

# Analysis of Coupled Scalar Systems by Displacement Convexity

Rafah El-Khatib\*, Nicolas Macris\*, Tom Richardson† and Ruediger Urbanke\*

\*School of Computer and Communication Sciences

EPFL, Lausanne, Switzerland

Emails: {rafah.el-khatib,nicolas.macris,ruediger.urbanke}@epfl.ch

† Qualcomm, USA

Email: tomr@qti.qualcomm.com

**Abstract**—Potential functionals have been introduced recently as an important tool for the analysis of coupled scalar systems (e.g. density evolution equations). In this contribution we investigate interesting properties of this potential. Using the tool of displacement convexity we show that, under mild assumptions on the system, the potential functional is *displacement convex*. Furthermore, we give the conditions on the system such that the potential is *strictly displacement convex* in which case the minimizer is unique.

## I. INTRODUCTION

Spatially coupled systems have been used recently in various frameworks, such as compressive sensing, statistical physics, coding, and random constraint satisfaction problems (see [1] and the references therein for a review of the literature). They have been shown to exhibit excellent performance, often optimal, under low complexity message passing algorithms. For example, spatially coupled codes achieve capacity under such algorithms [1].

The performance of these systems is assessed by the solutions of (coupled) Density Evolution (DE) update equations. In general, these equations can be viewed as the stationary point equations of a functional that is typically called “the potential”. It has already been recognized that this variational formulation is a powerful tool to analyze DE updates under suitable initial conditions [2]. There are various possible formulations of this potential functional [1], [2], [3], [4]. In this paper we will use the representation of [1] for scalar systems.

In a previous contribution [5], [6] we showed that the potential (in the form [2]) associated to a spatially coupled Low-Density Parity-Check (LDPC) code whose single system is the  $(\ell, r)$ -regular Gallager ensemble, with transmission over the BEC( $\epsilon$ ), has a convex structure called “displacement convexity”. This structure is well known in the theory of optimal transport [7]. In fact the potential we consider in [5], [6] is *not* convex in the usual sense but it *is* in the sense of displacement convexity. This in itself is an interesting property. When displacement convexity is strict one deduces that the minimum of the potential is unique — assuming it exists — and thus so is the solution of the DE equation.

The main purpose of the present note is to prove that a general class of scalar systems also exhibits the property of displacement convexity, and even strict displacement convexity

under rather mild assumptions. For this purpose we will use the potential in the representation of [1] which allows to obtain more transparent, general, and simpler proofs.

This manuscript is organized as follows. Section II introduces the model and the variational formulation. In Section III we prove rearrangement inequalities that allow us to naturally consider the potential as a functional of cumulative distribution functions. The potential is shown to be displacement convex in Section IV. Strict displacement convexity and unicity of the minimizer are proved in Section V.

## II. SET UP AND VARIATIONAL FORMULATION

The natural setting for displacement convexity is the continuum case which can be thought of as an approximation of the discrete system in the regime of large spatial length and window size. The continuum limit has already been introduced in the literature as a convenient means to analyze the behavior of the original discrete model [1], [3], [8], [9].

Consider a spatially coupled system with an averaging window  $w : \mathbb{R} \rightarrow \mathbb{R}$  which is always assumed to be bounded, non negative, even, integrable and normalized such that  $\int_{\mathbb{R}} dx w(x) = 1$  (as we will see, sometimes further assumptions will be necessary depending on the statements). We denote by  $\otimes$  the standard convolution on  $\mathbb{R}$  and express the “fixed point DE equations of a scalar continuous system” as follows:

$$g(x) = h_g((f \otimes w)(x)), \quad (1)$$

$$f(x) = h_f((g \otimes w)(x)), \quad (2)$$

where  $x \in \mathbb{R}$  is the spatial position. We will often use the shorthand notation  $f^w = f \otimes w$  and  $g^w = g \otimes w$ ; further, we will often refer to the functions  $f, g$  as *profiles* and to  $h_f, h_g$  as *update functions*.

