Electronic Research Archive

http://www.aimspress.com/journal/era

ERA, 30(5): 1632–1652.
DOI: 10.3934/era.2022082
Received: 22 August 2021
Revised: 12 February 2022
Accepted: 17 February 2022
Published: 24 March 2022

Research article

Some computable quasiconvex multiwell models in linear subspaces without rank-one matrices

Ke Yin[1,*] **and Kewei Zhang**[2,*]

[1] Center for Mathematical Sciences, Huazhong University of Science and Technology, Wuhan 430074, China

[2] School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK

* **Correspondence:** Email: kyin@hust.edu.cn, Kewei.Zhang@nottingham.ac.uk.

Abstract: In this paper we apply a smoothing technique for the maximum function, based on the compensated convex transforms, originally proposed by Zhang in [1] to construct some computable multiwell non-negative quasiconvex functions in the calculus of variations. Let $K \subseteq E \subseteq M^{m\times n}$ with K a finite set in a linear subspace E without rank-one matrices of the space $M^{m\times n}$ of real $m \times n$ matrices. Our main aim is to construct computable quasiconvex lower bounds for the following two multiwell models with possibly uneven wells:
i) Let $f : K \subseteq E \to E^{\perp}$ be an L-Lipschitz mapping with $0 \le L \le 1/\alpha$ and
$H_2(X) = \min\{|P_E X - A_i|^2 + \alpha|P_{E^\perp}X - f(A_i)|^2 + \beta_i :\ i = 1, 2, \dots, k\}$, where $\alpha > 0$ is a control parameter, and
ii) $H_1(X) = \alpha|P_{E^\perp}X|^2 + \min\{ \sqrt{|\mathcal{U}_i(P_E X - A_i)|^2 + \gamma_i} : i = 1, 2, \dots, k\}$, where $A_i \in E$ with $U_i : E \to E$ invertible linear transforms for $i = 1, 2, \dots, k$. If the control paramenter $\alpha > 0$ is sufficiently large, our quasiconvex lower bounds are 'tight' in the sense that near each 'well' the lower bound agrees with the original function, and our lower bound are of $C^{1,1}$. We also consider generalisations of our constructions to other simple geometrical multiwell models and discuss the implications of our constructions to the corresponding variational problems.

Keywords: multiwell models; vectorial calculus of variations; quasiconvex functions; quasiconvex envelope; quasiconvex lower bounds; computational lower boundes; translation method; maximum function; compensated convex transforms; $C^{1,1}$-smooth approximation

On the occasion of the 75th birthday of Professor E. N. Dancer

1. Introduction

In this paper we use the formula for the smooth approximations of the maximum function in $\mathbb{R}^n$ based on compensated convex transforms [1] (see Definition 2 below) to construct computable quasiconvex lower bounds with multiwell structure in the calculus of variations. By 'computable' we mean that the evaluation of the quasiconvex lower bounds only involves projection onto a linear subspace E without rank-one matrices in $M^{m\times n}$, and a projection onto the simplex Δ^k where $k \geq 2$ is the number of 'wells'. The complexity of the former is $O(mn)$, while the complexity of the latter is theoretically $O(k\log k)$ for the algorithm proposed in [2,3] or observed $O(k)$ in the algorithm proposed in [4].

The maximum function

$$f_m(x) = \max\{x_1, \dots, x_m\}, \quad x = (x_1, x_2, \dots, x_m) \in \mathbb{R}^m \tag{1.1}$$

plays an important role in many optimization problems, in the sense that many problems can be proposed as minimization of the maximum function composed with other functions [5]. In [1] a smooth approximation for the maximum function was introduced by the upper compensated convex transform $C^u_\lambda(f_m)$ for $\lambda > 0$, where the formula for the transform is given by

$$C^u_\lambda(f_m)(x) = \lambda|x|^2 - \lambda\mathrm{dist}^2\left(x,\ \mathrm{co}\left(\frac{\Delta_m}{2\lambda}\right)\right) + \frac{1}{4\lambda}, \quad x \in \mathbb{R}^m, \tag{1.2}$$

where $\mathrm{co}(\Delta_m/(2\lambda))$ is the convex hull of the set of the scaled basis vectors $\Delta_m/(2\lambda) = \{e_1/(2\lambda), e_2/(2\lambda), \dots, e_m/(2\lambda)\}$ with $e_j = (\delta_{j,1}, \delta_{j,1}, \dots, \delta_{j,m}) \in \mathbb{R}^m$ where $\delta_{j,k} = 1$ if $j = k$ and $\delta_{j,k} = 0$ if $j \neq k$ for $j = 1, 2, \dots, m$ are Kronecker delta. Let $P_{\mathrm{co}(\Delta_m/(2\lambda))}x$ be the projection of $x \in \mathbb{R}^m$ to the simplex $\mathrm{co}(\Delta_m/(2\lambda))$, we can write

$$C^u_\lambda(f_m)(x) = \lambda|x|^2 - \lambda\left|x - P_{\mathrm{co}\left(\frac{\Delta_m}{2\lambda}\right)}x\right|^2 + \frac{1}{4\lambda} \tag{1.3}$$

and the complexity for evaluatiing $C^u_\lambda(f_m)(x)$ is the same as finding the convex projection $P_{\mathrm{co}\left(\frac{\Delta_m}{2\lambda}\right)}x$. There are numerical schemes with complexity $O(m\log m)$ to compute the convex projection above [2]. Our computable quasiconvex functions rely on this fact. Some generalisations of approximations by compensated convex transforms to maximum-like functions can be found in [6].

It is known [1] that $C^u_\lambda(f_m)(x)$ is a $C^{1,1}$ approximation of $f_m(x)$ from above, that is, $f_m(x) \leq C^u_\lambda(f_m)(x)$ and $\lim_{\lambda\to+\infty} C^u_\lambda(f_m)(x) = f_m(x)$ [1], and in general compensated convex trnasforms are 'tight' approximations. For the special case of the maximum function, we will show in Proposition 2 below that $C^u_\lambda(f_m)(x)$ is a tight approximation in the sense that if the vector $x = (x_1, x_2, \dots, x_m)$ is sorted with the order $x_1 \geq x_2 \geq \cdots \geq x_m$, then $C^u_\lambda(f_m)(x) = f_m(x)$ if $x_1 - x_2 \geq 1/(2\lambda)$.

For the minimum function $g_m(x) = \min\{x_1, \dots, x_m\} = -\max\{-x_1, \dots, -x_m\}$ for $x \in \mathbb{R}^m$, it is easy to verify that the lower compensated convex transform of g_m is given by

$$C^l_\lambda(g_m)(x) = -C^u(f_m)(-x) = \lambda\mathrm{dist}^2\left(-x,\ \mathrm{co}\left(\frac{\Delta_m}{2\lambda}\right)\right) - \lambda|x|^2 - \frac{1}{4\lambda}. \tag{1.4}$$

In this paper we apply the lower compensated convex transforms to two multiwell models in the calculus of variations. Compensated convex transforms have been applied to problems involving geometric singular extraction, shape analysis and approximation and interpolations of sampled data [7–11].

