Displacement Convexity – A Useful Framework for the Study of Spatially Coupled Codes

Rafah El-Khatib
EPFL, Switzerland
Email: rafah.el-khatib@epfl.ch

Nicolas Macris
EPFL, Switzerland
Email: nicolas.macris@epfl.ch

Ruediger Urbanke
EPFL, Switzerland
Email: ruediger.urbanke@epfl.ch

***Abstract*—Spatial coupling has recently emerged as a powerful paradigm to construct graphical models that work well under low-complexity message-passing algorithms. Although much progress has been made on the analysis of spatially coupled models under message passing, there is still room for improvement, both in terms of simplifying existing proofs as well as in terms of proving additional properties.**

We introduce one further tool for the analysis, namely the concept of displacement convexity. This concept plays a crucial role in the theory of optimal transport and, quite remarkably, it is also well suited for the analysis of spatially coupled systems. In cases where the concept applies, displacement convexity allows functionals of distributions which are not convex in the usual sense to be represented in an alternative form, so that they are convex with respect to the new parametrization. As a proof of concept we consider spatially coupled (l, r)-regular Gallager ensembles when transmission takes place over the binary erasure channel. We show that the potential function of the coupled system is displacement convex. Due to possible translational degrees of freedom convexity by itself falls short of establishing the uniqueness of the minimizing profile. For the spatially coupled (l, r)-regular system strict displacement convexity holds when a global translation degree of freedom is removed. Implications for the uniqueness of the minimizer and for solutions of the density evolution equation are discussed.

I. Introduction

Spatially coupled codes were introduced in the form of low-density parity-check codes by Felstrom and Zigangirov in [1]. Such codes are constructed by spatially coupling nearby replicas of a code defined on a graph. It has been proven that such ensembles perform very well under low-complexity message-passing algorithms. Indeed, this combination achieves essentially optimal performance. More generally, the concept of spatial coupling is proving to be very useful not only for coding but also in compressive sensing, statistical physics, and random constraint solving problems. Given this range of applications, it is worth investigating basic properties of this construction in generality. Our aim is to introduce one further tool for the analysis of such systems – namely the concept of displacement convexity. Displacement convexity plays a crucial role in the theory of optimal transport. But it is also very well suited as a tool for the analysis of spatially coupled graphical models.

One of the most important properties of spatially coupled codes is that they exhibit the so-called threshold saturation phenomenon. That is, spatially coupled ensembles generically have a BP threshold which is as large as the maximum-a posteriori (MAP) threshold of the underlying ensemble, i.e., their threshold has *saturated* to the largest possible value. This result has been proved for transmission over the BEC in [2], [3], and [4] and for transmission over general binary-input memoryless output-symmetric (BMS) channels in [5] and [6].

The tool we introduce probably has a considerably larger range of applications for (coding) systems which are governed by a variational principle. We use (l, r)-regular Gallager ensembles as a proof of concept. We run the belief-propagation algorithm on this ensemble and express it in the variational form using the potential functional [3], [6]. The potential functional essentially is an average of the Bethe free energy on the ensemble and the channel output. The density evolution (DE) equations can be obtained by differentiating this potential. In particular, the minimizers of the potential are given by the "erasure probablity profiles" that are solutions of the DE equations.

For the (l, r)-regular ensemble, we use displacement convexity [7] to prove that the potential describing the system is convex with respect to an alternative structure of probability measures. Roughly speaking these are the probability measures associated to erasure probability profiles viewed as cumulative distribution functions (cdf's). We consider the static case, when the decoder phase transition threshold is equal to the maximum a posteriori (MAP) threshold. This displacement convexity property plays a fundamental role in characterizing the set of minimizing profiles in a suitable space of *increasing profiles*. For our case we get strict displacement convexity once a global translational degree of freedom of the profiles is removed. These results allow to conclude that in a suitable space of increasing profiles the minimizer of the potential functional is unique up to translations, and so is the solution of the DE equation.

We also look at the minimization problem in a more general space of profiles that are not necessarily increasing. We show that the potential functional satisfies rearrangement inequalities which allow to reduce the search for minimizers to a space of increasing profiles. However it is not clear if and when the rearrangement inequalities are strict. As a result the analysis falls short of establishing that there cannot exist non increasing profiles that are minimizers and solutions of the DE equation.

The paper is organized as follows: Section II introduces the framework for our analysis and our main results. We then give

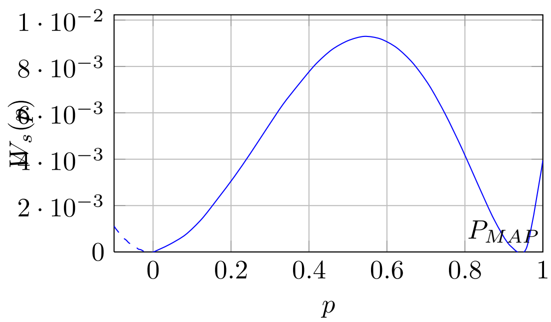

Fig. 1. The plot of the single system potential $W_s(p)$ as a function of the check-erasure probability p, for a $(3,6)$ uncoupled ensemble and $\epsilon = \epsilon_{\text{MAP}}$. There are two minima at $p = 0$ and $p = p_{\text{MAP}}$.

a quick introduction of the notion of displacement convexity in Section III. Finally, Section IV presents a proof of existence of the profile that minimizes the potential, and Section V proves that the functional is displacement convex. We discuss possible generalizations and open issues in the conclusion.

II. Setting and Main Results

In this section, we introduce the model and the associated variational problem to which we apply the displacement convexity proof technique, and we state our main result.

A. (l, r, L, w)-Regular Ensembles on the BEC

Consider the spatially coupled (l, r, L, w)-regular ensemble, described in detail in [5], where the parameters represent the left degree, right degree, system length, and coupling window size (or smoothing parameter), respectively. Specifically, the ensemble is constructed as follows: consider $2L+1$ replicas of a protograph of an (l, r)-regular ensemble. We couple these components by connecting every variable node to l check nodes, and every check node to r variable nodes. The connections are chosen randomly: for a variable node at position z, each of its l connections is chosen uniformly and independently in the range $[z, \dots, z+w-1]$, and for a check node at position z, each of its r connections is chosen uniformly and independently in the range $[z-w+1, \dots, z]$.

For the channel we take a BEC with parameter ϵ. Let $\check{x}_z;\ z \in [-L, \dots, L]$ denote the erasure probability of the variable node at position z. Consider the average over a window $x_z \equiv \frac{1}{w}\sum_{i=0}^{w-1} \check{x}_{z-k}$. Then the fixed-point (FP) condition implied by density evolution (DE) is

$$x_z = \frac{\epsilon}{w}\sum_{k=0}^{w-1}\Big(1 - \frac{1}{w}\sum_{i=0}^{w-1}(1 - x_{z-k+i})^{r-1}\Big)^{l-1}.$$

This FP condition can be obtained by minimizing a "potential functional", which is

$$\frac{1}{w}\sum_{z=-L}^{L}\Big\{ - x_z(1-x_z)^{r-1} + \frac{1}{r} - \frac{1}{r}(1-x_z)^r - \frac{\epsilon}{l}\Big(\frac{1}{w}\sum_{u=0}^{w-1}(1-(1-x_{z+u})^{r-1})\Big)^l\Big\}. \tag{1}$$

At this point the normalization $1/w$ is a convenience whose reason will immediately appear.

The natural setting for displacement convexity is the continuum case. We will therefore consider the continuum limit of (1). Extending our results to the discrete setting is one among various open problems. We define the rescaled variables $\tilde{z} = \frac{z}{w}$, $\tilde{u} = \frac{u}{w}$ and the rescaled function $\tilde{x}(\frac{z}{w}) \equiv x_z$. It is easy to see that (1) becomes a Riemann sum. When we take the limit $L \to +\infty$ first and then $w \to +\infty$, we find

$$\int_{\mathbb{R}} \mathrm{d}\tilde{z}\,\Big\{ - \tilde{x}(\tilde{z})(1-\tilde{x}(\tilde{z}))^{r-1} - \frac{1}{r}(1-\tilde{x}(\tilde{z}))^r + \frac{1}{r} - \frac{\epsilon}{l}\Big(\int_0^1 \mathrm{d}\tilde{u}\,(1-(1-\tilde{x}(\tilde{z}+\tilde{u}))^{r-1})\Big)^l\Big\}. \tag{2}$$

At this point the reader might wonder if the integrals converge. As explained in the introduction, we look in this paper at the decoder phase transition threshold $\epsilon = \epsilon_{\text{MAP}}$. We give at the end of this paragraph the conditions on the erasure probability profile needed to have a well defined problem. From now on, the reader should think of the noise level as fixed to the value $\epsilon = \epsilon_{\text{MAP}}$, although we abuse notation by simply writing ϵ in the formulas that follow.

It is more convenient to express (2) with the function $p(z) = 1-(1-\tilde{x}(z))^{r-1}$. Note that this function is interpreted as the erasure probability emitted by check nodes. Summarizing, the potential functional of interest is

$$\mathcal{W}[p(\cdot)] = \int_{\mathbb{R}} \mathrm{d}z\Big\{\Big(1-\frac{1}{r}\Big)\Big(1-p(z)\Big)^{\frac{r}{r-1}} - (1-p(z)) + \frac{1}{r} - \frac{\epsilon}{l}\Big(\int_0^1 \mathrm{d}u\, p(z+u)\Big)^l\Big\}. \tag{3}$$

A word about the notation here: we use *square* brackets for functionals i.e., "functions of functions" and usual *round* brackets for functions of a real variable. The continuum limit of the DE equation expressed in terms of $p(z)$ reads

$$1-(1-p(z))^{\frac{1}{r-1}} = \epsilon\int_0^1 dv\Big(\int_0^1 \mathrm{d}u\, p(z+u-v)\Big)^{l-1} \tag{4}$$

One can check that (4) gives the stationary points of (3).