We will also adopt a convenient normalization for all these functions. As will become clear in the example below it is always possible to adopt this normalization in specific applications. First, we assume that the profiles are bounded. Specifically,  $f, g : \mathbb{R} \rightarrow [0, 1]$ . Next, we assume that the update functions  $h_f$  and  $h_g$  are non-decreasing bounded functions  $h_{f,g} : [0, 1] \rightarrow [0, 1]$  normalized such that  $h_f(0) = h_g(0) = 0$

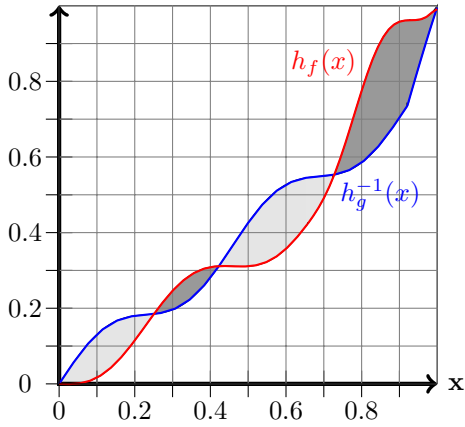


Fig. 1. An example of the systems we consider. The EXIT-like curves are  $h_f$  (in red) and  $h_g^{-1}$  (in blue). The signed area  $A(h_f, h_g; 1)$  from (3) is the sum of the light gray areas (positively signed) and the dark gray areas (negatively signed), and it is equal to 0.

and  $h_f(1) = h_g(1) = 1$ . We will think of them as EXIT-like curves  $(u, h_f(u))$  and  $(h_g(v), v)$  for  $u, v \in [0, 1]$  (see the generic plot).

Consider the signed area between the two curves, namely

$$A(h_f, h_g; u) = \int_0^u du' (h_g^{-1}(u') - h_f(u')). \quad (3)$$

This is a functional of  $h_f$ ,  $h_g$  and a function of  $u \in [0, 1]$ . We consider the case where  $A(h_f, h_g; u) > 0$  for all  $u \in ]0, 1[$  and  $A(h_f, h_g; 1) = 0$ . This is equivalent to the *strictly positive gap condition* of [1]. In [1] the condition was stated in terms of the function of  $u$  and  $v$

$$\phi(h_f, h_g; u, v) = \int_0^u du' h_g^{-1}(u') + \int_0^v dv' h_f^{-1}(v') - uv.$$

the condition being that  $\phi$  is positive for  $(u, v) \in [0, 1]^2$  except at  $(0, 0)$  and  $(1, 1)$  where it takes the value 0.  $A(h_f, h_g; u)$  is obtained by minimizing  $\phi(h_f, h_g; u, v)$  over  $v$ .

*Example* Take a spatially coupled LDPC code whose single system is the  $(\ell, r)$ -regular Gallager ensemble, with transmission over the BEC( $\epsilon$ ). Let  $u$  (resp.  $v$ ) be the erasure probabilities emitted by check (resp. variable) nodes. Let the functions  $\tilde{h}_f$  (resp.  $\tilde{h}_g$ ) give the usual erasure probabilities emitted on variable (resp. check) node sides. Explicitly,  $\tilde{h}_f(u) = \epsilon u^{\ell-1}$  and  $\tilde{h}_g(v) = 1 - (1-v)^{r-1}$ , and the usual DE equations are  $v = \tilde{h}_f(u)$ ,  $u = \tilde{h}_g(v)$ . We consider the special case  $\epsilon = \epsilon_{\text{MAP}}$ . Let  $(u_{\text{MAP}}, v_{\text{MAP}})$  be the unique non-trivial fixed point such that  $A(h_f, h_g; u_{\text{MAP}}) = 0$ . The *normalized* functions  $h_f$  and  $h_g$  are defined as  $h_f(u) = \tilde{h}_f(u_{\text{MAP}}u)/v_{\text{MAP}}$  and  $h_g(v) = \tilde{h}_g(v_{\text{MAP}}v)/u_{\text{MAP}}$ . The coupled DE fixed point equations are

$$u_{\text{MAP}}g(x) = 1 - (1 - f^w(x)v_{\text{MAP}})^{r-1}, \quad f(x) = (g^w(x))^{\ell-1}.$$

It is not difficult to check that the DE equations (1)-(2) are stationary point equations of a *potential functional* of both

profiles  $f$  and  $g$ ,

$$\mathcal{W}(f, g) = \int_{\mathbb{R}} dx \left\{ I_{f,g}(x) - f(x)g^w(x) \right\} \quad (4)$$

where

$$I_{f,g}(x) = \int_0^{g(x)} du h_g^{-1}(u) + \int_0^{f(x)} dv h_f^{-1}(v).$$

In order for the integral over  $\mathbb{R}$  in (4) to exist we have to impose some conditions on the profiles  $f$  and  $g$ . These will thus be taken in the spaces (here  $\epsilon > 0$ )

$$\mathcal{S}_f = \{f : \mathbb{R} \rightarrow [0, 1], \quad (5)$$

$$\lim_{x \rightarrow -\infty} x^{1+\epsilon} f(x) = 0, \quad \lim_{x \rightarrow +\infty} x^{1+\epsilon} (f(x) - 1) = 0\},$$

$$\mathcal{S}_g = \{g : \mathbb{R} \rightarrow [0, 1], \quad (6)$$

$$\lim_{x \rightarrow -\infty} x^{1+\epsilon} g(x) = 0, \quad \lim_{x \rightarrow +\infty} x^{1+\epsilon} (g(x) - 1) = 0\}.$$

