_{-\pi}^{\pi} a_2(\theta, x) \rho(\theta, x) d\theta dx. \end{aligned}$$

Loosely speaking S corresponds to the degree of orientational order of the molecules, and σ represents a coupling between the order of the molecules and the location of their centre of mass. If $S = \sigma = 0$, then the sample is in an isotropic phase. If $S \neq 0$ and $\sigma = 0$ then it is a nematic phase, and if $S \neq 0$ and $\sigma \neq 0$ then it is a smectic A phase. Before any analysis can be performed, the set of physical moments will be established.

Proposition 4.8. *The set \mathcal{Q} for the constraint functions of the McMillan model is*

given by

$$\mathcal{Q} = \left\{ (S, \sigma) \in \mathbb{R}^2 : S \in \left(-\frac{1}{2}, 1 \right), |\sigma| < \frac{S+2}{3} \right\}. \quad (4.37)$$

Proof. To see that the candidate set contains \mathcal{Q} , use Proposition 2.1 and test against $(\pm 1, 0)$ and $(-\frac{1}{3}, \pm 1)$. The maxima a_1 and $-a_1$ are 1 and $\frac{1}{2}$ respectively, which gives that $S \in (-\frac{1}{2}, 1)$. The maximum of $-\frac{1}{3}a_1(\theta, x) - a_2(\theta, x)$ can be found by noting that

$$\begin{aligned} -\frac{1}{3}a_1(\theta, x) - a_2(\theta, x) &= \frac{1}{2} (1 - 3 \cos(\theta)^2) \left(\frac{1}{3} + \cos(2\pi x) \right) \\ &\leq \frac{1}{2} \cdot \frac{4}{3} = \frac{2}{3}, \end{aligned} \quad (4.38)$$

which is attained at $\theta = \pi$ and $x = 0$. For $u = (-\frac{1}{3}, 1)$ the same argument is used. To show that \mathcal{Q} contains the candidate set is equivalent to showing that the closure of the candidate set is a subset of the closure of \mathcal{Q} since the sets are convex. To prove this, it is then sufficient to show that the four vertices $((1, \pm 1), (-\frac{1}{2}, \pm \frac{1}{2}))$ are contained in the closure of \mathcal{Q} . This can be done constructively, and here only one vertex will be proven with the rest being shown by the same method. Let $0 < \epsilon < 1$, and define the set $A_\epsilon = \{(\theta, x) : \cos(\theta)^2 > 1 - \epsilon, \cos(2\pi x) > 1 - \epsilon\}$. Define

$$\rho_\epsilon(\theta, x) = \left(2\pi \int_{A_\epsilon} \sin(\Theta) d\Theta \right)^{-1} 2\pi \sin(\theta) \chi_{A_\epsilon}(\theta, x). \quad (4.39)$$

This corresponds to a distribution uniform with respect to the measure on X on A_ϵ . Let the corresponding moments be denoted $S_\epsilon, \sigma_\epsilon$. Then it is immediate that

$$\begin{aligned} 1 > S_\epsilon &> \frac{1}{2}(3(1 - \epsilon) - 1) = 1 - \frac{3}{2}\epsilon \\ 1 > \sigma_\epsilon &> \frac{1}{2}(3(1 - \epsilon) - 1)(1 - \epsilon) = 1 - \frac{5}{2}\epsilon + \frac{3}{2}\epsilon^2. \end{aligned} \quad (4.40)$$

Therefore $(S_\epsilon, \sigma_\epsilon) \in \mathcal{Q}$, and by taking ϵ to 0, $(1, 1) \in \overline{\mathcal{Q}}$. \square

Using the formula given in Appendix Appendix A, the exact symbolic integration package in Maple gives the fourth order Taylor approximation to the singular potential as

$$\psi_s^4(S, \sigma) = \frac{425}{196}S^4 + \frac{50}{49}\sigma^2 S^2 + \frac{825}{196}\sigma^4 - \frac{25}{21}S^3 - \frac{50}{7}\sigma^2 S + \frac{5}{2}S^2 + 5\sigma^2 + \psi_s(0, 0).$$