Equation (3) can be expressed as a sum of two contributions $\mathcal{W}_{\text{single}}[p(\cdot)] + \mathcal{W}_{\text{int}}[p(\cdot)]$ which are defined as follows:

$$\mathcal{W}_{\text{single}}[p(\cdot)] = \int_{\mathbb{R}} \mathrm{d}z\Big\{\Big(1-\frac{1}{r}\Big)(1-p(z))^{\frac{r}{r-1}} - (1-p(z)) + \frac{1}{r} - \frac{\epsilon}{l}p(z)^l\Big\} \equiv \int_{\mathbb{R}} \mathrm{d}z\, W_s(p(z)), \tag{5}$$

$$\mathcal{W}_{\text{int}}[p(\cdot)] = \int_{\mathbb{R}} \mathrm{d}z\,\frac{\epsilon}{l}\Big\{p(z)^l - \Big(\int_0^1 \mathrm{d}u\, p(z+u)\Big)^l\Big\}. \tag{6}$$

We call (5) the "single system potential functional" and (6) the "interaction functional". The following remarks explain the interpretation suggested by these names. The term (6) vanishes when evaluated for a constant $p(z) = p$. Moreover the integrand of (5), namely $W_s(p(z)) = W_s(p)$ is just the potential of the underlying uncoupled code ensemble. This is easily seen by recognizing that the usual DE equation for the erasure probability of checks is recovered by setting the derivative of $W_s(p)$ to zero. We will call $W_s(p)$ the "single system potential". A plot of $W_s(p)$ for the $(3, 6)$-ensemble is shown in Figure I when $\epsilon = \epsilon_{\text{MAP}}$. The figure shows that the single potential vanishes at $p = 0$ and $p = p_{\text{MAP}}$, some positive value. This is a generic feature of all (l, r)-regular code ensembles as long as $l \geq 3$ (for cycle codes $l = 2$, we have $p_{\text{MAP}} = 0$). This shows, in particular, that in order for the integrals in (3) to be well defined, we have to consider profiles $p(z)$ that tend quickly enough to values, as $z \to \pm\infty$, such that the potential vanishes. Besides we will, for simplicity, restrict ourselves to continuous profiles.

The above remarks motivate us to define the following spaces of profiles. Let

$$\mathcal{S} = \{p(\cdot) : \mathbb{R} \to \mathbb{R}^+ \text{ continuous and s.t} \lim_{z\to-\infty} zp(z) = 0, \lim_{z\to+\infty} z(p(z) - p_{\text{MAP}}) = 0\}. \tag{7}$$

Note that the left limit is 0 and the right limit is p_{MAP}. We will also need a space of increasing profiles (which can be flat over some intervals)

$$\mathcal{S}' = \{p(\cdot) \in \mathcal{S} : \text{ increasing}\}$$

and a space of strictly increasing profiles,

$$\mathcal{S}'' = \{p(\cdot) \in \mathcal{S} : \text{ strictly increasing}\}.$$

As will become apparent it is useful to think of profiles in $\mathcal{S}'$ and $\mathcal{S}''$ as cdf's of measures over $\mathbb{R}$. Here these measures are normalized so that the measure of $\mathbb{R}$ is p_{MAP}. Note that for $p(\cdot) \in \mathcal{S}'$ the support of the associated measure is not necessarily the whole real line, while it is the whole real line for $p(\cdot) \in \mathcal{S}''$. Finally we introduce the space $\mathcal{S}''_0$ of strictly increasing profiles that are pinned at the origin, more precisely $p(\cdot) \in \mathcal{S}''_0$ if and only if $p(\cdot) \in \mathcal{S}''$ and $p(0) = p_{\text{MAP}}/2$.

B. Main Results

It is easy to see that $\mathcal{W}[p(\cdot)]$ is bounded from below, more precisely $\inf_{p(\cdot)\in\mathcal{S}} \mathcal{W}[p(\cdot)] \geq -\frac{1}{2}p^l_{\text{MAP}}$. Indeed $W_s(p) \geq 0$ as seen in Figure I and using Jensen's inequality one can show that $\mathcal{W}_{\text{int}}[p(\cdot)] \geq -\frac{1}{2}p^l_{\text{MAP}}$ (see Lemma 4.1). The first non-trivial question one may ask is whether the minimum is attained in $\mathcal{S}$. Using a rearrangement inequality of Brascamp-Lieb-Luttinger [12] this question can be reduced to the same one in $\mathcal{S}' \subset \mathcal{S}$. Existence in $\mathcal{S}'$ will be answered by means of the so-called "direct method", a standard strategy of calculus of variations [13], [14], (see Section IV-B). Once we know that a minimizer exists in $\mathcal{S}'$ we prove that it cannot be in $\mathcal{S}' \setminus \mathcal{S}''$. The following existence theorem is proven in Section IV-B.

Theorem 2.1: Let $\epsilon = \epsilon_{\text{MAP}}$. The functional $\mathcal{W}[p(\cdot)]$ achieves its minimum over $\mathcal{S}$ in the subspace $\mathcal{S}''$. There does not exist a minimum in $\mathcal{S}' \setminus \mathcal{S}''$.

The existence of a minimum in $\mathcal{S}' \setminus \mathcal{S}''$ is excluded, i.e., a minimizer that increases has to be strictly increasing. However we are not able to exclude the existence of a minimizer in $\mathcal{S} \setminus \mathcal{S}'$, in other words a minimizer that would have "oscillations". In order to exclude such minimizers we would have to study under what conditions, in our context, the Brascamp-Lieb-Luttinger inequality is strict; but we do not address this issue in the present work. In general this can be a difficult problem, see [15], [16].

When functionals are convex one obtains important information on the set of minimizers. For example strict convexity implies that the minimizer is unique. Thus the next natural question is whether or not the functional $\mathcal{W}[p(\cdot)]$ is (strictly) convex. This is in fact not true, but we will show that it is *displacement convex* in $\mathcal{S}'$ (hence also in $\mathcal{S}''$). Here displacement convexity refers to convexity under an interpolation path that is different from the usual linear combination. The displacement convexity in $\mathcal{S}'$ (and $\mathcal{S}''$) cannot be strict since the system certainly has at least one translational degree of freedom: indeed the functional is invariant under a global translation, i.e., we have $\mathcal{W}[p(\cdot+\tau)] = \mathcal{W}[p(\cdot)]$. To investigate if displacement convexity is strict it is convenient to remove this degree of freedom by pinning the profiles, say at the origin. In fact, we will prove strict displacement convexity of the functional only in the space $\mathcal{S}''_0$ of pinned and strictly increasing profiles.

Theorem 2.2: Let $\epsilon = \epsilon_{\text{MAP}}$. The functional $\mathcal{W}[p(\cdot)]$ is displacement convex on $\mathcal{S}'$, and strictly displacement convex on $\mathcal{S}''_0$.

This implies that there is a unique minimizer in $\mathcal{S}''_0$. But since the existence of a minimizer is excluded in $\mathcal{S}' \setminus \mathcal{S}''$ (by theorem 2.1), we can conclude that the only minimizers in $\mathcal{S}'$ are translates of the unique one in $\mathcal{S}''_0$. These consequences also translate into properties of solutions of the DE equation (4).

Corollary 2.3: Let $\epsilon = \epsilon_{\text{MAP}}$. In the space $\mathcal{S}''_0$ the functional $\mathcal{W}[p(\cdot)]$ has a unique minimizer. In $\mathcal{S}'$ all minimizers are translates of it. Similarly, in $\mathcal{S}''_0$ the DE equation (4) has a unique solution, and in $\mathcal{S}'$ all solutions are translates of it.

The proofs of Theorem 2.2 and Corollary 2.3 are given in Section VI.

We would like to point out that while displacement convexity itself is quite general and can presumably be generalized to the general potential functionals of [3], the issue of strict displacement convexity is more subtle. In fact T. Richardson [17] pointed out examples of system where "internal" translation degrees of freedom may exist (besides the global one) which would spoil the unicity up to global translations.

III. Displacement Convexity

Displacement convexity can be very useful in functional analysis. It goes back to McCann [7] and plays an important role in the theory of optimal transport [8]. It has been used in

[9] and [10] to study a functional governing a spatially coupled Curie-Weiss model, which bears close similarities with the coding theory model studied here (see [11]). In this section, we give a quick introduction to the tool of displacement convexity.

Recall first that the usual notion of convexity of a generic functional $\mathcal{F}[p(\cdot)]$ on a generic space $\mathcal{X}$ means that for all $p_0(\cdot), p_1(\cdot) \in \mathcal{X}$ and $\lambda \in [0,1]$,

$$\mathcal{F}[(1-\lambda)p_0(\cdot) + \lambda p_1(\cdot)] \leq \lambda\mathcal{F}[p_0(\cdot)] + (1-\lambda)\mathcal{F}[p_1(\cdot)].$$

Lemma 4.3 in Section IV shows that we can restrict the minimization problem to the space of increasing profiles. Thus, the discussion below assumes that we consider only such profiles. This is the correct setting for defining displacement convexity.

An increasing profile with left limit 0 and right limit p_{MAP} can be thought of as a cdf (up to scaling because the right limit is not 1). Further, such increasing functions have increasing inverse functions (which can also be thought of as cdfs, up to scaling). More precisely, consider the following bijective maps that associate (with an abuse of notation) to a cdf $p(\cdot)$ its inverse $z(\cdot)$:

$$z(p) = \inf\{\, z:\ p(z) > p\},$$
$$p(z) = \inf\{\, p:\ z(p) > z\}.$$

For any two increasing profiles $p_0(\cdot), p_1(\cdot) \in \mathcal{S}'$, we consider $z_0(\cdot), z_1(\cdot)$ their respective inverses under the maps defined above. Then for any $\lambda \in [0,1]$, the interpolated profile $p_\lambda(\cdot)$ is defined as follows:

$$z_\lambda(p) = (1-\lambda)z_0(p) + \lambda z_1(p),$$
$$p_\lambda(z) = \inf\{\, p:\ z_\lambda > z\}.$$

In words, the difference in interpolation under the alternative structure is that the linear interpolation is applied on the inverse of the profiles of interest, and the effect of such an interpolation is then mapped back into the space of profiles. It is not difficult to see that if $p_0(\cdot)$ and $p_1(\cdot)$ are in $\mathcal{S}'$, $\mathcal{S}''$ or $\mathcal{S}''_0$, then so are the interpolating profiles $p_\lambda(\cdot)$ for all $\lambda \in [0,1]$.