The fixed point profile solutions of the DE equations (1), (2) can be seen as the left side of the (symmetric) decoding waves. In this paper we assume that these solutions belong to the spaces  $\mathcal{S}_f$  and  $\mathcal{S}_g$ . This is achieved under some mild conditions on the slopes of the update functions  $h_f$  and  $h_g$  at the corner points, and if  $w$  decays fast enough. In particular this is true for the example of the BEC( $\ell, r$ ) with a finitely supported  $w$ , where the limiting values are approached at least exponentially fast.

### III. REARRANGEMENTS

Displacement convexity is defined on a space of probability measures. For measures on the real line it is most convenient to view displacement convexity on a space of cumulative distribution functions (cdf's). In this section we use the tool of *increasing rearrangements* to show that such rearrangements of  $f$  and  $g$  can only decrease the potential.

We first give a brief introduction to the notion of non-decreasing rearrangement, see [10]. Consider a profile  $p : \mathbb{R} \rightarrow [0, 1]$  such that  $\lim_{x \rightarrow -\infty} p(x) = 0$  and  $\lim_{x \rightarrow +\infty} p(x) = 1$ . Then, the *increasing rearrangement*<sup>1</sup> of  $p$  is the non-decreasing function  $\bar{p}$  so that this function has the same limits and the mass of each level set is preserved. More formally, let us represent  $p$  in layer cake form as  $p(x) = \int_0^{p(x)} dt = \int_0^{+\infty} dt \chi_t(x)$ , where  $\chi_t(x)$  is the indicator function of the level set  $E_t = \{x | p(x) > t\}$ . For each value  $t \in [0, +\infty[$ , 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  $\bar{E}_t = ]a_t - |A_t|, +\infty[$  and  $\bar{\chi}_t$  the indicator function of  $\bar{E}_t$ . Then,  $\bar{p}(x) = \int_0^{+\infty} dt \bar{\chi}_t(x)$ .

*Proposition 3.1:* Assume that the window function  $w$  is symmetric decreasing<sup>2</sup>. Let  $f$  and  $g$  be in  $\mathcal{S}_f$  and  $\mathcal{S}_g$  respectively, and let  $\bar{f}$  and  $\bar{g}$  be their respective increasing rearrangements. Then,  $\mathcal{W}(f, g) \geq \mathcal{W}(\bar{f}, \bar{g})$ .

<sup>1</sup>Note that an increasing rearrangement is not necessarily strictly increasing.

<sup>2</sup>We say that a function is *symmetric decreasing* if it is even and non-increasing on the positive half line.

To prove Proposition 3.1, we make use of the Riesz rearrangement inequality in one dimension [11].

Let  $p$  be a non-negative measurable function on  $\mathbb{R}$ . Then the *symmetric decreasing rearrangement*  $p^*$  of  $p$  is defined as an even function, that is decreasing on  $[0, +\infty[$  and such that the level sets  $\{x|p(x) \geq t\}$  and  $\{x|p^*(x) \geq t\}$  have equal Lebesgue measure. Note that the decrease on  $[0, +\infty[$  is not necessarily strict.

*Lemma 3.2 (Riesz's Inequality):* Let  $f_1, f_2,$  and  $f_3$  be any measurable non-negative functions on the real line, and  $f_1^*, f_2^*$ , and  $f_3^*$  be their *symmetric decreasing rearrangements*. Then<sup>3</sup>,

$$\begin{aligned} & \iint_{\mathbb{R}^2} dx dy f_1(x) f_2(x-y) f_3(y) \\ & \leq \iint_{\mathbb{R}^2} dx dy f_1^*(x) f_2^*(x-y) f_3^*(y). \end{aligned}$$

*Proof of Proposition 3.1:* Consider the expression of the potential in (4). In order to make use of the Riesz inequality we first ‘‘symmetrize’’ the profiles  $f$  and  $g$ , and rewrite the functional in terms of symmetric profiles. Choose  $R > 0$  very large but fixed. Eventually we will take  $R \rightarrow +\infty$ . Denote by  $\hat{f}$  the profile such that  $\hat{f}(x) = f(x)$ ,  $x < R$  and  $\hat{f}(x) = f(2R-x)$ ,  $x > R$ , and by  $\hat{g}$  the function such that  $\hat{g}(x) = g(x)$ ,  $x < R$  and  $\hat{g}(x) = \hat{g}(2R-x)$ ,  $x > R$ .