The first type of multiwell functions is defined as follows. Let $\alpha > 0$ be a control parameter. Let $E \subseteq M^{m\times n}$ be a linear subspace without rank-one matrices and let $f : K \subseteq E \to E^{\perp}$ be an L-Lipschitz mapping with $0 \leq L \leq 1/\alpha$, where $E^{\perp}$ is the orthogonal complement of E in $M^{m\times n}$ and $K = \{A_i : i = 1, 2, \dots, k\} \subseteq E$ is a finite set. We define the multiwell function $H_2(X)$ with quadratic growth by

$$H_2(X) = \min\{|P_E X - A_i|^2 + \alpha|P_{E^\perp} X - f(A_i)|^2 + \beta_i : i = 1, 2, \dots, k\}, \quad X \in M^{m\times n} \tag{1.5}$$

for $k \geq 2$. If we define the quadratic function $q_i(X) = |P_E X - A_i|^2 + \alpha|P_{E^\perp} X - f(A_i)|^2 + \beta_i$ for $i = 1, 2, \dots, k$, then we may write

$$H_2(X) = \min\{q_i(X) : i = 1, 2, \dots, k\}.$$

We see that the point $X_i = A_i + f(A_i)$ is a local minimum point of $H_2(\cdot)$ if for $j \neq i$,

$$q_i(A_i + f(A_i)) = \beta_i < q_j(A_i + f(A_i) = |A_i - A_j|^2 + \alpha|f(A_i) - f(A_j)|^2 + \beta_j.$$

A sufficient condition for every $X_i = A_i + f(A_i)$ to be a local minimum point is $|A_i - A_j|^2 > |\beta_i - \beta_j|$ for $1 \leq i \neq j \leq k$. So, under this additional restriction, $H_2(X)$ has a multiwell structure with each $X_i = A_i + f(A_i)$ a local minimum point.

The second type of multiwell functions is in the form

$$H_1(X) = \alpha|P_{E^\perp} X|^2 + \min\{ \sqrt{|\mathcal{U}_i(P_E X - A_i)|^2 + \gamma_i} : i = 1, 2, \dots, k\}, \tag{1.6}$$

where $A_i \in E$ and $\mathcal{U}_i : E \to E$ are invertible linear transforms and $\gamma_i \geq 0$ for $i = 1, 2, \dots, k$. Again we can show that every $A_i \in E$ is a local minimum point of H_1 if $|\mathcal{U}_j(A_i - A_j)|^2 > |\gamma_i - \gamma_j|$.

Our aim is to use the lower compensated convex transform $C^u_\lambda(f_k)$ to construct composite functions that are quasiconvex lower bounds of H_1 and H_2. We will also consider simple generalisations of these constructions.

Next we discuss backgrounds and motivations. The Quasiconvex function in the sense of Morrey [12] is an important tool in the vectorial calculus of variations [13], with applications in nonlinear elasticity [14] and material microstructure. For example, under the growth condition $0 \leq F(X) \leq C|X|^p + C_1$ for $X \in M^{m\times n}$, the variational integral $I(u) := \int_\Omega F(Du)\,dx$ is weakly lower-semicontinuous in the Sobolev space $W^{1,p}(\Omega, \mathbb{R}^m)$ if and only if F is quasiconvex [15], where $M^{m\times n}$ is the linear space of real $m \times n$ matrices.

A continuous function $F : M^{m\times n} \to \mathbb{R}$ is quasiconvex [12] if for every $X \in M^{m\times n}$, for every open set $\Omega \subseteq \mathbb{R}^n$ and for every $\phi \in C_0^\infty(\Omega, R^m)$, we have

$$\int_\Omega F(X + D\phi(x))\,dx \geq \int_\Omega F(X). \tag{1.7}$$

Convex functions are quasiconvex but the converse is not true. A large class of functions, called polyconvex functions [14] are quasiconvex. Quasiconvex functions satisfying certain geometric conditions have been constructed by studying the quasiconvex envelope [13] of a given function with the multiwell property which describes the macroscopic properties of material microstructure for non-quasiconvex multiwell models [16, 17]. A simplified multiwell model is in the form of an (Euclidean) p-distance function to a compact set K in $M^{m\times n}$ [18–20]. A typical example is the p-distance function $\mathrm{dist}^p(X, \{-I, I\})$ to the set $\{-I, I\}$ for $X \in M^{2\times 2}, 1 \le p < 2$, whose quasiconvex envelope $Q\mathrm{dist}^p(X, \{-I, I\})$ vanishes exactly at $-I$ and $I \in M^{2\times 2}$. So far a more explicit geometric lower bound of $Q\mathrm{dist}^p(X, \{-I, I\})$ above is still not known. Some more explicit quasiconvex functions can be obtained by establishing lower bounds for the quasiconvex envelope via the so-called 'translation method' (see e.g., [21]). In particular, the explicit formula of the quasiconvex envelope for a double-well model [20] and the systematic study of restrictions of microstructure [22] lead to the study of linear subspaces of matrices without rank-one matrices, which are simple linear 'elliptic' objects in $M^{m\times n}$. Here a linear subspace E is called simple linear elliptic if $\mathrm{div}P_{E^\perp}(\mathrm{grad}\cdot)$ is a linear elliptic operator satisfying the strong Legendre-Hadamard ellipticity condition for some constant coefficient.

The following is a motivating example for our construction of computable quasiconvex lower bounds for a class of multiwell functions [1].

Let $E \subseteq M^{m\times n}$ be a linear subspace without rank-one matrices, where $m, n \ge 2$. It is known [22] that there exists $\varepsilon > 0$ such that

$$|P_{E^\perp} a \otimes b|^2 \ge \varepsilon |a|^2 |b|^2, \qquad a \in \mathbb{R}^m, \quad b \in \mathbb{R}^n,$$

where $P_{E^\perp} : M^{m\times n} \to E^\perp$ is the Euclidean orthogonal projection of $M^{m\times n}$ to the orthogonal complement $E^\perp$ of E, $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$ are treated as column vectors and $a\otimes b = ab^T \in M^{m\times n}$ with b^T the transpose of b.

Now we optimise the above construction [1] by defining

$$\lambda_0 = \inf\{|P_{E^\perp} a \otimes b|^2 : \ a \in \mathbb{R}^m, \ b \in \mathbb{R}^n, \ |a| = |b| = 1\},$$

where $|a|$ and $|b|$ are the Euclidean norms of a and b in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively, so that we have $0 < \lambda_0 < 1$ and

$$|P_{E^\perp} a \otimes b|^2 \ge \lambda_0 |a \otimes b|^2 = \lambda_0 |P_E a \otimes b|^2 + \lambda_0 |P_{E^\perp} a \otimes b|^2$$

hence if we define $\lambda_E = \lambda_0/(1 - \lambda_0)$, we have

$$|P_{E^\perp} a \otimes b|^2 \ge \lambda_E |P_E a \otimes b|^2 \tag{1.8}$$

for all $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$. Thus the quadratic form $q_E(X) = |P_{E^\perp} X|^2 - \lambda_E |P_E X|^2$ is a rank-one convex quadratic form as $q_E(a \otimes b) \ge 0$ for all $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$, hence is a quasiconvex function (see, e.g., [23]).