Displacement convexity of $\mathcal{W}[p(\cdot)]$ on the space $\mathcal{S}'$ simply means that the following inequality holds:

$$\mathcal{W}[p_\lambda(\cdot)] \leq (1-\lambda)\mathcal{W}[p(\cdot)] + \lambda\mathcal{W}[p'(\cdot)] \tag{8}$$

for any $p(\cdot), p'(\cdot) \in \mathcal{S}'$ and $\lambda \in [0,1]$. Strict displacement convexity means that this inequality is strict as long as $p_0(\cdot)$ and $p_1(\cdot)$ are distinct and $\lambda \in]0,1[$. We will prove displacement convexity of $\mathcal{W}[p(\cdot)]$ by separately proving this property, in Sections V-A and V-B, for the two functionals (5) and (6), respectively. Moreover we will see that (6) is strictly displacement convex in $\mathcal{S}''_0$.

IV. Existence of Minimizing Profile

In this section, we prove that the functional $\mathcal{W}$ attains its minimum.

A. Preliminaries

We start by some preliminaries to show that one can restrict the search of minimizing profiles to those in $\mathcal{S}$ that are monotone increasing. The proofs of the lemmas can be found in Appendices A-E.

The first lemma states that the interaction potential is bounded from below.

Lemma 4.1: For any $p(\cdot)$ in $\mathcal{S}$,

$$\int_{\mathbb{R}} \mathrm{d}z \left\{ p(z)^l - \left(\int_0^1 \mathrm{d}u\, p(z+u) \right)^l \right\} \geq -\frac{1}{2}p_{\text{MAP}}^l.$$

Proof: See Appendix A. ■

We remark that any constant lower bound is sufficient for our purposes: finding profiles that minimize a potential is equivalent to find those that minimize a potential added to a constant. The following lemma states that a truncation of the profile at the value p_{MAP} decreases the potential functional, so we may restrict our search of minimizing profiles to those with range $p(z) \in [0, p_{\text{MAP}}]$.

Lemma 4.2: Define $\bar{p}(z) = \min\{p(z), p_{\text{MAP}}\}$. For all $p(\cdot) \in \mathcal{S}$ we have

$$\mathcal{W}[p(\cdot)] \geq \mathcal{W}[\bar{p}(\cdot)],$$

and the inequality is strict if $p(\cdot) \neq \bar{p}(\cdot)$.

Proof: See Appendix B. ■

We next restrict our search of minimizing profiles to increasing ones. In order to achieve this we will use rearrangement inequalities. Here we will need a notion of increasing rearrangement, see [18]. In words, an increasing rearrangement associates to any function $p(\cdot) \in \mathcal{S}$ with range $[0, p_{\text{MAP}}]$ an increasing function $p^*(\cdot) \in \mathcal{S}'$ so that the total mass is preserved. More formally, any non-negative function in $\mathcal{S}$ can be represented in layer cake form

$$p(z) = \int_0^{+\infty} dt\, \mathbb{1}_{E_t}(z),$$

where $\mathbb{1}_{E_t}(z)$ is the indicator function of the level set $E_t = \{z | p(z) > t\}$. For each t, the level set E_t can be written as the union of a bounded set A_t and a half line $]a_t, +\infty[$. We define the rearranged set $E_t^* =]a_t - |A_t|, +\infty[$. The increasing rearrangement of $p(\cdot)$ is the new function $p^*(\cdot)$ whose level sets are E_t^*. More explicitly,

$$p^*(z) = \int_0^{+\infty} dt\, \mathbb{1}_{E_t^*}(z).$$

Lemma 4.3: Take any $p(\cdot) \in \mathcal{S}$ and let $p^*(\cdot) \in \mathcal{S}'$ be its increasing rearrangement. Then,

$$\mathcal{W}[p(\cdot)] \geq \mathcal{W}[p^*(\cdot)].$$

Proof: See Appendix C. ■

We can thus restrict the search of minimizing profiles to the space of increasing profiles (but as already explained before we cannot exclude that there exist a $p(\cdot) \in \mathcal{S}$ such that $\mathcal{W}[p(\cdot)] = \mathcal{W}[p^*(\cdot)]$). In fact, the following lemma allows to further restricts the search of minimizing profiles to strictly increasing ones.

Lemma 4.4: Let $p(\cdot) \in \mathcal{S}'$ be a minimizer of the potential functional $\mathcal{W}[p(\cdot)]$ that is in $\mathcal{S}'$. Then it must be strictly increasing, i.e., $p(\cdot) \in \mathcal{S}''$.

Proof: See Appendix D. ■

The final step of these preliminaries concerns a necessary condition that any *minimizing sequence* in $\mathcal{S}_0''$ must satisfy. It is useful to think of such profiles as cdf's. A minimizing sequence in $\mathcal{S}_0''$ is by definition any sequence $p_n(\cdot) \in \mathcal{S}_0''$ such that

$$\lim_{n\to\infty} \mathcal{W}[p_n(\cdot)] = \inf_{p\in\mathcal{S}_0''} \mathcal{W}[p(\cdot)]. \tag{9}$$

Such a sequence exists as long as the functional is bounded from below. Since $W_s(p) \geq 0$ and due to Lemma 4.1, this is true. Consider the sequence of probability measures associated to the sequence of cdfs $p_n(\cdot)$. The following lemma states that this sequence of measures is *tight*.

Lemma 4.5: Let $p_n(\cdot) \in \mathcal{S}_0''$ be a minimizing sequence of cdfs. For any $\delta > 0$ we can find $M_\delta > 0$ (independent of n) such that

$$p_n(M_\delta) - p_n(-M_\delta) > (1-\delta)p_{\text{MAP}}$$

for all n.

Proof: See Appendix E. ■

B. The Direct Method

The direct method in the calculus of variations [13]-[14] is a standard scheme to prove that minimizers exist. We use this method to obtain the following theorem:

Proof of theorem 2.1: Let us take any minimizing sequence $p_n(\cdot)$ of cdfs i.e., a sequence that satisfies (9). By Lemma 4.5 the corresponding sequence of measures is tight. Thus by a simple version of Prokhorov's theorem for measures on the real line, we can extract a (point-wise) convergent subsequence of cdfs $p_{n_k}(\cdot) \to p_\ell(\cdot)$ as $k \to +\infty$ with $p_\ell(\cdot) \in \mathcal{S}_0''$. By Fatou's Lemma, one can check that the potential functional is lower-semi-continuous, which means

$$\mathcal{W}[p_\ell(\cdot)] \leq \liminf_{k\to+\infty} \mathcal{W}[p_{n_k}(\cdot)]. \tag{10}$$

Putting (9) and (10) together,

$$\mathcal{W}[p_\ell(\cdot)] \leq \liminf_{k\to+\infty} \mathcal{W}[p_{n_k}(\cdot)] = \lim_{n\to+\infty} \mathcal{W}[p_n(\cdot)] = \inf_{\mathcal{S}_0''} \mathcal{W}[p(\cdot)],$$

On the other hand $\inf_{\mathcal{S}_0''} \mathcal{W}[p(\cdot)] \leq \mathcal{W}[p_\ell(\cdot)]$. Thus, we conclude that $\inf_{\mathcal{S}_0''} \mathcal{W}[p(\cdot)] = \mathcal{W}[p_\ell(\cdot)]$.

We have shown that the minimum is achieved in $\mathcal{S}_0''$ the space of strictly increasing profiles pinned at the origin. Hence it is achieved in $\mathcal{S}''$ and $\mathcal{S}'$ (note that by translation invariance, translations of $p_\ell(\cdot)$ are minimizers in these spaces). Finally, Lemma 4.4 ensures that there is no minimum in $\mathcal{S}' \setminus \mathcal{S}''$. ■

V. Analysis of Displacement Convexity for the Functional $\mathcal{W}[p(\cdot)]$

This section contains the main results of the paper, namely that the potential functional $\mathcal{W}[p(\cdot)]$ is displacement convex in $\mathcal{S}'$ and strictly displacement convex in $\mathcal{S}_0''$.

A. Displacement Convexity of the Single-Potential Term

We first prove that the single-potential functional $\mathcal{W}_{\text{single}}[p(\cdot)]$ is displacement convex. Note that the single system potential $W_s(p)$ is *not convex* in the usual sense (see Figure I).

Proposition 5.1: Let $p_0(\cdot)$ and $p_1(\cdot)$ be in $\mathcal{S}'$ and let $p_\lambda(\cdot)$ the interpolating profile as defined in Section III. Then

$$\mathcal{W}_{\text{single}}[p_\lambda(\cdot)] = (1-\lambda)\mathcal{W}_{\text{single}}[p_0(\cdot)] + \lambda\mathcal{W}_{\text{single}}[p_1(\cdot)].$$

Proof: Recall that $\mathcal{W}_{\text{single}}[p(\cdot)] = \int_{\mathbb{R}} \mathrm{d}z W_s(p(z))$. Recall also that $p_\lambda(z)$ as defined in Section III is the inverse of $z_\lambda(p) = (1-\lambda)z_0(p) + \lambda z_1(p)$. Thus

$$\begin{aligned}\int_{\mathbb{R}} \mathrm{d}z W_s(p_\lambda(z)) &= \int_0^{p_{\text{MAP}}} \mathrm{d}z_\lambda(p) W_s(p)\\ &= (1-\lambda)\int_0^{p_{\text{MAP}}} \mathrm{d}z_0(p)W_s(p) + \lambda\int_0^{p_{\text{MAP}}} \mathrm{d}z_1(p)W_s(p)\\ &= (1-\lambda)\int_{\mathbb{R}} \mathrm{d}z W_s(p_0(z)) + \lambda\int_{\mathbb{R}} \mathrm{d}z W_s(p_1(z)).\end{aligned}$$

■

Thus, the function $\lambda \to \mathcal{W}_{\text{single}}[p_\lambda(\cdot)]$ is linear, hence convex.

B. Displacement Convexity of the Interaction-Potential Term

The proof of displacement convexity of the interaction potential term is more involved.