We now write the potential in (4) in terms of the symmetrized profiles  $\hat{f}$  and  $\hat{g}$ . We have

$$\begin{aligned} \mathcal{W}(f, g) &= \lim_{R \rightarrow +\infty} \int_{-\infty}^R dx \left\{ I_{f,g}(x) - f(x)g^w(x) \right\} \\ &= \lim_{R \rightarrow +\infty} \left\{ \int_{-\infty}^R dx I_{f,g}(x) - \int_{-\infty}^R dx f(x)g^w(x) \right\}. \end{aligned} \quad (7)$$

For the first term in the brackets above it is straightforward to see that

$$\int_{-\infty}^R dx I_{f,g}(x) = \frac{1}{2} \int_{\mathbb{R}} dx I_{\hat{f}, \hat{g}}(x). \quad (8)$$

For the second term slight care must be taken because of the convolution. We find

$$\begin{aligned} & \int_{-\infty}^R dx \int_{\mathbb{R}} dy f(x)w(x-y)g(y) \\ &= \frac{1}{2} \int_{\mathbb{R}} dx \int_{\mathbb{R}} dy \hat{f}(x)w(x-y)\hat{g}(y) + o\left(\frac{1}{R^2}\right). \end{aligned} \quad (9)$$

Note that  $\hat{f}$  and  $\hat{g}$  are integrable over  $\mathbb{R}$ . Then we can write

$$\mathcal{W}(f, g) = \frac{1}{2} \lim_{R \rightarrow +\infty} (\mathcal{U}_1(\hat{g}) + \mathcal{U}_2(\hat{f}) + \mathcal{U}_3(\hat{f}, \hat{g}))$$

where

$$\begin{aligned} \mathcal{U}_1(\hat{g}) &= \int_{\mathbb{R}} dx \int_0^{\hat{g}(x)} du h_g^{-1}(u), \\ \mathcal{U}_2(\hat{f}) &= \int_{\mathbb{R}} dx \int_0^{\hat{f}(x)} dv h_f^{-1}(v), \\ \mathcal{U}_3(\hat{f}, \hat{g}) &= - \int_{\mathbb{R}} dx \int_{\mathbb{R}} dy \hat{f}(x)w(x-y)\hat{g}(y). \end{aligned} \quad (10)$$

<sup>3</sup>If the left hand side is infinite so is the right hand side and the inequality is satisfied.

Now consider  $\hat{f}^*$  and  $\hat{g}^*$  the symmetric decreasing rearrangements of  $\hat{f}$ , and  $\hat{g}$ , respectively. Each of the above terms may only decrease upon rearrangements. Indeed each of the first two terms is a functional of a single monotone function and thus remains unchanged by rearrangement:  $\mathcal{U}_1(\hat{g}) = \mathcal{U}_1(\hat{g}^*)$  and  $\mathcal{U}_2(\hat{f}) = \mathcal{U}_2(\hat{f}^*)$  (see for example [11] p. 80 (3.3)). The term (10) decreases upon rearrangement as a direct application<sup>4</sup> of Lemma 3.2. We thus conclude that

$$\mathcal{W}(\hat{f}, \hat{g}) \geq \frac{1}{2} \lim_{R \rightarrow +\infty} \mathcal{W}(\hat{f}^*, \hat{g}^*).$$

To obtain  $\mathcal{W}(f, g) \geq \mathcal{W}(\bar{f}, \bar{g})$  it remains to remark that  $\frac{1}{2} \lim_{R \rightarrow +\infty} \mathcal{W}(\hat{f}^*, \hat{g}^*) = \mathcal{W}(\bar{f}, \bar{g})$ . This is achieved by reversing the steps (7)-(9). ■

From now on we therefore restrict the functional to the spaces of non-increasing profiles.

#### IV. DISPLACEMENT CONVEXITY

A generic functional  $\mathcal{F}(p)$  on a space  $\mathcal{X}$  (of ‘‘profiles’’ say) is said to be convex in the usual sense if for any pair  $p_0, p_1 \in \mathcal{X}$ , and for all  $\lambda \in [0, 1]$ , the inequality  $\mathcal{F}((1-\lambda)p_0 + \lambda p_1) \leq (1-\lambda)\mathcal{F}(p_0) + \lambda\mathcal{F}(p_1)$  holds. Displacement convexity, on the other hand, is defined as convexity under an alternative interpolation called displacement interpolation. The right setting for displacement convexity is a space of probability measures. For *measures over the real line* one can conveniently define the displacement interpolation in terms of the cdf's associated to the measures. This is the simplest setting and the one that we adopt here.