Proposition 4.9. *For the McMillan model, ψ_s^4 is coercive, so that*

$$\lim_{|(S, \sigma)| \rightarrow +\infty} \psi_s^4(S, \sigma) = +\infty.$$

Proof. It is sufficient to show that the fourth order terms are coercive. This can be written as a bilinear form in S^2 and σ^2 , since

$$\frac{425}{196}S^4 + \frac{50}{49}\sigma^2 S^2 + \frac{825}{196}\sigma^4 = \begin{bmatrix} S^2 \\ \sigma^2 \end{bmatrix} \cdot \begin{bmatrix} \frac{425}{196} & \frac{25}{49} \\ \frac{25}{49} & \frac{825}{196} \end{bmatrix} \begin{bmatrix} S^2 \\ \sigma^2 \end{bmatrix}.$$

The eigenvalues of the matrix are given by $\frac{625}{196} \pm \frac{25}{49}\sqrt{5}$, which are both positive (evaluating at approximately 2.04 and 4.33), so that the leading order terms are positive. \square

Proposition 4.10. *In the McMillan model, ψ_s^4 is not a convex function of S, σ , even when restricted to \mathcal{Q} .*

Proof. It is sufficient to show that the Hessian matrix of ψ_s^4 has a negative eigenvalue for some (S, σ) . The Hessian matrix is readily computed as

$$\begin{aligned} H(S, \sigma) &= \frac{1}{49} \begin{bmatrix} 1275S^2 + 100\sigma^2 - 350S + 245 & 200S\sigma - 700\sigma \\ 200S\sigma - 700\sigma & 100S^2 + 2475\sigma^2 - 700S + 490 \end{bmatrix}, \\ \Rightarrow H(1, 0) &= \frac{1}{49} \begin{bmatrix} 1275 - 350 + 245 & 0 \\ 0 & 100 - 700 + 490 \end{bmatrix} \\ &= \frac{1}{49} \begin{bmatrix} 1170 & 0 \\ 0 & -110 \end{bmatrix}. \end{aligned}$$

Therefore it has a negative eigenvalue, and ψ_s^4 is not convex in this case. In particular, it is not convex in a neighbourhood of $(1, 0)$, since the Hessian is continuous. By noting that the intersection of any neighbourhood of $(1, 0)$ with \mathcal{Q} is non-empty since $(1, 0) \in \partial\mathcal{Q}$, this implies that ψ_s^4 is not convex on \mathcal{Q} . \square

To conclude, contour plots of the singular potential and its approximations are given in Figure 2. The singular potential itself is given in Subfigure 2a. It should be noted that ψ_s is only given on $\{(S, \sigma) \in \mathcal{Q} : d((S, \sigma), \partial\mathcal{Q}) > 10^{-2}\}$ to avoid the difficulty in calculating the singular potential near $\partial\mathcal{Q}$. The fourth order Taylor approximation is given in Subfigure 2b. Also included on the plot is a white dashed line, which is the boundary of the set where the Hessian matrix has a negative eigenvalue. The Yosida-Moreau approximation for $J = 100$ is given in Subfigure 2c. The numerical values for ψ_s and ψ^J were obtained via a steepest descent algorithm on the dual optimisation problem.

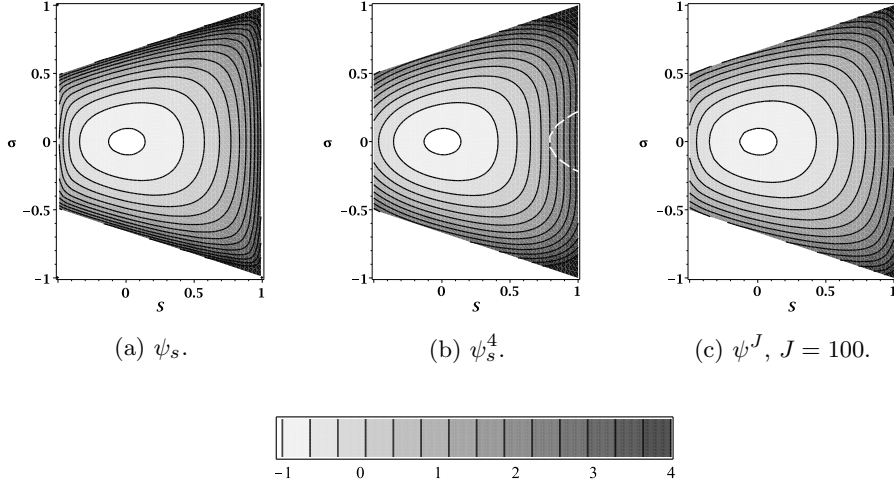


Fig. 2: Contour plots of the singular potential and its approximations.