Proposition 5.2: Let $p_0(\cdot)$ and $p_1(\cdot)$ be in $\mathcal{S}'$ and let $p_\lambda(\cdot)$ the interpolating profile as defined in Section III. Then

$$\mathcal{W}_{\text{int}}[p_\lambda(\cdot)] \leq (1-\lambda)\mathcal{W}_{\text{int}}[p_0(\cdot)] + \lambda\mathcal{W}_{\text{int}}[p_1(\cdot)]. \tag{11}$$

Proof: Since p can be seen as a cdf we associate with it its probability measure μ such that $p(z) = p_{\text{MAP}} \int_{-\infty}^{z} \mathrm{d}\mu(x)$. Let us rewrite the interaction functional in the form

$$\mathcal{W}_{\text{int}}[p_\lambda(\cdot)] = \int_{\mathbb{R}} \mathrm{d}\mu_\lambda(x_1)\ldots\mathrm{d}\mu_\lambda(x_l)V(x_1,\ldots,x_l), \tag{12}$$

where $V(x_1,\ldots,x_l)$ is a totally symmetric "kernel function" that we will compute. There is an argument (see [8]) that allows to conclude (11) whenever V is jointly convex (in the usual sense). Let us briefly explain this argument here. Consider the measures μ_0, μ_1 associated to cdfs $p_0(\cdot)$, $p_1(\cdot)$. Then there exists a unique increasing map $T : \mathbb{R} \to \mathbb{R}$ such that $\mu_1 = T\#\mu_0$. Here $T\#\mu_0$ is the push-forward[1] of μ_0 under T. Then from $x_\lambda(p) = (1-\lambda)x_0(p) + \lambda x_1(p)$ we have that $\mu_\lambda = T_\lambda\#\mu_0$ where $T_\lambda(x) = \lambda x + (1-\lambda)T(x)$. Equation (12) can be written as

$$\begin{aligned}\mathcal{W}_{\text{int}}[p_\lambda(\cdot)] &= \int_{\mathbb{R}} \mathrm{d}\mu_0(x_1)\ldots\mathrm{d}\mu_0(x_l)V(T_\lambda(x_1),\ldots,T_\lambda(x_l))\\ &= l!\int_{S_x} \mathrm{d}\mu(x_1)\ldots\mathrm{d}\mu(x_l)V(T_\lambda(x_1),\ldots,T_\lambda(x_l)).\end{aligned} \tag{13}$$

In the second equality we restrict the integrals over the sector $S_x = \{\mathbf{x} = (x_1,\cdots,x_l) : x_i \geq x_j \text{ if } i < j\}$,

[1]Given a measurable map $T : \mathbb{R} \to \mathbb{R}$, the push-forward of μ under T is the measure $T\#\mu_0$ such that, for any bounded continuous function ϕ, $\int_{\mathbb{R}} \phi(T(x))\mathrm{d}\mu(x) = \int_{\mathbb{R}} \phi(x)\mathrm{d}(T\#\mu)(x)$.

which is possible since V is totally symmetric. Now it is important to notice that since T is an increasing map we have $T_\lambda(x_1) \geq \cdots \geq T_\lambda(x_l)$ for any $\lambda \in [0,1]$. Moreover the λ dependence in the kernel function is linear. Thus the proof of displacement convexity ultimately rests on checking that the kernel function is jointly convex in one sector, say S_x.[2] In fact the kernel function is translation invariant and can be expressed as a function of the distances $d_{1i} \equiv x_1 - x_i$, $i = 1, \ldots, l$. We will prove joint convexity of V as a function of these distances.

Now it remains to compute V and to investigate its joint convexity. With appropriate usage of Fubini's theorem and after some manipulations, we find

$$\mathcal{W}_{\rm int}[p(\cdot)] = \frac{\epsilon p_{\rm MAP}^l}{l} \int_{\mathbb{R}^l} \prod_{i=1}^{l} \mathrm{d}\mu_0(x_i) \Big\{ \int_{[0,1]^l} \prod_{i=1}^{l} \mathrm{d}u_i$$
$$\int_{\mathbb{R}} \mathrm{d}z \Big(\prod_{i=1}^{l} \theta(z - x_i) - \prod_{i=1}^{l} \theta(z - (x_i - u_i)) \Big) \Big\}, \quad (14)$$

where $\theta(x)$ denotes the Heaviside step function. So the kernel $V(x_1, \ldots, x_l)$ in (12) is the integrand of the first l integrals in (14). Our goal henceforth is to prove that V is convex in the usual sense. We will prove that in fact $V_{\mathbf{u}}$ is convex for all fixed $\mathbf{u}$, where $\mathbf{u} = (u_1, \ldots, u_l)$,

$$V_{\mathbf{u}}(\mathbf{x}) = \int_{\mathbb{R}} dz \Big(\prod_{i=1}^{l} \theta(z - x_i) - \prod_{i=1}^{l} \theta(z - (x_i - u_i)) \Big).$$

We recall here that we restrict our analysis to the sector of the space of variables S_x. Also, we remark that $\prod_{i=1}^{l} \theta(a_i) = \theta(\max_{i=1\ldots l} a_i)$. We observe that $V_{\mathbf{u}}$ can be written in terms of the distances $d_{1i} = x_1 - x_i$, $i = 2, \ldots, l$ as (here $d_{1i} \equiv 0$)

$$\begin{aligned} V_{\mathbf{u}}(\mathbf{x}) &= \int_{\mathbb{R}} \mathrm{d}z \{ \theta(z - x_1) - \theta(\max_{i=1\ldots l}(z - (x_i - u_i))) \} \\ &= - \min_{i=1\ldots l}(x_1 - x_i + u_i) = - \min_{i=1\ldots l}(d_{1i} + u_i) \end{aligned}$$

Lemma 5.3 below states that $V_{\mathbf{u}}$ is jointly convex in S_x for all fixed $\mathbf{u}$. This implies that $V(\mathbf{x})$ is jointly convex in S_x. This completes the proof. ■

Lemma 5.3: The function $f_{\mathbf{u}}(\mathbf{d}) = \min_i(d_{1i} + u_i)$ is concave in $\mathbf{d}$, where $\mathbf{d} = (d_{12}, \ldots, d_{1l})$ and $d_{1i} \equiv 0$.

Proof: Let $\mathbf{d}$ and $\mathbf{d}'$ be two instances of the argument of $f_{\mathbf{u}}$. Then, for $\lambda \in [0,1]$,

$$\begin{aligned} f_{\mathbf{u}}((1-\lambda)\mathbf{d} + \lambda \mathbf{d}') &= \min_i((1-\lambda)d_{1i} + \lambda d'_{1i} + u_i) \\ &= \min_i((1-\lambda)(d_{1i} + u_i) + \lambda(d'_{1i} + u_i)) \\ &\geq (1-\lambda)\min_i(d_{1i} + u_i) + \lambda \min_i(d'_{1i} + u_i) \\ &= (1-\lambda) f_{\mathbf{u}}(\mathbf{d}) + \lambda f_{\mathbf{u}}(\mathbf{d}'). \end{aligned}$$

This shows concavity.

■

[2]By symmetry, convexity in one sector implies convexity in other sectors. However this does not mean that convexity holds if arguments are taken in different sectors. And, indeed in the present problem one can check that convexity only holds within each sector.

C. Strict displacement convexity

We now prove that for $p_0(\cdot)$ and $p_1(\cdot)$ in $\mathcal{S}_0''$ the inequality (11) is strict whenever $\lambda \neq 0, 1$.

Since we already know that the kernel function $V(\mathbf{x})$ is convex in the sector S_x, it is sufficient to show it is *strictly convex as a function of the distances* $\mathbf{d} = (d_{12}, \ldots, d_{1l})$ on some *subset of positive measure* of the $\mathbf{d}$-space. This is sufficient because in (13) the kernel function is integrated against measure μ_0 with full support $\mathbb{R}$. Recall that μ_0 has full support because the profile $p_0(\cdot)$ belongs to $\mathcal{S}''$.

In Appendix F we give explicit formulas for V in terms of the distances d_{1i}. These formulas allow to prove that V is strictly convex in a subset of non-zero measure. Concretely, this subset is a small enough neighborhood of the origin $d_{1i} = 0$, $i = 2, \ldots, l$.

We remark that it is not easy to see that V is convex in the whole $\mathbf{d}$-space directly from these formulas (however we already know that V is convex by the method of proof of the previous section). In fact the formulas show that it is certainly not strictly convex when some of the distances become greater than 1.

VI. Proofs of Theorem 2.2 and Corollary 2.3

In this paragraph, for completeness, we wrap up the proofs of Theorem 2.2 and Corollary 2.3.

Proof of Theorem 2.2: Displacement convexity of $\mathcal{W}[p(\cdot)]$ in $\mathcal{S}'$ follows from Propositions 5.1 and 5.2. In section V-C we have shown that $\mathcal{W}_{\rm int}[p(\cdot)]$ is strictly displacement convex in $\mathcal{S}_0''$. Combining this with Proposition 5.1 immediately yields strict displacement convexity of $\mathcal{W}[p(\cdot)]$ in $\mathcal{S}_0''$. ■

Proof of Corollary 2.3: Uniqueness of the minimizer of $\mathcal{W}[p(\cdot)]$ in $\mathcal{S}_0''$ follows from strict displacement convexity. Indeed suppose there are two distinct minimizers $p_1(\cdot)$ and $p_2(\cdot)$ with $\mathcal{W}[p_1(\cdot)] = \mathcal{W}[p_2(\cdot)]$ and let $p_\lambda(\cdot)$ be the displacement interpolant. Then $\mathcal{W}[p_\lambda(\cdot)] < (1-\lambda)\mathcal{W}[p_1(\cdot)] + \lambda \mathcal{W}[p_2(\cdot)]$ (for $\lambda \neq 0, 1$) which implies $\mathcal{W}[p_\lambda(\cdot)] < \mathcal{W}[p_1(\cdot)]$. We must also have $\mathcal{W}[p_1(\cdot)] \leq \mathcal{W}[p_\lambda(\cdot)]$, hence $\mathcal{W}[p_1(\cdot)] < \mathcal{W}[p_\lambda(\cdot)]$ which is a contradiction.