Now we consider a simple multiwell model. Let $E \subseteq M^{m\times n}$ be a linear subspace without rank-one matrices and let $K \subseteq E$ be a closed set. Consider the (Euclidean) squared-distance function $\mathrm{dist}^2(X, K) = \mathrm{dist}^2(P_E(X), K) + |P_{E^\perp} X|^2$, then (see [1])

$$F_{\lambda_E}(X) = \mathsf{co}(\mathrm{dist}^2(P_E(\cdot), K) + \lambda_E |P_E(\cdot)|^2)(X) + q_E(X)$$

is a quasiconvex lower bound of the squared-distance function as $\mathsf{co}(\mathrm{dist}^2(P_E(\cdot), K) + \lambda_E |P_E(\cdot)|^2)(X)$ is a convex function and $q_E(X) = |P_{E^\perp} X|^2 - \lambda_E |P_E X|^2$ is a rank-one convex quadratic form.

However, in the above construction, even when $K \subseteq E$ is a finite set, in general, the computation of the convex envelope $\mathrm{co}(\mathrm{dist}^2(P_E(\cdot), K) + \lambda_E|P_E(\cdot)|^2)(X)$ is not a simple task whose complexity is generally not known. As a general reference, it is known that to determine the convexity of a quartic polynomial is an NP-hard question [24]. The computation of the values of quasiconvex functions whose existence are known is one of the motivations for us to find computable quasiconvex functions with multiwell structure. For numerical computation of the rank-one convex envelope for general functions, we refer to [25]. On the other hand, for numerical computation of compensated convex transforms in the discrete setting, there are efficient methods [26].

Before we state our main results, let us first introduce some preliminary notions and results.

Definition 1. *(Quasiconvex functions [12–14]) Suppose $f : M^{m\times n} \to \mathbb{R}$ is continuous. Then f is quasiconvex if*

$$\int_G f(X + D\phi(x))dx \geq \int_G f(X)dx$$

$\forall X \in M^{m\times n}$, $\forall G \subseteq \mathbb{R}^n$ open and $\forall \phi \in C_0^\infty(G, \mathbb{R}^m)$.

For a continuous function $f : M^{m\times n} \to \mathbb{R}$ bounded below, the *quasiconvex envelope* $Qf : M^{m\times n} \to \mathbb{R}$ is the largest quesiconvex function satisfying $Q(f) \leq f$. For the precise definition, we refer to [13]. In this paper we only consider quasiconvex lower bound g of a given function f.

The *translation method* (see, e.g., [21]) is a simple and effective method for finding quasiconvex lower bounds. Suppose $f : M^{m\times n} \to \mathbb{R}$ is continuous and bounded below and assume that $g : M^{m\times n} \to \mathbb{R}$ is quasiconvex, then the function h defined by

$$h(X) := \mathrm{co}[f - g](X) + g(X) \tag{1.9}$$

is quasiconvex and satisfies $h(X) \leq f(X)$ for all $X \in M^{m\times n}$. We call h a translation lower bound which is quasiconvex as the sum of a convex function and a quasiconvex function remains quasiconvex.

The following example, which is a generalisation of a double-well function in [20] using a different method, can be used to show how the translation method applies to multiwell models when the wells are contained in subspaces without rank-one matrices [1]. This example is one of the motivations for the definition of compensated convex transforms [1].

Example 1. Let $E \subseteq M^{m\times n}$ be a linear subspace without rank-one matrices where m, $n \geq 2$. Let $K \subseteq E$ be a non-empty closed set.

Consider the squared Euclidean distance function to the set K

$$H(X) = \mathrm{dist}^2(X, K), \quad X \in M^{m\times n},$$

which is not quasiconvex. We see that

$$\mathrm{dist}^2(X, K) = |P_{E^\perp}X|^2 + \mathrm{dist}^2(P_E X, K).$$

Let $\lambda_E > 0$ be defined by

$$\lambda_E = \frac{\lambda_0}{1-\lambda_0} \quad \text{with} \quad \lambda_0 = \inf\{|P_{E^\perp} a \otimes b|^2 : \ a \in \mathbb{R}^m, \ b \in \mathbb{R}^n, \ |a| = |b| = 1\}. \tag{1.10}$$

Then the translation bound by the rank-one convex quadratic form $q_E(X) := |P_{E^\perp}X|^2 - \lambda_E|P_EX|^2$ given by

$$G(X) := \mathsf{co}[\mathrm{dist}^2(P_E\cdot,\ K) + \lambda_E|P_E\cdot|^2](X) + [|P_{E^\perp}X|^2 - \lambda_E|P_EX|^2] \tag{1.11}$$

is the quasiconvex envelope of $\mathrm{dist}^2(X,\ K)$. Note that rank-one convex quadratic functions are quasiconvex [23]. For detailed calculations, we refer to [1]. □

Next we briefly describe the main tool used in this paper: the compensated convex transforms.

Motivated from the translation method, in particular, formula (1.9), compensated convex transforms were introduced in [1].

Definition 2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function with at most quadratic growth and let $\lambda > 0$ be large if needed, we define λ-parametrised convexity-based transforms.*

The ***Lower compensated convex transform*** *is defined by*

$$C^l_\lambda(f)(x) := \mathsf{co}\left[\lambda|\cdot|^2 + f(\cdot)\right](x) - \lambda|x|^2, \quad x \in \mathbb{R}^n.$$

The ***Upper compensated convex transform*** *is defined by*

$$C^u_\lambda(f)(x) := \lambda|x|^2 - \mathsf{co}\left[\lambda|\cdot|^2 - f(\cdot)\right](x), \quad x \in \mathbb{R}^n$$

where $\mathsf{co}[g]$ = *convex envelope of g, and* $|\cdot|$ *denotes the Euclidean norm.*

We also have $C^u_\lambda(f)(x) = -C^l_\lambda(-f)(x)$, and both $C^l_\lambda(f)$ and $C^u_\lambda(f)$ are 'tight' approximations of f from below and above respectively as $\lambda \to +\infty$ in the sense that if f is of $C^{1,1}$ in a neiighbourhood of $x_0 \in \mathbb{R}^n$, then when $\lambda > 0$ is sufficiently large, we have $f(x_0) = C^l_\lambda(f)(x_0) = C^u_\lambda(f)(x)$.

Next we briefly describe the properties of the smooth approximation of the finite maximum function $f_m(x) = \max\{x_1, x_2, \ldots, x_m\}$ for $x = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$ by the upper compensated convex transform given by [1]

$$C^u_\lambda(f_m)(x) = \lambda|x|^2 - \lambda\mathrm{dist}^2\left(x,\ \mathsf{co}\left(\frac{\Delta_m}{2\lambda}\right)\right) + \frac{1}{4\lambda}.$$

If we consider the minimum function $g_m(x) = \min\{x_i\ :\ 1 \le i\} = -f_m(-x)$, then the lower compensated convex transform $C^l_\lambda(g_m)(x)$ satisfies

$$C^l_\lambda(g_m)(x) = -C^u_\lambda(f_m)(-x) = \lambda\mathrm{dist}^2\left(-x,\ \frac{\Delta_m}{2\lambda}\right) - \lambda|x|^2 - \frac{1}{4\lambda}.$$

It is known that there are numerical schemes to compute the convex projection $P_{\Delta_m/(2\lambda)}$ with complexity $O(m\log m)$, e.g., [2].

The following are some properties of $C^u_\lambda(f_m)$ and $C^l_\lambda(g_m)$. We will discuss their proofs in Section 5.