Appendix A. Appendix: The Fourth Order Taylor Expansion for Shannon Entropy

For this section consider the pseudo-Haar constraint functions $(a_i)_{i=1}^k$ to be orthogonal to the constant function so that $\int_X a_i d\mu = 0$, but not necessarily orthogonal to each other. Furthermore in this section the convention that Greek indices are summed over, whilst Latin indices are free, is used. Given $n \in \mathbb{N}$ and $\rho \in \mathcal{P}(X)$ define the tensor $M^n \in (\mathbb{R}^k)^n$ component-wise by

$$M_{i_1 i_2 \dots i_n}^n(\rho) = \int_X a_{i_1} a_{i_2} \dots a_{i_n} \rho d\mu. \quad (\text{A.1})$$

for $i_j = 1, \dots, k, j = 1, \dots, n$. Consider the function $b \mapsto M^n(\rho_b)$, which when unambiguous will simply be denoted M^n , with ρ_b maximal entropy under the Shannon entropy $\phi(x) = x \ln x$. This allows any maximal entropy $\rho \in \mathcal{P}(X)$ to be written as

$$\rho(t) = \frac{1}{Z} \exp(\lambda \cdot a(t)). \quad (\text{A.2})$$

The map $b \mapsto M^n(\rho_b)$ is differentiable, so by applying the chain rule the derivative can be found as

$$\begin{aligned} \frac{\partial M_{i_1 i_2 \dots i_n}^n}{\partial b_{i_{n+1}}} &= \frac{\partial M_{i_1 i_2 \dots i_n}^n}{\partial \lambda_\alpha} \frac{\partial \lambda_\alpha}{\partial b_{i_{n+1}}} \\ &= \left(\int_X a_{i_1} a_{i_2} \dots a_{i_n} a_\alpha \rho_b d\mu - \frac{1}{Z} M_{i_1 i_2 \dots i_n}^n \frac{\partial Z}{\partial \lambda_\alpha} \right) ((M^2 - b \otimes b)^{-1})_{\alpha i_{n+1}} \\ &= (M_{i_1 i_2 \dots i_n i_\alpha}^{n+1} - M_{i_1 i_2 \dots i_n}^n b_\alpha) V_{\alpha i_{n+1}} \end{aligned} \quad (\text{A.3})$$

where for brevity $V = (M^2 - b \otimes b)^{-1}$. That the derivatives can be seen to satisfy this by applying Propositions 3.3 and 3.4 with some elementary calculus. From here it is now possible to perform a Taylor expansion of ψ_s to low order. Immediately it holds that $\psi_s(0) = -\frac{1}{\mu(X)} \ln(\mu(X))$. For the first derivative it was shown in Proposition 3.4 that

$$\frac{\partial \psi_s}{\partial b_i}(b) = \lambda_i. \quad (\text{A.4})$$

Using the strict convexity of ϕ , it is immediate that the uniform distribution is the global entropy minimiser of $\mathcal{P}(X)$, so that in particular $\frac{1}{\mu(X)} = \rho_b(t)$ for $b = 0$ and almost every $t \in X$. By rewriting the uniform distribution on X as $\frac{1}{\mu(X)} = (\phi')^{-1}(\alpha + \lambda \cdot a(t))$ for $\lambda = 0$ and $\alpha = (\phi')^{-1}(\mu(X)^{-1})$ demonstrates that $\lambda(0) = 0$, so that

$$\frac{\partial \psi_s}{\partial b_i}(0) = 0.$$

Furthermore this gives that $b = 0$ is always a critical point for ψ_s and for any truncated Taylor expansion of ψ_s about $b = 0$ with order greater than or equal to 1. For the second derivative it has been seen in Prop. 3.5 that

$$\begin{aligned} \frac{\partial^2 \psi_s}{\partial b_{i_1} \partial b_{i_2}}(b) &= \left((M^2(b) - b \otimes b)^{-1} \right)_{i_1 i_2} \\ \Rightarrow \frac{\partial^2 \psi_s}{\partial b_{i_1} \partial b_{i_2}}(0) &= \left((M^2(0))^{-1} \right)_{i_1 i_2}. \end{aligned} \quad (\text{A.5})$$