We think of the increasing profiles  $f$  and  $g$  as cdf's of some underlying measures over the real line. Consider two pairs  $(f_0, g_0)$  and  $(f_1, g_1)$ , and define two (pushforward) maps  $T_f$  and  $T_g$  as

$$T_f(x) = f_1^{-1}(f_0(x)), \quad T_g(x) = g_1^{-1}(g_0(x)).$$

Consider the linear interpolation between points on  $\mathbb{R}$ ,

$$x_{f,\lambda} = (1-\lambda)x + \lambda T_f(x), \quad x_{g,\lambda} = (1-\lambda)x + \lambda T_g(x).$$

The displacement interpolants  $(f_\lambda, g_\lambda)$  are defined so that the following equalities hold for all  $\lambda \in [0, 1]$  and  $x \in \mathbb{R}$

$$f_\lambda(x_{f,\lambda}) = f_0(x), \quad g_\lambda(x_{g,\lambda}) = g_0(x).$$

We now state the main result of this section.

*Proposition 4.1:* The potential  $\mathcal{W}(f, g)$  is displacement convex which means that for all  $\lambda \in [0, 1]$

$$\mathcal{W}(f_\lambda, g_\lambda) \leq (1-\lambda)\mathcal{W}(f_0, g_0) + \lambda\mathcal{W}(f_1, g_1).$$

We define the two following quantities

$$\Omega(x) = \int_{-\infty}^x dz w(z), \quad V(x) = \int_{-\infty}^x dz \Omega(z), \quad (11)$$

and call  $V$  the *kernel* for reasons which will appear shortly.

<sup>4</sup>We apply this inequality for  $f_1 = f$ ,  $f_2 = w$ ,  $f_3 = g$ . As assumed in Proposition 3.1  $w$  is symmetric decreasing window so that for us  $w(x) = w^*(x)$  for all  $x \in \mathbb{R}$ .

Before proving the proposition let us first note

$$\begin{aligned} \int_{\mathbb{R}} dx (f(x) - f(+\infty))g^w(x) &= \int_{\mathbb{R}} dx (f^w(x) - f(+\infty))g(x) \\ &= \int_{\mathbb{R}} dx \int_{\mathbb{R}} dy (f(y) - f(+\infty))w(y-x)g(x) \\ &= - \int_{\mathbb{R}} \int_{\mathbb{R}} df(y) V(y-x) dg(x), \end{aligned}$$

where we have used integration by parts and  $g(-\infty) = f(-\infty) = 0$  for the last step.

*Proof of Proposition 4.1:* Using the last identity we rewrite the potential in (4) as follows.

$$\begin{aligned} \mathcal{W}(f, g) &= \int_{\mathbb{R}} dx \left\{ \int_0^{g(x)} du h_g^{-1}(u) + \int_0^{f(x)} dv h_f^{-1}(v) \right. \\ &\quad \left. - f(+\infty)g(x) \right\} + \iint_{\mathbb{R}^2} df(y) V(y-x) dg(x). \end{aligned} \quad (12)$$

We now express the potential as the sum  $\mathcal{W}(f, g) = \mathcal{W}_1(f, g) + \mathcal{W}_2(f, g)$ , where  $\mathcal{W}_2(f, g)$  consists of the last double integral in (12).

We first consider  $\mathcal{W}_1(f, g)$  and write

$$\begin{aligned} \mathcal{W}_1(f_\lambda, g_\lambda) - \mathcal{W}_1(f_0, g_0) &= \int_{\mathbb{R}} dx \left\{ \int_0^{g_\lambda(x)} du h_g^{-1}(u) \right. \\ &\quad - \int_0^{g_0(x)} du h_g^{-1}(u) + \int_0^{f_\lambda(x)} dv h_f^{-1}(v) - \int_0^{f_0(x)} dv h_f^{-1}(v) \\ &\quad \left. - (f_\lambda(+\infty)g_\lambda(x) + f_0(+\infty)g_0(x)) \right\}. \end{aligned} \quad (13)$$