Proposition 2. *i) The function $C^u_\lambda(f_m)$ is a $C^{1,1}$ approximation of f_m with linear growth and*

$$|DC^u_\lambda(f_m)(x) - DC^u_\lambda(f_m)(y)| \le 8\lambda|x - y|.$$

ii) ***Tight approximation****: Assume $x = (x_1, x_2, \dots, x_m)$ is 'sorted' in the increasing order: $x_1 \geq x_2 \geq \cdots \geq x_m$, then $f_m(x) = C^u_\lambda(f_m)(x)$ if and only if $x_1 - x_2 \geq \frac{1}{2\lambda}$.*

iii) Similar to ii), if $x = (x_1, x_2, \dots, x_m)$ is 'sorted' in the decreasing order: $x_1 \leq x_2 \leq \cdots \leq x_m$, Then $g_m(x) = C^l_\lambda(g_m)(x)$ iff $x_2 - x_1 \geq \frac{1}{2\lambda}$.

iv) ***Error estimate****:*

$$f_m(x) \leq C^u_\lambda(f_m)(x) \leq f_m(x) + \frac{1}{2\lambda}.$$

vi) $C^u_\lambda(f_m)$ is a 'monotone' convex function [1]:
if $x \geq y$ in the sense that $x_i \geq y_i$ for $i = 1, 2, \dots, m$, then $C^u_\lambda(f_m)(x) \geq C^u_\lambda(f_m)(y)$.

The plan for the rest of the paper is as follows. In Section 2 we consider the multiwell function H_2 with quadratic growth. We give conditions on parameters involved to construct quasiconvex lower bounds H_2 by $C^l_\lambda(g_m)$ so that the resulting quasiconvex functions are computable. In Section 3 we consider the multiwell function H_1 with mixed linear-quadratic growth and take a slightly different approach from the method used for H_2. We present a generalisation of H_1 in Section 4 and conclude the paper with the proof of tight approximation of $C^u_\lambda(f_m)$.

2. A construction of quasiconvex lower bounds for H_2

In this section we construct quasiconvex lower bounds for the multiwell function H_2 by the lower compensated convex transform $C^l_\lambda(g_k)(\cdot)$ where $g_k(x) = \min\{x_1, x_2, \dots, x_k\}$ is the minimum function defined on $\mathbb{R}^k$.

Let $E \subseteq M^{m\times n}$ be a linear subspace without rank-one matrices. Let $K = \{A_1, A_2, \cdots, A_k\} \subseteq E$ and let $f : K \to E^\perp$ be a Liptschitz function with a small Lipschitz constant $L \geq 0$.

Recall the definition of $H_2 : M^{m\times n} \to \mathbb{R}$:

$$H_2(X) = \min\{|P_E(X - A_i)|^2 + \alpha|P_{E^\perp}X - f(A_i)|^2 + \beta_i : \ i = 1, 2, \dots, k\}, \qquad X \in M^{m\times n}. \tag{2.1}$$

In this section we give conditions on K, $\alpha > 0$, β_i and $L > 0$ so that the quasiconvex lower bound G_2 defined below preserves the shape of H_2 near the graph of f defined by $\Gamma_f = \{(A_i + f(A_i) : \ i = 1, 2, \dots, k\}$, i.e., $H_2(X) = G_2(X)$ if X is near Γ_f.

We define

$$q_i(X) := |P_E(X - A_i)|^2 + \alpha|P_{E^\perp}X - f(A_i)|^2 + \beta_i, \tag{2.2}$$

so that

$$H_2(X) = q(X) + \min\{\ell_i(X) : \ i = 1, 2, \dots, k\}, \qquad X \in M^{m\times n}, \tag{2.3}$$

where

$$\begin{aligned} q(X) &:= |P_E(X)|^2 + \alpha|P_{E^\perp}(X)|^2 - 2A_1 \cdot P_E(X) - 2\alpha f(A_1) \cdot P_{E^\perp}(X), \\ \ell_i(X) &= -2(A_i - A_1) \cdot P_E(X) - 2\alpha(f(A_i) - f(A_1)) \cdot P_{E^\perp}(X) + c_i, \end{aligned} \tag{2.4}$$

with $c_i = |A_i|^2 + \alpha|f(A_i))|^2 + \beta_i$. Note that $\ell_1(X) = c_1$ is a constant.

We also define

$$L_k(X) = (\ell_1(X), \ldots, \ell_k(X)). \tag{2.5}$$

The following is the main result of this section.

Theorem 3. *Suppose $E \subseteq M^{m\times n}$ is a linear subspace without rank-one matrices with the ellipticity constant $\lambda_E > 0$ defined by (1.10). Let*

$$D = \max_{i\neq j} |A_i - A_j|, \quad d = \min_{i\neq j} |A_i - A_j|, \quad \beta = \max_{i\neq j} |\beta_i - \beta_j|$$

Suppose $f : K \to E^\perp$ is a Lipschitz mapping with Lipschitz constant $0 \le L \le 1/\alpha$. If $d^2 > 2\beta$, $\alpha > 1$,

$$1 + \lambda_E\alpha > \frac{32(1+\lambda_E)(k-1)D^2}{d^2 - 2\beta} \quad \text{and } \lambda = \frac{1+\lambda_E\alpha}{8(1+\lambda_E)(k-1)D^2}. \tag{2.6}$$

Then

$$G_2(X) := q(X) - C^u_\lambda(f_k(-L_k(X))), \qquad X \in M^{m\times n} \tag{2.7}$$

is a quasiconvex lower bound of $H_2(X)$, where f_k is the maximum function in $\mathbb{R}^k$.

Also $G_2(X) = H_2(X)$ if

$$|P_E X - A_i|^2 + |P_{E^\perp} X - f(A_i)|^2 \le \frac{1}{4\lambda}, \quad i = 1, 2, \ldots, k. \tag{2.8}$$

Remark 4. *The assumption $d^2 > 2\beta$ implies that every 'energy well' $X_i := A_i + f(A_i)$ for $i = 1, 2, \ldots, k$ is a local minimum point of H_2.*

Also if X is close to the 'well' X_i, the quasiconvex lower bound $G_2(X)$ agrees with $H_2(X)$.

The computation of $G_2(X)$ requires the computation of $C^u_\lambda(f_k(-L_k(X))$ which has the complexity of $O(k\log k)$.

Proof of Theorem 3 By the the formula for the upper transform in the definition of $G_2(X)$, we have

$$G_2(X) := [q(X) - \lambda| - L_k(X)|^2] + \left(\lambda \mathrm{dist}^2\left(-L_k(X), \mathrm{co}\left(\frac{\Delta_k}{2\lambda}\right)\right) - \frac{1}{4\lambda}\right).$$

Since the function $\lambda \mathrm{dist}^2\left(-L_k(X), \mathrm{co}\left(\frac{\Delta_k}{2\lambda}\right)\right) - \frac{1}{4\lambda}$ is convex as $-L_k(X)$ is an affine mapping, hence this function is convex. So, we need to show:

i) the quadratic function $q(X) - \lambda|L_k(X)|^2$ is a rank one convex quadratic function, which implies that $G_2(X)$ is quasiconvex.

ii) near $A_i + f(A_i)$ we have $\ell_j(X) \ge \ell_i(X) + 1/(2\lambda)$ so that $-C^u_\lambda(f_k)(-L_k(X)) = \ell_i(X)$ for all $j \ne i$ hence $G_2(X) = H_2(X)$.