As seen before, this is necessarily a positive definite matrix, and consequently $b = 0$ is always a local minimum for ψ_s and any Taylor expansion of order greater than or equal to 2. For the third derivative first note that the derivative of V will have to be taken and since this involves a matrix inverse a more complex expression appears,

$$\begin{aligned} \frac{\partial V_{i_1 i_2}}{\partial b_{i_3}} &= -V_{i_1 \alpha_1} V_{i_2 \alpha_2} \frac{\partial M_{\alpha_1 \alpha_2}^2 - b_{\alpha_1} b_{\alpha_2}}{\partial b_{i_3}} \\ &= -V_{\alpha_1 i_1} V_{\alpha_2 i_2} V_{\alpha_3 i_3} (M_{\alpha_1 \alpha_2 \alpha_3}^3 - b_{\alpha_1} M_{\alpha_2 \alpha_3}^2 - b_{\alpha_2} M_{\alpha_1 \alpha_3}^2 - b_{\alpha_3} M_{\alpha_1 \alpha_2}^2). \end{aligned} \quad (\text{A.6})$$

This gives that the third derivative at $b = 0$ can be given by

$$\frac{\partial^3 \psi_s}{\partial b_{i_1} \partial b_{i_2} \partial b_{i_3}}(0) = -V_{i_1 \alpha_1}(0) V_{i_2 \alpha_2}(0) V_{i_3 \alpha_3}(0) M_{\alpha_1 \alpha_2 \alpha_3}^3(0). \quad (\text{A.7})$$

Since the Taylor expansion will only be taken to fourth order, and the expansion is to be performed around $b = 0$, all terms with factors of b will be denoted as Ab , which can later be neglected in calculating the fourth derivative of ψ_s at $b = 0$. It

can be given as

$$\begin{aligned}
 \frac{\partial^4 \psi_s}{\partial b_{i_1} \partial b_{i_2} \partial b_{i_3} \partial b_{i_4}} &= V_{i_1 \beta_1} V_{i_2 \alpha_2} V_{i_3 \alpha_3} V_{i_4 \beta_3} V_{\alpha_1 \beta_2} M_{\alpha_1 \alpha_2 \alpha_3}^3 M_{\beta_1 \beta_2 \beta_3}^3 \\
 &+ V_{i_1 \alpha_1} V_{i_2 \beta_1} V_{i_3 \alpha_3} V_{i_4 \beta_3} V_{\alpha_2 \beta_2} M_{\alpha_1 \alpha_2 \alpha_3}^3 M_{\beta_1 \beta_2 \beta_3}^3 \\
 &+ V_{i_1 \alpha_1} V_{i_2 \alpha_2} V_{i_3 \beta_1} V_{i_4 \beta_3} V_{\alpha_3 \beta_2} M_{\alpha_1 \alpha_2 \alpha_3}^3 M_{\beta_1 \beta_2 \beta_3}^3 \\
 &- (V_{i_1 \alpha_1} V_{i_2 \alpha_2} V_{i_3 \alpha_3} V_{i_4 \alpha_4}) \\
 &\times (M_{\alpha_1 \alpha_2 \alpha_3 \alpha_4}^4 - M_{\alpha_1 \alpha_2}^2 M_{\alpha_3 \alpha_4}^2 - M_{\alpha_1 \alpha_3}^2 M_{\alpha_2 \alpha_4}^2 - M_{\alpha_1 \alpha_4}^2 M_{\alpha_2 \alpha_3}^2) \\
 &+ Ab.
 \end{aligned} \tag{A.8}$$

At the isotropic state it must hold that $Ab = 0$, so that the fourth derivative at $b = 0$ is given by the remaining terms.

Acknowledgement

The author would like to thank John Ball for insightful discussions relating to the work in this paper. Also the author would like to thank Apala Majumdar and Tim Sluckin for interesting related discussions that took place during the MLC Young Researchers Meeting at the Isaac Newton Institute in Cambridge, as well as the organisers of the MLC programme for providing the forum for the discussions. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 291053.