We remark that

$$\begin{aligned} \int_{\mathbb{R}} dx \left( \int_0^{g_\lambda(x)} du h_g^{-1}(u) - \int_0^{g_0(x)} du h_g^{-1}(u) \right) \\ = \int_{\mathbb{R}} dx \int_0^1 du \left( \Theta(g_\lambda(x) - u) - \Theta(g_0(x) - u) \right) h_g^{-1}(u), \end{aligned} \quad (14)$$

where  $\Theta$  is the Heaviside step function. One can check by considering the two cases  $g_\lambda(x) > g_0(x)$  and  $g_0(x) > g_\lambda(x)$  that

$$\int_0^1 du \left| \Theta(g_\lambda(x) - u) - \Theta(g_0(x) - u) \right| = |g_\lambda(x) - g_0(x)|.$$

This observation allows us to use Fubini's theorem to swap the integrals in (14). We then write (14) as

$$\begin{aligned} \int_0^1 du h_g^{-1}(u) \int_{\mathbb{R}} dx \left( \Theta(g_\lambda(x) - u) - \Theta(g_0(x) - u) \right) \\ = \int_0^1 du h_g^{-1}(u) \int_{\mathbb{R}} dx \left( \Theta(x - g_\lambda^{-1}(u)) - \Theta(x - g_0^{-1}(u)) \right) \\ = \int_0^1 du h_g^{-1}(u) (g_0^{-1}(u) - g_\lambda^{-1}(u)) \\ = \lambda \int_{\mathbb{R}} dg_0(x) (x - T_g(x)) h_g^{-1}(g_0(x)) \end{aligned}$$

In the last line we used a change of variables  $u = g_0(x)$ . Using a similar analysis for the other terms in (13) we find

$$\begin{aligned} \mathcal{W}_1(f_\lambda, g_\lambda) - \mathcal{W}_1(f_0, g_0) \\ = \lambda \left( \int_{\mathbb{R}} dg_0(x) (x - T_g(x)) h_g^{-1}(g_0(x)) \right. \\ \left. + df_0(x) (x - T_f(x)) h_f^{-1}(f_0(x)) + dg_0(x) (x - T_g(x)) \right). \end{aligned}$$

We thus conclude that  $\mathcal{W}_1(f_\lambda, g_\lambda)$  is linear, and hence convex, in  $\lambda$ .

We now consider the double integral term  $\mathcal{W}_2(f, g)$  in (12). Using again a change of variables write

$$\begin{aligned} \mathcal{W}_2(f_\lambda, g_\lambda) &= \iint_{\mathbb{R}^2} df_\lambda(y) V(y-x) dg_\lambda(x) \\ &= \iint_{\mathbb{R}^2} df_0(y) V((1-\lambda)(y-x) \\ &\quad + \lambda(T_f(y) - T_g(x))) dg_0(x). \end{aligned}$$

This is convex in  $\lambda$  because the kernel  $V$  is (see (11)).  $\blacksquare$

## V. UNICITY OF MINIMIZER

In this section we prove that the potential is strictly displacement convex under the strictly positive gap condition. This implies that it admits a unique minimizer.

Under this condition and assuming that  $w$  is even and regular<sup>5</sup>, the existence of increasing fixed point solutions was established in [1]. It was also shown in [1] that existence of such a fixed point implies a positive gap condition<sup>6</sup>. It was shown that if  $A(h_f, h_g; u) = 0$  for some  $u \in ]0, 1[$  then there may be an infinite family of fixed point solutions not equivalent under translation. The proof of unicity relies on the potential function formulation so the regularity conditions (5) and (6) are required. They can be shown to be necessary under mild assumptions on the scalar recursion stability of the fixed points (0, 0) and (1, 1) and on the decay of  $w$ .

From the results in the preceding section it follows that all increasing fixed points must have the same potential and this potential is minimal.

Let  $f_0, g_0$  and  $f_1, g_1$  both be non-decreasing fixed points. We claim that they must be translates of each other, i.e.,  $y - T_f(y)$  is constant  $df_0$ -almost everywhere (a.e.) and  $x - T_g(x)$  takes the same constant value  $dg_0$ -a.e. Note that one of these conditions implies the other since both pairs are fixed points. We will show that  $y - T_f(y)$  is constant  $df_0$ -a.e. The method of proof is to show that if this is not the case then  $\mathcal{W}_2(f_\lambda, g_\lambda)$  is strictly convex at  $\lambda = 0$  which contradicts the minimality of  $f_1, g_1$ .

The proof relies on results from [1] that relate the strictly positive gap condition to the positivity of certain integrals of spatial fixed points, which will be shown to imply the strict convexity of  $\mathcal{W}_2(f_\lambda, g_\lambda)$  if  $y - T_f(y)$  is not constant  $df_0$ -a.e.