Proof of i) Let $\tau_k = 8(k-1)D^2$. We may write

$$
\begin{aligned}
& q(X) - \lambda| - L_k(X)|^2 = q(X) - \lambda|L_k(X)|^2 \\
& = |P_E X|^2 + \alpha|P_{E^\perp} X|^2 \\
& -4\lambda \textstyle\sum_{i=2}^k [((A_i - A_1)\cdot P_E X)^2 + \alpha^2((f(A_i) - f(A_1))\cdot P_{E^\perp}X)^2] \\
& -8\lambda\alpha \textstyle\sum_{i=2}^k ((A_i - A_1)\cdot P_E X)((f(A_i) - f(A_1))\cdot P_{E^\perp}X) \\
& +\text{affine terms} \\
& \geq (1-\tau_k\lambda)|P_E X|^2 + (\alpha - \tau_k\lambda)|P_{E^\perp}X|^2 + \text{affine terms}, \quad \text{where } \tau_k = 8(k-1)D^2.
\end{aligned}
$$

We observe that

$$\alpha - \tau_k\lambda = \frac{\alpha - 1}{1+\lambda_E} > 0$$

and

$$1 - \tau_k\lambda = \frac{\lambda_E(1-\alpha)}{1+\lambda_E} < 0.$$

Now let $X \in M^{m\times n}$ be a rank-one matrix, we have

$$(1-\tau_k\lambda)|P_E X|^2 + (\alpha - \tau_k\lambda)|P_{E^\perp}X|^2 \geq (\lambda_E\alpha + 1 - (1+\lambda_E)\tau_k\lambda)|P_E X|^2 = 0.$$

Thus $q(X) - \lambda|L_k(X)|^2$ is a rank-one convex quadratic function, hence is quasiconvex. Therefore $G_2(X)$ is quasiconvex on $M^{m\times n}$. □

Proof of ii) We need to show that for each i, if (2.8) holds, i.e.,

$$|P_E X - A_i|^2 + |P_{E^\perp}X - f(A_i)|^2 \leq \frac{1}{4\lambda},$$

then for each $j \neq i$, we have $\ell_j(X) \geq \ell_i(X) + \frac{1}{2\lambda}$. Equivalently we need to prove that

$$q(X) + \ell_i(X) + \frac{1}{2\lambda} \leq q(X) + \ell_j(X) \iff q_i(X) + \frac{1}{2\lambda} \leq q_j(X).$$

We have

$$
\begin{aligned}
& q_j(X) - q_i(X) - \tfrac{1}{2\lambda} \\
& = |A_j - A_i|^2 + \alpha|f(A_j) - f(A_i)|^2 - 2(P_E X - A_i)\cdot(A_j - A_i) \\
& -2\alpha(P_{E^\perp}X - f(A_i))\cdot(f(A_j) - f(A_i)) + \beta_j - \beta_i - \tfrac{1}{2\lambda} \\
& \geq |A_j - A_i|^2 - 2|A_j - A_i||P_E X - A_i| \\
& -2\alpha|f(A_j) - f(A_i)||P_{E^\perp}X - f(A_i)| - \beta - \tfrac{1}{2\lambda} \\
& \geq |A_j - A_i|^2 - \tfrac{\sqrt{2}}{\sqrt{\lambda}}|A_j - A_i| - \beta - \tfrac{1}{2\lambda}.
\end{aligned}
$$

Next we show that

$$|A_j - A_i|^2 - \frac{\sqrt{2}}{\sqrt{\lambda}}|A_j - A_i| - \beta - \frac{1}{2\lambda} > 0 \tag{2.9}$$

if (2.6) and (2.8) hold.

Let $x_1 < 0 < x_2$ be the two roots of the quadratic polynomial $x^2 - \frac{\sqrt{2}}{\sqrt{\lambda}}x - \beta - \frac{1}{2\lambda}$. It remains to show that $|A_j - A_i| > x_2$. From (2.6) we see that

$$\lambda > \frac{4}{d^2 - 2\beta},$$

so

$$
\begin{aligned}
x_2 =& \frac{1}{\sqrt{2\lambda}} + \sqrt{\beta + \frac{1}{\lambda}} \\
<& \frac{1}{\sqrt{2}} \sqrt{\frac{d^2 - 2\beta}{4}} + \sqrt{\frac{d^2 + 2\beta}{4}} \\
<& \frac{1}{2}\sqrt{d^2 - 2\beta} + \frac{1}{2}\sqrt{d^2 + 2\beta} \\
<& d < |A_j - A_i|.
\end{aligned}
$$

Note that by our assumption on f, we have $|f(A_j) - f(A_i)| \leq |A_i - A_j|/\alpha$. Thus by Property iii) in Proposition 2, we have $-C^u_\lambda(f_k(-L_k(X)) = C^l_\lambda(g_k(L_k(X)) = g_k(L_k(X)) = \ell_i(X)$ as $\ell_j(X) \geq \ell_i(X) + \frac{1}{2\lambda}$, hence $H_2(X) = G_2(X)$. □

Remark 5. *Due to Property iv) of the upper compensated convex transform in Propostion 2, we have the error estimate*

$$G_2(X) \leq H_2(X) \leq G_2(X) + \frac{1}{2\lambda}, \qquad X \in M^{m\times n}.$$

Therefore if $\alpha > 0$ is large, then $G_2(X)$ is a very good quasiconvex lapproximation of $H_2(X)$.

The structure of the quasiconvex lower bound $G_2(\cdot)$ of the multiwell function $H_2(\cdot)$ suggests that $G_2(\cdot)$ is of $C^{1,1}(M^{m\times n})$ and is of quadratic growth for $X \in M^{m\times n}$. Therefore when we consider the variational integral

$$I_2(u) = \int_\Omega G_2(Du(x))\,dx$$

for minimisers or more general stationary points [27] the natural space would be $W^{1,2}(\Omega, \mathbb{R}^m)$. This is in contrast with quasiconvex lower bounds G_1 for H_1 which will be discussed in the next section, where we will see that G_1 has a mixed growth which might lead to some more challenges for us to choose a proper function space to accommodate such energy density (integrands).

The following figure (Figure 1) gives an illustration of an example of a three well model H_2 restricted to $E \simeq \mathbb{R}^2$, where $f \equiv 0$:

$$H_2(x, y) = \min\{(x + 2)^2 + y^2 + 1,\ x^2 + 9(y + 2)^2 - 1,\ (x - 2)^2 + (y - 2)^2\}$$

with the three wells at $(-2, 0)$, $(0, -2)$, $(2, 2)$, with different heights and a quasiconvex lower bound ($\lambda = 0.25$) restricted on E.

3. A construction of quasiconvex lower bounds for H_1

In this section we construct quasiconvex lower bounds for the multiwell function H_1 defined in (1.6) by taking the lower compensated convex transform $C^l_\lambda(g_k)(\cdot) = -C^u_\lambda(-f_k)(\cdot)$, where $g_k(x)$ and $f_k(x)$ are the minimum and maximum functions defined on $\mathbb{R}^k$ respectively.