Let us now show that all minimizers $p_1(\cdot) \in \mathcal{S}'$ are translates of the unique minimizer $p_0(\cdot) \in \mathcal{S}_0''$. We know from Theorem 4.4 that $p_1(\cdot) \in \mathcal{S}''$, i.e., it has to be strictly increasing. Thus there is a unique position, say z_1 such that $p_1(z_1) = p_{\rm MAP}/2$. Consider the set of profiles $\mathcal{S}_{z_1}''$ obtained by translating the set $\mathcal{S}_0''$ by the vector z_1. Clearly $p_1(\cdot)$ is the unique minimizer in $\mathcal{S}_{z_1}''$. But it is also clear that $\mathcal{W}[p_0(\cdot - z_1)] = \mathcal{W}[p_1(\cdot)]$. Thus, since $p_0(\cdot - z_1) \in \mathcal{S}_{z_1}''$, we must have $p_0(\cdot - z_1) = p_1(\cdot)$ as announced.

Finally let us discuss the consequences for the solutions of the DE equation (4). We show that in the space $\mathcal{S}'$ a solution of the DE equation is necessarily a minimum of $\mathcal{W}[p(\cdot)]$. This implies the statement of the theorem. Let $p_0(\cdot) \in \mathcal{S}'$ a solution of the DE equation. Consider any other profile $p_1(\cdot) \in \mathcal{S}'$, and consider the displacement interpolant $p_\lambda(\cdot)$. A computation of the derivative shows that $\frac{d}{d\lambda}\mathcal{W}[p_\lambda(\cdot)]|_{\lambda=0} = 0$ because $p_0(\cdot)$ is a solution of the DE equation. Since the map

$\lambda \to \mathcal{W}[p_\lambda(\cdot)]$ is convex, $\lambda = 0$ must be a minimum of this map. Thus $\mathcal{W}[p_\lambda(\cdot)] \geq \mathcal{W}[p_0(\cdot)]$ and in particular with $\lambda = 1$ we get $\mathcal{W}[p_1(\cdot)] \geq \mathcal{W}[p_0(\cdot)]$. Thus $p_0(\cdot)$ is a minimum of the functional in $\mathcal{S}'$. ∎

VII. Conclusion

In this paper, we demonstrate a new tool for the analysis of spatially coupled codes, namely the concept of displacement convexity. This tool makes use of an alternative structure of probability distributions and hence applies to an appropriate space of increasing profiles. We prove that the potential functional governing the (l, r)-regular ensemble is (strictly) convex under the alternative structure. This result implies that the potential functional admits a unique minimizing profile, or equivalently, that the DE equations governing the system admit a unique FP solution, in an appropriate space of profiles.

There are several questions that can be posed in this context. First, we recall that the original potential functional governing the system at hand is in discrete form. Can one extend the displacement convexity framework to the discrete setting? Displacement convexity can presumably be used to analyze a large range of problems with flavors similar to the present one. The generalization to irregular LDPC ensembles is immediate for pure displacement convexity (the question of strict convexity is however more subtle). It is interesting to consider more general one dimensional scalar recursions as in [3] and find out what are the general restrictions on the single system potential that still allow to prove displacement (strict) convexity. It also remains to be seen if these techniques can be applied to general BMS channels, the random K-SAT and Q-coloring problems to name a few. We plan to come back to these problems in the future.

Appendix A
Lower Boundedness of the Interaction Potential

Proof of Lemma 4.1: From Jensen's inequality,

$$\int_0^1 \mathrm{d}u\, p(z+u)^l \geq \Big(\int_0^1 \mathrm{d}u\, p(z+u)\Big)^l. \tag{15}$$

Further,

$$\int_{-M}^{M} \mathrm{d}z \int_0^1 \mathrm{d}u\, p(z+u)^l \overset{(a)}{=} \int_0^1 \mathrm{d}u \int_{-M+u}^{M+u} \mathrm{d}z'\, p(z')^l$$
$$= \int_0^1 \mathrm{d}u \Big(\int_{-M+u}^{-M} \mathrm{d}z' p(z')^l + \int_{-M}^{M} \mathrm{d}z' p(z')^l + \int_{M}^{M+u} \mathrm{d}z' p(z')^l \Big),$$

where (a) is obtained by first changing the order of integration (which is admissible since the integral converges) and then making the change of variable $z' = z + u$. And so, by combining this identity with (15) we obtain

$$\int_{-M}^{M} \mathrm{d}z \left\{ p(z)^l - \Big(\int_0^1 \mathrm{d}u\, p(z+u)\Big)^l \right\}$$
$$+ \int_0^1 \mathrm{d}u \int_{-M+u}^{-M} \mathrm{d}z'\, p(z')^l + \int_0^1 \mathrm{d}u \int_{M}^{M+u} \mathrm{d}z'\, p(z')^l \geq 0.$$

Now we take the limit $M \to +\infty$ for each term of this inequality. By an application of Lebesgue's dominated convergence theorem, the last two terms tend to zero and $\frac{1}{2} p_{\text{MAP}}^l$, respectively. Therefore the limit of the first term is bounded from below by $-\frac{1}{2} p_{\text{MAP}}^l$, which concludes the proof. ∎

Appendix B
Truncation of Profiles

Proof of Lemma 4.2: It is easy to prove that a truncation of $p(z)$ at p_{MAP} yields a smaller value for the single system potential $\mathcal{W}_s(p(z))$ (see e.g. the Figure I for an intuition). Therefore we have $\mathcal{W}_{\text{single}}[p(\cdot)] \geq \mathcal{W}_{\text{single}}[\bar{p}(\cdot)]$.

We now treat the functional corresponding to the interaction term. We define the function g as $g(z) = p(z) - \bar{p}(z)$ and notice that:

$$\begin{cases} p(z) \leq p_{\text{MAP}} \Rightarrow g(z) = 0 \text{ and } \bar{p}(z) = p(z), \\ p(z) > p_{\text{MAP}} \Rightarrow g(z) > 0 \text{ and } \bar{p}(z) = p_{\text{MAP}}. \end{cases} \tag{16}$$

We need to show that $\mathcal{W}_{\text{int}}[\bar{p}(\cdot)] \leq \mathcal{W}_{\text{int}}[p(\cdot)]$, or equivalently that:

$$\int_{\mathbb{R}} \mathrm{d}z \left\{ \bar{p}(z)^l - \Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u) \Big)^l \right\}$$
$$\leq \int_{\mathbb{R}} \mathrm{d}z \left\{ (\bar{p}(z) + g(z))^l - \Big(\int_0^1 \mathrm{d}u\, (\bar{p}(z+u) + g(z+u)) \Big)^l \right\}.$$

Using the binomial expansion this is equivalent to

$$\sum_{i=0}^{l-1} \binom{l}{i} \int_{\mathbb{R}} \mathrm{d}z \left\{ \bar{p}(z)^i g(z)^{l-i} \right.$$
$$\left. - \Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u) \Big)^i \Big(\int_0^1 \mathrm{d}u\, g(z+u) \Big)^{l-i} \right\} \geq 0.$$

In the following steps, we show that the integral inside the summation above is positive for any fixed value of i; the inequality follows directly. We see that:

$$\Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u) \Big)^i \Big(\int_0^1 \mathrm{d}u\, g(z+u) \Big)^{l-i}$$
$$\leq p_{\text{MAP}}^i \Big(\int_0^1 \mathrm{d}u\, g(z+u) \Big)^{l-i} \leq p_{\text{MAP}}^i \int_0^1 \mathrm{d}u\, g(z+u)^{l-i},$$

where the first inequality is due to the property $\bar{p}(z) \leq p_{\text{MAP}}$ and the second is using the convexity of the function $f(g) =$

$g^l;\ g \geq 0$. We integrate over z, then make the change of variable $z' = z + u$ on the right-hand side to obtain:

$$\int_{\mathbb{R}} \mathrm{d}z \Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u)\Big)^i \Big(\int_0^1 \mathrm{d}u\, g(z+u)\Big)^{l-i}$$
$$\leq \int_{\mathbb{R}} \mathrm{d}z\, p^i_{\text{MAP}} \int_0^1 \mathrm{d}u\, g(z+u)^{l-i} = \int_{\mathbb{R}} \mathrm{d}z'\, p^i_{\text{MAP}} g(z')^{l-i} \tag{17}$$

Using the properties of g in (16), we remark that $\int_{\mathbb{R}} \mathrm{d}z\, p^i_{\text{MAP}} g(z)^{l-i} = \int_{\mathbb{R}} \mathrm{d}z\, \bar{p}(z)^i\, g(z)^{l-i}$ and so the difference of quantities in the inequality (17) is integrable, and thus we obtain:

$$\int_{\mathbb{R}} \mathrm{d}z \Big(\int_0^1 \mathrm{d}u\, \bar{p}(z+u)\Big)^i \Big(\int_0^1 \mathrm{d}u\, g(z+u)\Big)^{l-i} \leq \int_{\mathbb{R}} \mathrm{d}z\, \bar{p}(z)^i g(z)^{l-i}$$

for any i. This yields the desired result $\mathcal{W}_{\text{int}}[p(\cdot)] \geq \mathcal{W}_{\text{int}}[\bar{p}(\cdot)]$. ■

Appendix C
Rearrangement of Profiles

Before proceeding with the proof of Lemma 4.3, we state a general rearrangement inequality of Brascamp, Lieb and Luttinger [12].

Theorem C.1: Let $f_j,\ 1 < j < k$ be nonnegative measurable functions on $\mathbb{R}$, and let $a_{jm},\ 1 < j < k,\ 1 < m < n$, be real numbers. Then, if f^* is the symmetric decreasing rearrangement of f, we have:

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \prod_{j=1}^k f_j\left(\sum_{m=1}^n a_{jm} x_m\right) \leq \int_{\mathbb{R}^n} \mathrm{d}^n x \prod_{j=1}^k f_j^*\left(\sum_{m=1}^n a_{jm} x_m\right) \tag{18}$$

Remark Theorem C.1 is nontrivial only if $k > n$. Otherwise, both integrals diverge and the inequality trivially holds. We will see in this section that $k > n$ in our case.

Proof of Lemma 4.3: It is sufficient to prove that the increasing rearrangement of a profile decreases $\mathcal{W}_{\text{int}}[p(\cdot)]$, since $\mathcal{W}_{\text{single}}[p(\cdot)]$ is invariant under rearrangement.