<sup>5</sup>Regularity of  $w$  means that it is strictly positive on an interval  $(-W, W)$ ,  $W \leq +\infty$  and 0 off of  $[-W, W]$ .

<sup>6</sup>The strictly positive gap condition requires that  $A(h_f, h_g; u) > 0$  whereas the positive gap condition requires only that  $A(h_f, h_g; u) \geq 0$ .

Since we have no further need for explicit use of  $f_1, g_1$  we will simplify notation and refer to  $f_0, g_0$  as  $f, g$ .

Let us introduce the following functional from [1],

$$\begin{aligned} \xi_\phi(w; f, g; x_1, x_2) = & \quad (15) \\ & \int_0^{+\infty} dx w(x) \int_{]x_2, x_1+x]} df(y) (g(x_1+) - g(y-x)) \\ & + \int_0^{+\infty} dx w(x) \int_{]x_1, x_2+x]} dg(y) (f(x_2+) - f(y-x)) \end{aligned}$$

where the integrals are Lebesgue-Stieltjes integrals. If  $x' < x$  then  $\int_{]x, x']} dg(y) f(y)$  is defined to be  $-\int_{]x', x]} dg(y) f(y)$ . Note that  $\xi_\phi$  is non-negative; this is closely related to the strictly positive gap condition.

One of the main results in [1] is that for a non-decreasing fixed-point solution we have

$$\xi_\phi(w; f, g; x_1, x_2) = \phi(h_f, h_g; g(x_1+), f(x_2+)). \quad (16)$$

Note that if  $x_1 = x_2 = x$  then the right hand side is (the right continuous version of)  $I_{f,g}(x) - f(x)g(x)$ . By the strictly positive gap condition we now have  $\xi_\phi(w; f, g; x_1, x_2) > 0$  for all  $x_1, x_2$  where  $f(x_1), g(x_2) \in ]0, 1[$ . Let the support of  $w$  be  $[-W, W]$  (we may have  $W = +\infty$ ). We define  $A_x = (x - W, x + W)$ . From (15) it is easy to see that  $dg(A_x) = 0$  implies  $\xi_\phi(w; f, g; x, x) = 0$  so  $dg(A_x) > 0$  for all  $x$  with  $g(x) \in ]0, 1[$  and, similarly  $df(A_x) > 0$  for all  $x$  with  $f(x) \in ]0, 1[$ . These conditions imply  $\dot{f}^w(x) > 0$  and  $\dot{g}^w(x) > 0$  for  $f(x) \in ]0, 1[$  and  $g(x) \in ]0, 1[$  respectively, where the dot denotes differentiation.

We express  $\xi_\phi$  in a more useful form for our current purpose. We assume that  $x_1 = x_2$  since we need the result only for this case. We claim that (15) is equal to

$$\iint df(x) dg(y) \mathbb{1}_{\{(x-x_1)(y-x_1) \leq 0\}} \Omega(-|x-y|). \quad (17)$$

To see this note that

$$\begin{aligned} & \int_0^{+\infty} dx w(x) \int_{]x_1, x_1+x]} df(y) (g(x_1+) - g(y-x)) \\ & = \int_0^{+\infty} dx w(x) \int_{]x_1, x_1+x]} df(y) \int_{]y-x, x_1]} dg(z) \\ & = \int_0^{+\infty} dx w(x) \iint dg(z) df(y) \mathbb{1}_{\{0 \leq y-z \leq x\}} \mathbb{1}_{\{(z-x_1)(y-x_1) \leq 0\}} \\ & = \iint dg(z) df(y) \mathbb{1}_{\{0 \leq y-z\}} \Omega(-(y-z)) \mathbb{1}_{\{(z-x_1)(y-x_1) \leq 0\}} \end{aligned}$$

where for the last step we integrate by parts, using  $\frac{d}{dx}(-\Omega(-x)) = w(x)$ . The other term can be handled similarly and we obtain (17).

Let  $A_g$  denote the support of  $dg$  and let  $A_f$  denote the support of  $df$ . Given a real valued  $s$  define measures  $df^+$  as  $df^+(A) = df(A \cap \{x : x - T_f(x) \geq s\})$  and  $df^-$  as  $df^-(A) = df(A \cap \{x : x - T_f(x) < s\})$ . Let  $A_f^+$  denote the support of  $df^+$  and  $A_f^-$  denote the support of  $df^-$ . Clearly  $A_f = A_f^+ \cup A_f^-$ .