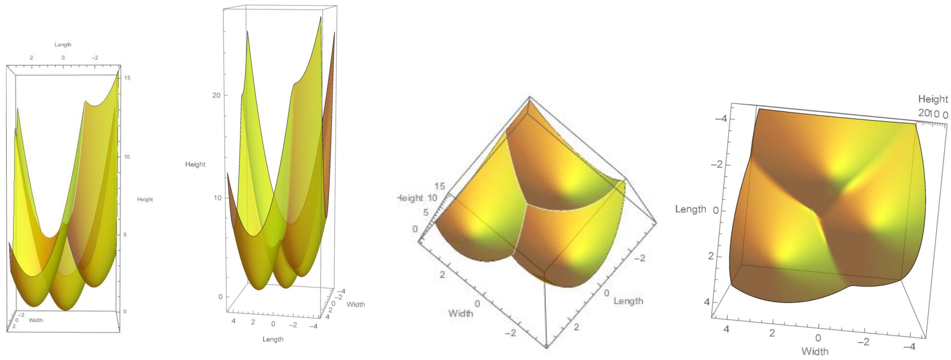

Figure 1. Views of $H_2(x, y)$ from two angles and a quasiconvex lower bound ($\lambda = 0.25$).

Similar to what we have defined in Section 2, let $E \subseteq M^{m\times n}$ be a linear subspace without rank-one matrices and let $K = \{A_1, A_2, \cdots, A_k\} \subseteq E$ be a finite set with k distinct elements.

We recall the definition of $H_1 : M^{m\times n} \to \mathbb{R}$:

$$H_1(X) = \alpha|P_{E^\perp}X|^2 + \min\{h_i(P_E X) : i = 1, 2, \ldots, k\}, \qquad X \in M^{m\times n}, \tag{3.1}$$

where

$$h_i(Y) = \sqrt{|\mathcal{U}_i(Y - A_i)|^2 + \gamma_i}, \quad , Y \in E$$

with $U_i : E \to E$ an invertible linear transform and $\gamma_i \geq 0$ for $i = 1, 2, \ldots, k$. We also define

$$q_i(X) = \alpha|P_{E^\perp}X|^2 + h_i(P_E X), \quad X \in M^{m\times n}$$

for $i = 1, 2, \ldots, k$.

Let $F_k(Y) = (h_1(Y),\ h_2(Y),\ \ldots,\ h_k(Y))$ for $Y \in E$. We define for $\lambda > 0$.

$$\begin{aligned} G_1(X) &= \alpha|P_{E^\perp}X|^2 - C^u_\lambda(f_k)(-F_k(P_E X)) \\ &= \lambda|\text{dist}^2\left(-F_k(P_E X),\ \tfrac{\Delta_k}{2\lambda}\right) + [\alpha|P_{E^\perp}X|^2 - \lambda|F_k(P_E X)|^2]. \end{aligned}$$

Note that $\alpha|P_{E^\perp}X|^2 - \lambda|F_k(P_E X)|^2$ is a quadratic function.

Let $d = \min_{1\leq i\neq j\leq k} |A_i - A_j|$, $\gamma = \max_{1\leq i\leq k} \gamma_i$, and $u_{\max} = \max_{1\leq i\leq k} \|\mathcal{U}_i\|_{op}$, where $\|\mathcal{U}_i\|_{op}$ is the operator norm of $\mathcal{U}_i$. Let $u_j = \inf\{|U_j Y| : Y \in E, |Y| = 1\}$ for $j = 1, 2, \ldots, k$. Since $U_j : E \to E$ is an invertible linear transform for $j = 1, 2, \ldots, k$, we see that $u_j > 0$ for all $j = 1, 2, \ldots, k$. We define $u_{\min} = \min\{u_j : j = 1, 2, \ldots, k\}$. Clearly, $u_{\min} > 0$. If γ is small enough and $\alpha > 0$ is sufficiently large, we can establish a similar result for H_1 as we have done for H_2.

Theorem 6. *Suppose $u_{\min} d > \sqrt{2\gamma}$. Let $\lambda_E > 0$ be defined by (1.10). If*

$$\lambda := \frac{\lambda_E \alpha}{k u_{\max}^2} > \frac{(1 + \sqrt{2})u_{\max}^2 + 1/\sqrt{2}}{u_{\min} d - \sqrt{2\gamma}},$$

then

$$G_1(X) = \alpha|P_{E^\perp}X|^2 - C^u_\lambda(f_k)(-F_k(P_E X))$$

is a quasiconvex lower bound of $H_1(X)$ *for* $X \in M^{m\times n}$ *and* $G_1(X) = H_1(X)$ *if* $|P_E X - A_i| \le 1/\lambda$ *for* $i = 1, 2, \ldots, k$.

The condition for α is a rough sufficient condition as we have just used the fact

$$\frac{\sqrt{a}+\sqrt{b}}{\sqrt{2}} \le \sqrt{a+b} \le \sqrt{a}+\sqrt{b}, \quad \text{for } a,\ b \ge 0$$

in our estimates for $g_j(P_E X) \ge g_i(P_E X) + 1/(2\lambda)$ for $j \ne i$.

The condition $u_{\min} d > \sqrt{2\gamma}$ is essentially a sufficient geometric assumption which means the minimum distance among the wells $|A_j - A_i|$ is larger that the maximum height $\sqrt{\gamma}$ so that any given well A_i is a genuine 'well' with A_i a local minimum point of the energy density $G_1(X)$ not swallowed by other wells.

Proof of Theorem 6 Observe that by the formula for the upper transform of the maximum function we have

$$G_1(X) = [\alpha|P_{E^\perp}X|^2 - \lambda|F_k(P_E X)|^2] + \lambda \mathrm{dist}^2\left(-F_k(P_E X),\ \mathrm{co}\left(\frac{\Delta_k}{2\lambda}\right)\right) - \frac{1}{4\lambda}.$$

We need to show that $\alpha|P_{E^\perp}X|^2 - \lambda|F_k(P_E X)|^2$ is a rank-one convex quadratic function and

$$X \mapsto \mathrm{dist}^2\left(-F_k(P_E X),\ \mathrm{co}\left(\frac{\Delta_k}{2\lambda}\right)\right)$$

is convex, that is, for $Y \in E$,

$$Y \mapsto \mathrm{dist}^2\left(-F_k(Y),\ \frac{\Delta_k}{2\lambda}\right)$$

is convex.

Note that $h_j(Y) = \sqrt{|\mathcal{U}_j(Y - A_j)|^2 + \gamma_j^2}$ is a non-negative convex function for $Y \in E$ and for $j = 1, 2, \ldots, k$. We define, for $u \in \mathbb{R}^k$, the function

$$f(u) = \mathrm{dist}^2\left(-u,\ \frac{\Delta_k}{2\lambda}\right).$$

Clearly $f(u)$ is convex. The key observation is that $f(u)$ is also positively increasing in the sense that if $u \ge v \ge 0$, i.e., $u_i \ge v_i \ge 0$ for all of the corresponding components, then $f(u) \ge f(v)$. This requires some more detailed structural properties of the convex projection $P_{\Delta_k/(2\lambda)}$. We have

Lemma 7. *Let*

$$f(u) = \mathrm{dist}^2\left(-u,\ \frac{\Delta_m}{2\lambda}\right), \quad u \in \mathbb{R}^m.$$

If $u = (u_1, \ldots, u_m)$, $h = (h_1, \ldots, h_m) \in \mathbb{R}^m$ *satisfy that* $u \ge 0$ *and* $h \ge 0$ *componentwise in the sense that* $u_i \ge 0$ *and* $h_i \ge 0$ *for* $i = 1, 2, \ldots, m$, *then* $Df(u) \cdot h \ge 0$. *Consequently,* $f(u) \ge f(v)$ *if* $u \ge v \ge 0$ *componentwise.*