Theorem C.1 applies to symmetric decreasing rearrangements. Therefore it is convenient to first "symmetrize" the profile and the functional. Consider a profile $p(\cdot) \in \mathcal{S}$ such that $p(z) \in [0, p_{\text{MAP}}]$ (due to Lemma 4.2) and denote by $\hat{p}(\cdot)$ the function such that $\hat{p}(z) = p(z),\ z < R$ and $\hat{p}(z) = \hat{p}(2R - z),\ z > R$. The value R is chosen (large enough) so that $p(R)$ is arbitrarily close to p_{MAP}. Note that $\hat{p}(\cdot)$ is integrable over $\mathbb{R}$.

We recall the expression of $\mathcal{W}_{\text{int}}[p(\cdot)]$ in (6) and rewrite it as:

$$\mathcal{W}_{\text{int}}[p(\cdot)] = \frac{\epsilon}{l} \lim_{R\to+\infty} \Big\{ \int_{-\infty}^R \mathrm{d}z p(z)^l - \int_{-\infty}^R \mathrm{d}z \big(\int_0^1 \mathrm{d}u\ p(z+u)\big)^l \Big\}. \tag{19}$$

We now express both integrals in the bracket in terms of the symmetrized profile. For the first one, this is immediate

$$\int_{-\infty}^R \mathrm{d}z p(z)^l = \frac{1}{2} \int_{\mathbb{R}} \mathrm{d}z \hat{p}(z)^l. \tag{20}$$

For the second one, some care has to be taken with the averaging over u when z is near R. One has

$$\int_{-\infty}^R \mathrm{d}z \big(\int_0^1 \mathrm{d}u\ p(z+u)\big)^l = \frac{1}{2} \int_{\mathbb{R}} \mathrm{d}z \big(\int_0^1 \mathrm{d}u\ \hat{p}(z+u)\big)^l + o\big(\frac{1}{R^l}\big). \tag{21}$$

Replacing these two formulas in (19) we have the representation

$$\mathcal{W}_{\text{int}}[p(\cdot)] = \frac{\epsilon}{2l} \int_{\mathbb{R}} \mathrm{d}z \hat{p}(z)^l - \frac{\epsilon}{2l} \int_{\mathbb{R}} \mathrm{d}z \big(\int_0^1 \mathrm{d}u\ \hat{p}(z+u)\big)^l. \tag{22}$$

Now consider $\hat{p}^*(\cdot)$, the symmetric decreasing rearrangement of $\hat{p}(\cdot)$. The first term in (22) is invariant under rearrangement. It remains to prove that the second term in (22) increases upon rearrangement. We express it as follows (dropping $\epsilon/2l$):

$$\begin{aligned}
&\int_{\mathbb{R}} \mathrm{d}z \Big(\int_0^1 \mathrm{d}u\ \hat{p}(z+u)\Big)^l \\
&= \int_{\mathbb{R}} \mathrm{d}z \int_{\mathbb{R}^l} \prod_{i=1}^l \mathrm{d}u_i\ \hat{p}(z+u_i) \mathbb{1}_{[0,1]}(u_i) \\
&\overset{(b)}{=} \int_{\mathbb{R}} \mathrm{d}z' \int_{\mathbb{R}^l} \prod_{i=1}^l \mathrm{d}u_i'\ \hat{p}(z' + R + u_i' + \frac{1}{2})\, \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(u_i') \\
&\overset{(c)}{\leq} \int_{\mathbb{R}} \mathrm{d}z' \int_{\mathbb{R}^l} \prod_{i=1}^l \mathrm{d}u_i'\ \hat{p}^*(z' + R + u_i' + \frac{1}{2})\, \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(u_i') \\
&\overset{(d)}{=} \int_{\mathbb{R}} \mathrm{d}z \int_{\mathbb{R}^l} \prod_{i=1}^l \mathrm{d}u_i\ \hat{p}^*(z+u_i)\, \mathbb{1}_{[0,1]}(u_i) \\
&= \int_{\mathbb{R}} \mathrm{d}z \Big(\int_0^1 \mathrm{d}u\ \hat{p}^*(z+u)\Big)^l,
\end{aligned}$$

where the equality in (b) is due to the changes of variables $z' = z - R$ and $u_i' = u_i - \frac{1}{2};\ i = 1 \ldots l$, the inequality in (c) is due Theorem C.1, and the equality in (d) is obtained by first remarking that the indicator function $\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(u_i')$ is unchanged upon rearrangement and then by making the reverse changes of variables $z = z' + R$ and $u_i = u_i' + \frac{1}{2};\ i = 1 \ldots l$. So far we have obtained

$$\mathcal{W}_{\text{int}}[p(\cdot)] \geq \frac{\epsilon}{2l} \int_{\mathbb{R}} \mathrm{d}z \hat{p}^*(z)^l - \frac{\epsilon}{2l} \int_{\mathbb{R}} \mathrm{d}z \Big(\int_0^1 \mathrm{d}u\ \hat{p}^*(z+u)\Big)^l.$$

To obtain $\mathcal{W}_{\mathrm{int}}[p(\cdot)] \geq \mathcal{W}_{\mathrm{int}}[p^*(\cdot)]$ it remains to reverse the steps (19)-(22). ∎

Appendix D
Strict Monotonicity of Profile

We establish Lemma 4.4 as a corollary of the following lemmas.

Lemma D.1: If $p(\cdot)$ minimizes $\mathcal{W}[p(\cdot)]$, then it satisfies the DE equation.

Proof: Consider a profile $p(\cdot)$ and a function $\nu(\cdot)$ such that $\lim_{z\to\pm\infty} \nu(z) = 0$. We compute the directional derivative of the potential $\mathcal{W}[p(\cdot)]$ in the direction of $\nu(\cdot)$,

$$\mathrm{d}\mathcal{W}[p(\cdot)][\nu] = \lim_{\delta\to 0} \frac{\mathcal{W}[p(\cdot)+\delta\nu(\cdot)] - \mathcal{W}[p(\cdot)]}{\delta}.$$

A calculation gives

$$\mathrm{d}\mathcal{W}[p(\cdot)][\nu] = \int_{\mathbb{R}} \mathrm{d}z\, \nu(z) \Big\{ 1 - (1-p(z))^{\frac{1}{r-1}} - \epsilon \int_0^1 \mathrm{d}v \Big(\int_0^1 \mathrm{d}u\, p(z+u-v) \Big)^{l-1} \Big\}.$$

Now consider the function

$$\nu_p(z) = -\Big\{ 1 - (1-p(z))^{\frac{1}{r-1}} - \epsilon \int_0^1 \mathrm{d}v \Big(\int_0^1 \mathrm{d}u\, p(z+u-v) \Big)^{l-1} \Big\}.$$

The directional derivative of $\mathcal{W}[p(\cdot)]$ in the direction of $\nu_p(\cdot)$ satisfies

$$\mathcal{W}[p(\cdot)][\nu_p] \leq 0 \tag{23}$$

because the integrand is a square. Now assume that $p(\cdot)$ is a minimizing profile. In the case where equality is met in (23), $p(\cdot)$ satisfies the DE equation. Consider the case where the inequality is strict. Then,

$$\mathrm{d}\mathcal{W}[p(\cdot)][\nu_p] = \lim_{\delta\to 0} \frac{\mathcal{W}[p(\cdot)+\delta\nu_p(\cdot)] - \mathcal{W}[p(\cdot)]}{\delta} < 0,$$

and we can find δ_0 small enough such that

$$\mathcal{W}[p(\cdot)+\delta_0\nu_p(\cdot)] < \mathcal{W}[p(\cdot)].$$

So $p(\cdot)$ cannot be a minimizing profile, and this concludes the proof by contradiction. ∎

Lemma D.2: If $p(\cdot)$ is increasing and satisfies the DE equation, then it cannot be strictly flat on an interval $]a,b[\subset \mathbb{R}$.

Proof: If $p(\cdot)$ satisfies the DE equation, then

$$1-(1-p(z))^{\frac{1}{r-1}} = \epsilon \int_0^1 \mathrm{d}v \Big(\int_0^1 \mathrm{d}u\, p(z+u-v) \Big)^{l-1}$$

By taking the derivative on each side, we find that

$$(1-p(z))^{\frac{1}{r-1}-1} \frac{1}{r-1} p'(z) = \epsilon \int_0^1 \mathrm{d}v\, (l-1) \times \Big(\int_0^1 \mathrm{d}u\, p(z+u-v) \Big)^{l-2} \int_0^1 \mathrm{d}w\, p'(z+w-v). \tag{24}$$

Notice that $\int_0^1 \mathrm{d}w\, p'(z+w-v) = \int_0^1 \mathrm{d}w\, \frac{\mathrm{d}}{\mathrm{d}w} p(z+w-v) = p(z+1-v) - p(z-v)$.

Now assume that there exists a flat spot of $p(\cdot)$ for $z \in]a,b[$ where it takes some value p_{flat}. We consider "maximal" intervals $]a,b[$ such that $p(\cdot)$ takes values different than p_{flat} for all $z \notin]a,b[$. On this flat spot, (24) becomes

$$0 = \int_0^1 \mathrm{d}v \Big(\int_0^1 \mathrm{d}u\, p(z+u-v) \Big)^{l-2} (p(z+1-v) - p(z-v)). \tag{25}$$

We will now show that this equality cannot be satisfied.