Since  $V(x)$  is strictly convex on  $] -W, W[$ , strict convexity of  $\mathcal{W}_2(f_\lambda, g_\lambda)$  at  $\lambda = 0$  follows if there exists  $x_g \in A_g$ ,

$x_f^+ \in A_f^+$ , and  $x_f^- \in A_f^-$ , such that  $|x_g - x_f^+| < W$  and  $|x_g - x_f^-| < W$ .

Assume now the existence of an  $s$  so that  $A_f^+$  and  $A_f^-$  are both non-empty. We will show that this implies strict convexity by establishing the existence of an  $x_g$  as above. Since this is a contradiction we conclude that  $x - T_f(x)$  is constant  $df$ -a.e., giving the desired result.

If there exists  $z \in A_f^- \cap A_f^+$  then we take  $x_f^+ = x_f^- = z$ . Since the strictly positive gap condition implies  $dg(z-W, z+W) > 0$  (by (16)) we have  $A_g \cap (z-W, z+W) \neq \emptyset$  and we can find a suitable  $x_g$ . Assume now that  $A_f^+ \cap A_f^- = \emptyset$ . Let  $z^- \in A_f^-$  and  $z^+ \in A_f^+$ . We shall assume that  $z^- < z^+$ , the argument being the same if the order is reversed. Define  $x_f^- = \max\{A_f^- \cap [z^-, z^+]\}$  and  $x_f^+ = \min\{A_f^+ \cap [x_f^-, z^+]\}$ . It follows that  $]x_f^-, x_f^+[\cap A_f = \emptyset$ . Since  $A_f = A_f^+ \cup A_f^-$  it follows from the strictly positive gap condition (16) that  $x_f^+ - x_f^- < 2W$ . Define  $z = (x_f^+ + x_f^-)/2$ .

Setting  $x_1 = x_2 = z$ , it follows from the form (17), the strictly positive gap condition, and (16) that the  $df(x)dg(y)$ -measure of at least one of the following

$$T_1 = \{(x, y) : x \geq z, y \leq z, x - y < W\}$$

$$T_2 = \{(x, y) : x \leq z, y \geq z, y - x < W\}$$

is strictly positive. Let us assume  $df dg(T_1) > 0$ . Note that  $x_f^+ = \min A_f^+ \cap [z, z+W]$  so it follows that there exists  $x_g \in [z-W, z]$  such that  $0 \leq x_f^+ - x_g < W$ . This clearly gives  $x_f^- - x_g < W$ . In addition we have  $x_f^- - x_g > x_f^- - z > -W$  so we obtain  $|x_f^- - x_g| < W$ . The argument assuming  $df dg(T_2) > 0$  is similar.

**Acknowledgments.** R. E. thanks Vahid Aref for many interesting discussions.

## REFERENCES

- [1] S. Kudekar, T. Richardson, and R. L. Urbanke, "Wave-Like Solutions of General One-Dimensional Spatially Coupled Systems," <http://arxiv.org/abs/1208.5273>, 2012.
- [2] A. Yedla, Y.-Y. Jian, P. S. Nguyen, and H. D. Pfister, "A Simple Proof of Threshold Saturation for Coupled Scalar Recursions," *7th Symp. on Turbo Codes and Iterative Information Processing*, pp. 51–55, 2012.
- [3] S. H. Hassani, N. Macris, and R. Urbanke, "Coupled Graphical Models and Their Thresholds," in *Proc. Inf. Theory Workshop IEEE (Dublin)*, 2010, pp. 1–5.
- [4] S. Kumar, A. J. Young, N. Macris, and H. D. Pfister, "A Proof of Threshold Saturation for Spatially-Coupled LDPC Codes on BMS Channels," <http://arxiv.org/abs/1301.6111>, 2013.
- [5] R. El-Khatib, N. Macris, and R. Urbanke, "Displacement Convexity – A Useful Framework for the Study of Spatially Coupled Codes," *Proc. Inf. Theory Workshop IEEE (Sevilla)*, pp. 1–5, 2013.
- [6] —, "Displacement Convexity – A Useful Framework for the Study of Spatially Coupled Codes," <http://arxiv.org/abs/1304.6026>, 2013.
- [7] C. Villani, *Topics in Optimal Transportation*. AMS, 2003, vol. 58.
- [8] S. H. Hassani, N. Macris, and R. L. Urbanke, "Chains of mean field models," *J. Stat. Mech: Theory and Exp*, vol. P02011, 2012.
- [9] D. L. Donoho, A. Javanmard, and A. Montanari, "Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*. IEEE, 2012, pp. 1231–1235.
- [10] G. E. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*. London and New York: Cambridge University Press, 1952.
- [11] E. Lieb and M. Loss, *Analysis*. Providence, Rhode Island: American Mathematical Society, 1997.