References

1. AHMADI, A. A., OLSHEVSKY, A., PARRILO, P. A., AND TSITSIKLIS, J. N. NP-hardness of deciding convexity of quartic polynomials and related problems. *Mathematical Programming* 137, 1-2 (2013), 453–476.
2. AKHIEZER, N. I., AND KEMMER, N. *The classical moment problem: and some related questions in analysis*, vol. 5. Oliver & Boyd Edinburgh, 1965.
3. AMANN, H., ESCHER, J., LEVY, S., AND CARGO, M. *Analysis III*. Springer, 2009.
4. BALL, J. M., AND MAJUMDAR, A. Equilibrium order parameters of liquid crystals in the Q-tensor framework. *In preparation*.
5. BALL, J. M., AND MAJUMDAR, A. Nematic liquid crystals: from Maier-Saupe to a continuum theory. *Mol Cryst Liq Cryst* 525, 1 (2010), 1–11.
6. BÉRARD, P. H. *Spectral geometry direct and inverse problems*. Springer, 1986.
7. BORWEIN, J. M., AND LEWIS, A. S. Duality relationships for entropy-like minimization problems. *SIAM J. Control Optim.* 29, 2 (1991), 325–338.
8. BORWEIN, J. M., AND LEWIS, A. S. Partially finite convex programming, part I: Quasi relative interiors and duality theory. *Math. Program.* 57, 1-3 (1992), 15–48.
9. BREZIS, H. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, vol. 5. Elsevier, 1973.
10. CALDERER, M. C., GARAVITO, C. A., AND YAN, B. A Landau-de Gennes theory of liquid crystal elastomers. *arXiv preprint arXiv:1303.6220* (2013).
11. DACOROGNA, B. *Direct methods in the calculus of variations*, vol. 78. Springer, 2007.

12. DECARREAU, A., HILHORST, D., LEMARÉCHAL, C., AND NAVAZA, J. Dual methods in entropy maximization. application to some problems in crystallography. *SIAM Journal on Optimization* 2, 2 (1992), 173–197.
13. FATKULLIN, I., AND SLASTIKOV, V. Critical points of the Onsager functional on a sphere. *Nonlinearity* 18, 6 (2005), 2565–2580.
14. FEIREISL, E., ROCCA, E., SCHIMPERNA, G., AND ZARNESCU, A. Evolution of non-isothermal Landau-de Gennes nematic liquid crystals flows with singular potential. *arXiv preprint arXiv:1207.1643* (2012).
15. HAN, J., LUO, Y., WANG, W., AND ZHANG, P. From microscopic theory to macroscopic theory: a systematic study on static modeling for liquid crystals. *arXiv preprint arXiv:1305.4889* (2013).
16. HARRIS, S. *An introduction to the theory of the Boltzmann equation*. Courier Dover Publications, 2004.
17. JAYNES, E. T. Information theory and statistical mechanics. *Phys. Rev.* 106, 4 (1957), 620–630.
18. KATRIEL, J., KVENTSEL, G., LUCKHURST, G., AND SLUCKIN, T. Free energies in the Landau and molecular field approaches. *Liquid Crystals* 1, 4 (1986), 337–355.
19. LEWIS, A. S. Consistency of moment systems. *Technical Report CORR 92-23* (1993).
20. LEWIS, A. S. Consistency of moment systems. *Canadian Journal of Mathematics* 47, 5 (1995), 995–1006.
21. MAJUMDAR, A. Equilibrium order parameters of nematic liquid crystals in the Landau-de Gennes theory. *Eur. J. Appl. Math* 21 (2010), 181–203.
22. McMILLAN, W. L. Simple molecular model for the smectic A phase of liquid crystals. *Physical Review A* 4, 3 (1971), 1238.
23. MOREAU, J.-J. Proximité et dualité dans un espace hilbertien. *Bulletin de la Société mathématique de France* 93 (1965), 273–299.
24. MOTTRAM, N. J., AND NEWTON, C. J. P. Introduction to Q-tensor theory. *arXiv:1409.3542* (2014).
25. ONSAGER, L. The effects of shape on the interaction of colloidal particles. *Annals of the New York Academy of Sciences* 51, 4 (1949), 627–659.
26. ROCKAFELLAR, R. T. *Convex analysis*, vol. 28. Princeton university press, 1997.
27. SHVEDOV, A. S. Coconvex approximation of functions of several variables by polynomials. *Matematicheskii Sbornik* 157, 4 (1981), 577–589.
28. STRALEY, J. P. Ordered phases of a liquid of biaxial particles. *Physical Review A* 10, 5 (1974), 1881.
29. TOLÉDANO, J.-C., AND TOLÉDANO, P. *The Landau theory of phase transitions*. World Scientific, 1987.
30. TRELOAR, L. R. G. *The physics of rubber elasticity*. Oxford University Press, 1975.
31. WILKINSON, M. Strict physicality of global weak solutions of a Navier-Stokes Q-tensor system with singular potential. *arXiv preprint arXiv:1211.6083* (2012).