Let us first consider the case when a and b are finite. Since $]a,b[$ is maximal and $p(\cdot)$ is increasing, we know that $0 < p_{\mathrm{flat}} < p_{\mathrm{MAP}}$, $p(z) < p_{\mathrm{flat}}$ for all $z < a$ and $p(z) > p_{\mathrm{flat}}$ for all $z > b$. Now let us fix $z \in [b-1+\delta_0, b]$, where $0 < \delta_0 < 1$. For such a z, the equality (25) holds. But for such z and for all $0 \leq v < 1$, $\int_0^1 \mathrm{d}u\, p(z+u-v) \geq \int_{u>v} \mathrm{d}u\, p(z+u-v) \geq (1-v)p_{\mathrm{flat}} > 0$. Thus we should have $p(z+1-v) = p(z-v)$ for a.e $v \in [0,1]$. This is not possible. Indeed, take $v \in [0,\delta_1]$ with $0 < \delta_1 < \delta_0$ is small enough so that $z-v < b < z+1-v$ and thus $p(z+1-v) - p(z-v) > p(b+\delta_0-\delta_1) - p(b) > 0$. These arguments prove that an increasing solution of DE cannot be flat for $z \in [b-1+\delta_0, b]$. We repeat the argument on $[b-k+\delta_0, b-(k-1)+\delta_0]$ for all $1 < k < K$ such that $b-K+\delta_0 < a$ and find that an increasing solution of DE cannot be flat on each of those intervals, and thus on $[b-(K-1)+\delta_0, b]$. Finally, we repeat the argument on the last interval $[a, b-(K-1)+\delta_0]$ and deduce that an increasing solution of DE cannot be flat on $[a,b]$.

Next, we consider the case when $a = -\infty$. In this case, we have that $p(z) = 0$ for all $z \leq b$ and $p(z) > 0$ for all $z > b$. The analysis is similar than the preceding one. First fix z in the interval $\in [b-1+\delta_0, b]$, $0 < \delta_0 < 1$. Equation (25) is satisfied for such z. Now take $v \in [0,\delta_1]$ with $0 < 2\delta_1 < \delta_0$. Then $p(z+1-v) - p(z-v) = p(z+1-v) > p(b+\delta_0-\delta_1) > 0$, and for $1-\delta_0+2\delta_1 < u < 1$ we have $p(z+u-v) > p(b+\delta_1)$ so $\int_0^1 du\, p(z+u-v) > (1+\delta_0-2\delta_1)p(b+\delta_1) > 0$. Thus the right hand side of (25) does not vanish which is a contradiction. We carry out the same analysis as above for $z \in [b-k+\delta_0, b-(k-1)+\delta_0]$ for $k \in \mathbb{N}$, and thus deduce that $p(\cdot)$ cannot be flat on $[-\infty, b]$.

Finally, consider the case when $b = +\infty$. The analysis is essentially symmetric to the preceding one. In this case we have $p(z) = p_{\mathrm{MAP}}$ for $z > a$ and $p(z) < p_{\mathrm{MAP}}$ for $z < a$. First fix z in the interval $[a, a+1-\delta_0]$. For such z (25) holds. For $v \in [1-\delta_1, 1]$ with $\delta_1 < \delta_0$ we have $p(z+1-v) - p(z-v) > p_{\mathrm{MAP}} - p(a-\delta_0+\delta_1) > 0$. Moreover it is clear that $\int_0^1 du\, p(z+u-v) > 0$. So the right hand side of (25) cannot vanish, and we arrive at a contradiction. We repeat the argument for $z \in [a+k-\delta_0, a+(k+1)-\delta_0]$ for $k \in \mathbb{N}$, and conclude that $p(\cdot)$ cannot be flat on $[a, +\infty]$.

∎

Proof of Lemma 4.4: The proposition follows directly from Lemmas D.1 and D.2. ∎

We now discuss a lemma that is interesting in itself but not necessary for our results.

Lemma D.3: If $p(\cdot)$ minimizes $\mathcal{W}[p(\cdot)]$ then it cannot have a flat spot, i.e., $p(z) = p_{\text{flat}}$ with $0 < p_{\text{flat}} < p_{\text{MAP}}$ in a bounded interval $z \in [a, b]$ such that $b - a > 1$.

Proof: Suppose that $p(\cdot)$ is an increasing minimizing profile and that it has a constant value $0 < p_{\text{flat}} < p_{\text{MAP}}$ on a bounded interval of length greater than 1. We will construct another profile that has strictly less energy.

We start by expressing the single potential as follows

$$\begin{aligned}\mathcal{W}_{\text{single}}[p(\cdot)] &= \int_{-\infty}^{a} \mathrm{d}z\, W_{\mathrm{s}}(p(z)) + \int_{a}^{b-1} \mathrm{d}z\, W_{\mathrm{s}}(p(z)) \\ &\quad + \int_{b-1}^{+\infty} \mathrm{d}z\, W_{\mathrm{s}}(p(z)) \\ &= \int_{-\infty}^{a} \mathrm{d}z\, W_{\mathrm{s}}(p(z)) + (b-1-a) W_{\mathrm{s}}(p(a)) \\ &\quad + \int_{b-1}^{+\infty} \mathrm{d}z\, W_{\mathrm{s}}(p(z)).\end{aligned}$$

By applying the change of variables $z' = z - (b-1) + a$ on the rightmost integral, we express it as

$$\int_{b-1}^{+\infty} \mathrm{d}z\, W_{\mathrm{s}}(p(z)) = \int_{a}^{+\infty} \mathrm{d}z\, W_{\mathrm{s}}(p(z+b-1-a)).$$

We define the profile $\tilde{p}$ by

$$\tilde{p}(z) = \begin{cases} p(z) & \text{if} \quad z \le a, \\ p(z+b-1-a) & \text{if} \quad z > a, \end{cases}$$

We thus obtain

$$\mathcal{W}_{\text{single}}[p(\cdot)] = \mathcal{W}_{\text{single}}[\tilde{p}(\cdot)] + (b-1-a) W_{\mathrm{s}}(p_{\text{flat}}).$$

Note that since $0 < p_{\text{flat}} < p_{\text{MAP}}$ we have $W_{\mathrm{s}}(p_{\text{flat}}) > 0$. Thus $\mathcal{W}_{\text{single}}[p(\cdot)] > \mathcal{W}_{\text{single}}[\tilde{p}(\cdot)]$.

For the interaction potential, we prove that $\mathcal{W}_{\text{int}}[p(\cdot)] = \mathcal{W}_{\text{int}}[\tilde{p}(\cdot)]$. Indeed,

$$\begin{aligned}\mathcal{W}_{\text{int}}[p(\cdot)] &= \frac{\epsilon}{l} \int_{\mathbb{R}} \mathrm{d}z \Big\{ p(z)^l - \Big(\int_0^1 \mathrm{d}u\, p(z+u) \Big)^l \Big\} \\ &= \frac{\epsilon}{l} \int_{-\infty}^{a} \mathrm{d}z \Big\{ p(z)^l - \Big(\int_0^1 \mathrm{d}u\, p(z+u) \Big)^l \Big\} && (26) \\ &+ \frac{\epsilon}{l} \int_{a}^{b-1} \mathrm{d}z \Big\{ p(z)^l - \Big(\int_0^1 \mathrm{d}u\, p(z+u) \Big)^l \Big\} && (27) \\ &+ \frac{\epsilon}{l} \int_{b-1}^{+\infty} \mathrm{d}z \Big\{ p(z)^l - \Big(\int_0^1 \mathrm{d}u\, p(z+u) \Big)^l \Big\}. && (28)\end{aligned}$$

We use the same definition of $\tilde{p}(\cdot)$ as above, and denote the functionals in (26), (27), and (28) by $\mathcal{T}_1[p(\cdot)]$, $\mathcal{T}_2[p(\cdot)]$, and $\mathcal{T}_3[p(\cdot)]$ respectively. Observe that

- $\mathcal{T}_1[p(\cdot)] = \mathcal{T}_1[\tilde{p}(\cdot)]$ since

$$\begin{aligned} p(z) &= \tilde{p}(z) \quad \text{if} \quad z < a \\ p(z+u) &= \tilde{p}(z+u) \quad \text{if} \quad z < a, \text{ and } 0 < u < 1. \end{aligned}$$

- $\mathcal{T}_2[p(\cdot)] = 0$ since $p(z) = p(z+u) = p_{\text{flat}}$ when $a < z < b-1$.
- $\mathcal{T}_3[p(\cdot)] = \mathcal{T}_2[\tilde{p}(\cdot)] + \mathcal{T}_3[\tilde{p}(\cdot)]$ since, by the change of variables $z' = z - (b-1) + a$, we have

$$\begin{aligned}\mathcal{T}_3[p(\cdot)] &= \frac{\epsilon}{l} \int_{a}^{+\infty} \mathrm{d}z \Big\{ p(z+b-1-a)^l \\ &\qquad - \Big(\int_0^1 \mathrm{d}u\, p(z+b-1-a+u) \Big)^l \Big\} \\ &= \frac{\epsilon}{l} \int_{a}^{+\infty} \mathrm{d}z \Big\{ \tilde{p}(z)^l - \Big(\int_0^1 \mathrm{d}u\, \tilde{p}(z+u) \Big)^l \Big\}.\end{aligned}$$

Thus $\mathcal{T}_1[p(\cdot)] + \mathcal{T}_2[p(\cdot)] + \mathcal{T}_3[p(\cdot)] = \mathcal{T}_1[\tilde{p}(\cdot)] + \mathcal{T}_2[\tilde{p}(\cdot)] + \mathcal{T}_3[\tilde{p}(\cdot)]$

Combining these results we get

$$\mathcal{W}[p(\cdot)] = \mathcal{W}[\tilde{p}(\cdot)] + (b-1-a)\, W_{\mathrm{s}}(p_{\text{flat}}) > \mathcal{W}[\tilde{p}(\cdot)]$$

■

Appendix E
Tightness of the Minimizing Sequence

Proof of Lemma 4.5: Consider a minimizing sequence of cdfs $p_n(\cdot)$, i.e., satisfying (9). Fix any $\delta > 0$ and suppose that

$$p_n(M) - p_n(-M) < (1-\delta) p_{\text{MAP}}. \qquad (29)$$

We will show that (29) implies that necessarily $M \le c/\delta^2$ for a fixed constant $c > 0$. Taking the contrapositive we find that: choosing $M_\delta = c'/\delta^2$ with $c' > c$ implies that any minimizing sequence satisfies $p_n(M_\delta) - p_n(-M_\delta) > (1-\delta) p_{\text{MAP}}$.