Proof We have

$$f(u) = \text{dist}^2\left(-u, \frac{\Delta_m}{2\lambda}\right) = \left|-u - P_{\frac{\Delta_m}{2\lambda}}(-u)\right|^2 = \left|u + P_{\frac{\Delta_m}{2\lambda}}(-u)\right|^2,$$

where the convex projection is in the form

$$P_{\frac{\Delta_m}{2\lambda}}(-u) = \sum_{j=1}^{m} \frac{\lambda_j}{2\lambda} e_j$$

with $\lambda_j \geq 0$ for $j = 1, 2, \ldots, m$ and $\sum_{j=1}^m \lambda_j = 1$. If $u \geq 0$ and $h \geq 0$ componentwise, then

$$Df(u) \cdot h = D\text{dist}^2\left(-u, \tfrac{\Delta_m}{2\lambda}\right) \cdot h = 2\left(u + P_{\frac{\Delta_m}{2\lambda}}(-u)\right) \cdot h$$

$$= 2\textstyle\sum_{j=1}^m u_j h_j + 2\sum_{i=1}^m \frac{\lambda_i}{2\lambda} e_i \cdot h = 2\sum_{j=1}^m u_j h_j + 2\sum_{i=1}^m \frac{\lambda_i}{2\lambda} h_i \geq 0$$

as $u_i h_i \geq 0$, $\lambda_i \geq 0$ and $e_i \cdot h = h_i \geq 0$ for $i = 1, 2, \ldots, m$ with $e_1, e_2, \ldots, e_m$ the standard Euclidean basis vectors. The last claim follows from the fundamental theorem of calculus that if $u \geq v \geq 0$, i.e., $u_i \geq v_i \geq 0$ for $i = 1, 2, \ldots, m$, then $f(u) - f(v) = \int_0^1 Df(tu + (1-t)v) \cdot (u - v)\, dt \geq 0$ as $tu + (1-t)v \geq 0$ and $u - v \geq 0$ componentwise. □

Proof of Theorem 6 (continued) Now we can show that

$$V_1(X) := \text{dist}^2\left(-F_k(P_E X), \text{co}\left(\frac{\Delta_k}{2\lambda}\right)\right) = f(F_k(P_E X)), \quad X \in M^{m\times n}$$

is convex, where $f : \mathbb{R}^k \to R$ is defined as in Lemma 7 with $m = k$. Let $0 < t < 1$ and $X, Y \in M^{m\times n}$, since $F_k(P_E X) = (h_1(P_E(X)), h_2(P_E(X)), \ldots, h_k(P_E(X)))$ and every component function $h_j(P_E(X)) \geq 0$ is convex, we first have

$$h_j(P_E(tX + (1-t)Y)) = h_j(tP_E X + (1-t)P_E Y)) \leq t h_j(P_E X) + (1-t) h_j(P_E Y), \quad j = 1, 2, \ldots, k$$

hence $F_k(P_E(tX + (1-t)Y)) \leq tF_k(P_E X) + (1-t)F_k(P_E Y)$ componentwise.

Since f is positively increasing, we have

$$V_1(tX + (1-t)Y) = f(F_k(P_E(tX + (1-t)Y))) \leq f(tF_k(P_E X) + (1-t)F_k(P_E Y)).$$

Also f is convex, which implies

$$f(tF_k(P_E X) + (1-t)F_k(P_E Y)) \leq tf(F_k(P_E X)) + (1-t)f(F_k(P_E Y)) = tV_1(X) + (1-t)V_1(Y).$$

Therefore $V_1(X)$ is a convex function in $M^{m\times n}$.

Next we show that the quadratic function $Q_1(X) = \alpha|P_{E^\perp}X|^2 - \lambda|F_k(P_E X)|^2$ is a rank-one convex quadratic function in $M^{m\times n}$. We have for every $X \in M^{m\times n}$, we have

$$Q_1(X) = \alpha|P_{E^\perp}X|^2 - \lambda\textstyle\sum_{j=1}^k h_j^2(P_E X) = \alpha|P_{E^\perp}X|^2 - \lambda\sum_{j=1}^k(|\mathcal{U}_i(P_E X - A_i)|^2 + \gamma_i)$$
$$\alpha|P_{E^\perp}X|^2| - (\lambda\textstyle\sum_{j=1}^k(|\mathcal{U}_i(P_E X)|^2) + L(P_E X),$$

where

$$L(P_EX) = -\lambda \sum_{j=1}^{k} (-2\mathcal{U}_i(P_EX) \cdot \mathcal{U}_i(A_i) + |\mathcal{U}_i(A_i)|^2 + \gamma_j)$$

is an affine function of P_EX hence is a convex function. So, we only need to show that the quadratic form

$$\alpha|P_{E^\perp}X|^2 - \lambda \sum_{j=1}^{k} |\mathcal{U}_i(P_EX)|^2$$

is rank-one convex. We have, for every rank-one matrix $X \in M^{m\times n}$,

$$\alpha|P_{E^\perp}X|^2| - (\lambda \textstyle\sum_{j=1}^{k} |\mathcal{U}_i(P_EX)|^2 \geq \alpha\lambda_E|P_EX|^2| - \lambda u_{\max}^2 k|P_EX|^2 = (\alpha\lambda_E - \lambda u_{\max}^2 k)|P_EX|^2 = 0,$$

as $\alpha\lambda_E = \lambda u^2 k$.

Next we prove that $G_1(X) = H_1(X)$ if $|P_EX - A_i| \leq 1/\lambda$ for each $i = 1, 2, \ldots, k$. By definition of the upper transform, $G_1(X) = H_1(X)$ if and only if $-C_\lambda^u(f_k)(-F_k(P_EX)) = \min\{h_1(P_EX), h_2(P_EX), \ldots, h_k(P_EX)\}$. By Property iii) in Proposition 2, if $|P_EX - A_i| \leq 1/\lambda$, we show that $h_j(P_EX) \geq h_i(P_EX) + 1/(2\lambda)$ for all $j = 1, 2, \ldots, k$ with $j \neq i$ so that $-C_\lambda^u(f_k)(-F_k(P_EX)) = h_i(P_EX)$ hence $G_1(X) = H_1(X)$. If we write $P_EX = A_i + Y \in E$, the assumption $|P_EX - A_i| \leq 1/\lambda$ implies $|Y| \leq 1/\lambda$. We have

$$h_j(P_EX) \geq h_i(P_EX) + \tfrac{1}{2\lambda} \iff \sqrt{|\mathcal{U}_j(P_EX - A_j)|^2 + \gamma_j} \geq \sqrt{|\mathcal{U}_i(P_EX - A_i)|^2 + \gamma_i} + \tfrac{1}{2\lambda}$$

$$\Longleftarrow \tfrac{|\mathcal{U}_j(P_EX - A_j)| + \sqrt{\gamma_j}}{\sqrt{2}} \geq |\mathcal{U}_i(P_EX - A_i)| + \sqrt{\gamma_i} + \tfrac{1}{2\lambda}$$

$$\iff \tfrac{|\mathcal{U}_j(A_i - A_j + Y)| + \sqrt{\gamma_j}}{\sqrt{2}} \geq |\mathcal{U}_i(Y)| + \sqrt{\gamma_i} + \tfrac{1}{2\lambda}$$