From Lemma 4.1 we have

$$\mathcal{W}[p_n(\cdot)] \ge \mathcal{W}_{\text{single}}[p_n(\cdot)] - \frac{\epsilon p_{\text{MAP}}^l}{2l} \ge \int_{-M}^{M} dz\, W_s(p_n(z)) - \frac{\epsilon p_{\text{MAP}}^l}{2l}. \qquad (30)$$

Now, assuming (29) there must be a mass at least δp_{MAP} outside of the interval $[-M, M]$. Thus we have $p_n(-M) \ge \delta p_{\text{MAP}}/2$ or $p_{\text{MAP}} - p_n(M) \ge \delta p_{\text{MAP}}/2$. Recall that $p_n \in \mathcal{S}_0''$ so $p_n(0) = p_{\text{MAP}}/2$. Therefore in $[-M, 0]$ or in $[0, M]$ the profile $p_n(z)$ must be $\delta p_{\text{MAP}}/2$ away from the minima 0 and p_{MAP} of W_s. Moreover, one can check that $W_s(p)$ has a parabolic shape near the minima at 0 and p_{MAP} so that away from these minima $W_s(p) \ge C\delta^2 p_{\text{MAP}}^2/4$ for a constant $C > 0$ depending only on l. These remarks imply

$$\int_{-M}^{M} dz\, W_s(p_n(z)) \ge \frac{1}{4} M C \delta^2 p_{\text{MAP}}^2. \qquad (31)$$

Since $p_n(\cdot)$ is a minimizing sequence, for n large enough its cost must be smaller than the cost of a fixed reference profile, say $\rho(z) = 0, z \le 0$, $\rho(z) = p_{\text{MAP}}, z > 0$. More formally,

$$\mathcal{W}[p_n(\cdot)] < \mathcal{W}[\rho(\cdot)] = -\frac{\epsilon}{l(l+1)} p_{\text{MAP}}^l. \qquad (32)$$

Finally, combining (30), (31) and (32) we find that

$$M \le 2 \frac{\epsilon(l-1)}{l(l+1)} \frac{p_{\text{MAP}}^{l-2}}{C\delta^2}. \qquad (33)$$

■

APPENDIX F
EXPLICIT EXPRESSIONS OF KERNEL FUNCTION

In this section, we compute the kernel function V of section V-B and illustrate some of its properties. In particular we show that it is strictly convex in a set of positive measure.

Recall that the function is totally symmetric under permutations. It is therefore enough to compute it in a fixed sector $S_x = \{\mathbf{x} = (x_1, \cdots, x_l) : x_i \geq x_j \text{ if } i < j\}$. We express V in terms of the distances d_{1i}, which are ordered such that $d_{1i} < d_{1j}$ if $i < j$.

Let us first discuss the explicit examples $l = 2$ and $l = 3$. For $l = 2$ an explicit computation yields,

$$V_{(l=2)}(d_{12}) = \begin{cases} -\frac{1}{2} & \text{if} \quad d_{12} \geq 1, \\ -\frac{1}{2} + \frac{1}{6}(1 - d_{12})^3 & \text{if} \quad d_{12} < 1. \end{cases}$$

By taking the second derivative it is easy to see that $V_{(l=2)}$ is convex everywhere, and strictly convex for $d_{12} < 1$.

For $l = 3$, we have $d_{12} < d_{13}$ and the computation yields,

$$V_{(l=3)}(d_{12}, d_{13}) = \begin{cases} V_{(l=2)}(d_{12}) & \text{if } d_{13} \geq 1, \\ \\ -\frac{1}{2} + \frac{1}{6}(1 - d_{12})^3 & \\ +\frac{1}{12}(1 - d_{13})^4 & \text{if } d_{13} < 1. \\ +\frac{1}{6}d_{12}(1 - d_{13})^3 & \end{cases}$$

For $d_{13} < 1$ the Hessian is

$$\begin{pmatrix} 1 - d_{12} & -\dfrac{1}{2}(d_{13} - 1)^2 \\ -\dfrac{1}{2}(d_{13} - 1)^2 & -(d_{13} - 1)(1 + d_{12} - d_{13}) \end{pmatrix}$$

and the corresponding eigenvalues are:

$$\lambda_{1,2} = \frac{1}{2}\{2 - 2d_{13} - d_{12}d_{13} + d_{13}^2 \pm \sqrt{\Delta}\},$$

where

$$\begin{aligned} \Delta = 1 &+ 4d_{12}^2 - 4d_{13} - 8d_{12}d_{13} - 4d_{12}^2 d_{13} + 10 d_{13}^2 \\ &+ 8d_{12}d_{13}^2 + d_{12}^2 d_{13}^2 - 8d_{13}^3 - 2d_{12}d_{13}^3 + 2d_{13}^4. \end{aligned}$$

A plot of the eigenvalues shows that they are non-negative in the region $0 \leq d_{12} \leq d_{13} \leq 1$. In fact, one eigenvalue is strictly positive everywhere in this region, and the other is strictly positive everywhere in this region except at the boundary $d_{13} = 1$, where it becomes equal to zero. This is consistent with the fact that $V_{(l=3)}(d_{12}, d_{13}) = V_{(l=2)}(d_{12})$ when $d_{13} \geq 1$. For $d_{13} \geq 1$, the Hessian always has a vanishing eigenvalue, and a strictly positive one when $d_{12} < 1$. For $d_{12} \geq 1$, the kernel $V_{(l=3)}(d_{12}, d_{13})$ is constant and both eigenvalues vanish. To summarize the kernel is always convex, and strictly convex for $0 \leq d_{13} < 1$.

These results can be generalized for all l. We find the general expression of the kernel

$$V_l(d_{12}, \cdots, d_{1l}) = \sum_{k=2}^{l} \sum_{m=k}^{l} \frac{(1 - d_{1m})^{m-k+3}}{(m-k+3)(m-k+2)} \times \Big(\sum_{\mathcal{S} \subseteq \{2, \dots, m-1\}} \prod_{n \in \mathcal{S}} d_{1n} \Big).$$

The corresponding Hessian (H_{ij}) is a symmetric matrix of dimension $(l-1) \times (l-1)$ with matrix elements that are polynomials in d_{1i}, $i = 1, \dots, l$. In particular, at $d_{1i} = 0$ for all i we have $H_{ii} = 1$ and $H_{ij} = -\frac{1}{j+1} + \sum_{m=j+1}^{l} \frac{1}{(m-1)(m-2)} = -\frac{1}{l-1}$; $j > i$. Defining $\mathbf{v}$ as the $(l-1)$-dimensional vector of 1's and denoting by $\mathbb{1}$ the $(l-1)$-dimensional identity matrix, we remark that H at the origin can be expressed as

$$H = (1 + \frac{1}{l-1})\mathbb{1} - \frac{1}{l-1}\mathbf{v}\mathbf{v}^{\mathrm{T}}.$$

The eigenvalues of this matrix are $1 + \frac{1}{l-1}$ and 1 (with $1 + \frac{1}{l-1}$ having degeneracy $l-2$). Since these eigenvalues are strictly positive, the Hessian is strictly positive definite at the origin, and thus (by continuity) also in a small neighborhood of the origin. Thus V is a strictly convex function of d_{1i}, $i = 2, \dots, l$ in a small neighborhood of the origin.

Acknowledgment. We thank Marc Vuffray for discussions on rearrangement inequalities and for bringing reference [12] to our attention. We in particular thank Tom Richardson for pointing out the subtle way in which uniqueness can fail in a system despite the strict displacement convexity of its underlying functional.

REFERENCES

[1] A.J. Felstrom and K.S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," IEEE Trans. Inform. Theory, vol. 45, no. 5, (1999) pp. 2181-2190.

[2] S. Kudekar, T. Richardson, R. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC", in *Information Theory Proceedings (ISIT)*, (2010) pp. 684-688.

[3] A. Yedla, Y. Jian, P. Nguyen, and H. Pfister, "A simple proof of threshold saturation for coupled scalar recursions," e-print: arXiv: 1204.5703[cs] (2012).

[4] S. Kudekar, T. Richardson, R. Urbanke, "Wave-Like Solutions of General One-Dimensional Spatially Coupled Systems," arXiv: 1208.5273[cs] (2012)

[5] S. Kudekar, T. Richardson, and R. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," arXiv: 1201.2999. (2012)

[6] S. Kumar, A.J. Young, N. Macris, H.D. Pfister, "A Proof of Threshold Saturation for Spatially-Coupled LDPC Codes on BMS Channels," in *Allerton,* Monticello, Il, (2012).

[7] R. McCann, "A convexity principle for interacting gases," Adv. Math. 128(1), (1997) pp. 153-179.

[8] C. Villani, "Displacement Interpolation and Displacement Convexity," in *Topics in Optimal Transportation,* Rhode Island: American Mathematical Society (2003), chapter 5, pp. 143-165.

[9] G. Alberti and G. Bellettini, "A nonlocal anisotropic model for phase transitions Part I: The optimal profile problem," Math. Annalen. 310, (1998) pp. 527-560.

[10] E.A. Carlen, M.C. Carvalho, R. Esposito, J.L. Lebowitz, and R. Marra, "Displacement Convexity and Minimal Fronts at Phase Boundaries," Arch. Rational Mech. Anal. 194, (2009) pp. 823-847.

[11] S.H. Hassani, N. Macris, R. Urbanke, "Chains of Mean-Field Models," Journal of Statistical Mechanics: Theory and Experiment (2012) (02), P02011.

[12] H.J. Brascamp, E.H. Lieb, and J.M. Luttinger. “A general rearrangement inequality for multiple integrals,” J. Funct. Anal., vo. 17, (1974) pp. 227-237.
[13] B. Dacorogna, *Direct Methods in the Calculus of Variations*, 2nd ed. New York: Springer-Verlag, (1992).
[14] I. Fonseca, G. Leoni, *Modern Methods in the Calculus of Variations: Lp Spaces*, 1st ed. New York: Springer, (2007).
[15] A. Burchard, ”Cases of equality in the Riesz rearrangement inequality,” Annals of Mathematics, 143, (1996) pp. 499-527.
[16] M. Christ, T. Flock, ”Cases of equality in certain multilinear inequalities of Hardy-Riesz-Brascamp-Lieb-Luttinger type,“ arXiv:1307.8439[math.CA] (2013)
[17] T. Richardson, Private communication, July 2013.
[18] G. Alberti, ”Some remarks about a notion of rearrangement“, ann. Scuola Norm. Sup. Pisa Cl. Sci, XXIX, pp. 457-472 (2000)