$$\Longleftarrow \tfrac{|\mathcal{U}_j(A_i - A_j)| - |\mathcal{U}_j(Y)|}{\sqrt{2}} \geq |\mathcal{U}_i(Y)| + \sqrt{\gamma_i} + \tfrac{1}{2\lambda}$$

$$\Longleftarrow \tfrac{u_{\min}|A_i - A_j|}{\sqrt{2}} \geq \tfrac{|\mathcal{U}_j(Y)|}{\sqrt{2}} + |\mathcal{U}_i(Y)| + \sqrt{\gamma} + \tfrac{1}{2\lambda}$$

$$\Longleftarrow u_{\min}d \geq \sqrt{2}\left(\tfrac{u_{\max}|Y|}{\sqrt{2}} + u_{\max}|Y| + \sqrt{\gamma} + \tfrac{1}{2\lambda}\right)$$

$$\Longleftarrow u_{\min}d - \sqrt{2\gamma} \geq \tfrac{(1+\sqrt{2})u_{\max}}{\lambda} + \tfrac{1}{\sqrt{2}\lambda} \qquad \text{as } |Y| \leq 1/\lambda$$

$$\iff \lambda \geq \tfrac{(1+\sqrt{2})u_{\max} + 1/\sqrt{2}}{u_{\min}d - \sqrt{2\gamma}}.$$

The last inequality holds as we have assumed that $u_{\min}d > \sqrt{2\gamma}$ and $\lambda = \dfrac{\lambda_E\alpha}{ku_{\max}^2}$ is large if $\alpha > 0$ is large.

□

Remark 8. *In Theorem 6, if we consider the special case that $\mathcal{U}_i = I$ - the identity transform, then we have $u_{\max} = u_{\min} = 1$ and in this special case the assumptions will be much simpler as we can assume that*

$$\lambda = \frac{\lambda_E\alpha}{k} > \frac{1 + \sqrt{2} + 1/\sqrt{2}}{d - \sqrt{2\gamma}}.$$

Remark 9. *We may generalise Theorem 6 to deal with more complicated multiwell models. Even for a single non-elliptic well model in the form*

$$H_1(X) = \alpha|P_{E^\perp}X|^2 + \min\{|\mathcal{U}_i P_E X| : \ i = 1, 2, \ldots, k\}, \quad X \in M^{m\times n}, \tag{3.2}$$

where $U_i : E \to E$ is an invertible linear transform, we see that under the assumption that

$$\lambda = \frac{\lambda_E \alpha}{k}, \tag{3.3}$$

we can show that the corresponding lower bound $G_1(X)$ is still a quasiconvex lower bound. However, at $X = 0$ we have $H_1(0) = 0$ but $G_1(0) < 0$. This is due to the fact that

$$-C^u_\lambda(f_k(-F_k(0))) = \lambda \mathrm{dist}^2(0,\ \Delta_{k/(2\lambda)}) - \tfrac{1}{4\lambda} = \lambda|P_{\Delta_{k/(2\lambda)}}(0)|^2 - \tfrac{1}{4\lambda}$$

$$= \lambda \left|\textstyle\sum_{j=1}^k \frac{e_j}{2k\lambda}\right|^2 - \tfrac{1}{4\lambda} = -\tfrac{k-1}{4k\lambda} < 0.$$

Here we have used the fact that the distance between 0 *and the simplex $\Delta_k/(2\lambda)$ is attained at the center of the simplex* $\displaystyle\sum_{j=1}^{k} \frac{e_j}{2k\lambda}$.

At any point where $|\mathcal{U}_j P_E X| \geq |\mathcal{U}_i P_E X| + 1/(2\lambda)$ for all $j \neq i$, we still have $H_1(X) = G_1(X)$. Thus if we just wish to construct a quasiconvex function with the 'desired' geometric feature which is differentiable except at 0, *then we can make a simple lift by considering $G_1(X) + (k-1)/(4k\lambda)$.*

Remark 10. *From Theorem 6 we see that both the multiwell function H_1 and its quasiconvex lower bound G_1 are of mixed growth. In the subspace $E^\perp \subseteq M^{m\times n}$ both H_1 and G_1 are of quadratic growth. In the subspace E, both H_1 and G_1 are of linear growth.*

If the height $\gamma_j > 0$ for all $j = 1, 2, \ldots, k$, then we see that G_1 is at least of $C^{1,1}_{loc}(M^{m\times n})$. However, if for some j, $\gamma_j = 0$, then G_1 is not differentiable at A_j.

For the variational integral $I(u) = \int_\Omega G_1(Du)\,dx$, if we consider the Dirichlet problem, say $u = 0$ on $\partial\Omega$, the natural space is $W^{1,2}_0(\Omega,\ \mathbb{R}^m)$ where $\Omega \subseteq \mathbb{R}^n$ is, say, a bounded Lipchitz domain. The main reason is that $I(u)$ is coercive in $W^{1,2}_0(\Omega,\ \mathbb{R}^m)$ because we have

$$\int_\Omega |P_{E^\perp} D\phi(x)|^2\,dx \geq \lambda_E \int_\Omega |P_E D\phi(x)|^2\,dx$$

for $\phi \in W^{1,2}_0(\Omega,\ R^m)$. As $G_1(P_E X) \geq c_0$ for some $c_0 \in \mathbb{R}$ we have both $\int_\Omega G_1(Du)\,dx \geq \alpha|P_{E^\perp}Du|^2\,dx + c_0 \mathrm{meas}(\Omega)$ and $\int_\Omega G_1(Du)\,dx \geq \alpha\lambda_E|P_E Du|^2\,dx + c_0\mathrm{meas}(\Omega)$.

If we consider the natural boundary value problem for $I(u) = \int_\Omega G_1(Du)\,dx$ under the constraint, say $\int_\Omega u(x)\,dx = 0$, then $W^{1,2}(\Omega,\ \mathbb{R}^m)$ does not seem to be the right space to study such a variational integral as $I(u)$ is not coercive in this space. We have $\alpha|P_{E^\perp}X|^2 - C^u_\lambda(f_k)(-F_k(P_E X))$ and it can be verified that there are $c_0 > 0$, $c_1 > 0$, $C_0 > 0$ and $C_1 > 0$ such that $c_0|P_E X| - c_1 \leq -C^u_\lambda(f_k)(-F_k(P_E X)) \leq C_0|P_E X| + C_1$. However, the integral $\int_\Omega |P_{E^\perp}Du|^2\,dx$ does not contribute to the coercivity of $I(u)$ in the subspace

E. An example is in [28] that if we consider the two dimensional conformal subspace $E_\partial \subseteq M^{2\times 2}$ which does not have rank-one matrices, then $E_\partial^\perp = E_{\bar\partial}$ is the two-dimensional anti-conformal subspace. If $u = u_1 + iu_2$ is a holomophic function in Ω and we define $v := (u_1, u_2) : \Omega \to \mathbb{R}^2$, then we have $|P_{E_{\bar\partial}} Dv|^2 = |\bar\partial u|^2 = 0$ by the Cauchy-Riemann equations. Therefore $\int_\Omega |P_{E_{\bar\partial}} Dv|^2\, dx$ does not contribute to the coercivity of $I(v)$ under the condition $\int_\Omega v\, dx = 0$.

It seems that some spaces with mixed growth condition might be the correct spaces to accommodate such variational integrals. Furthermore the study of (partial) regularity of minimisers and more general critical points for $I(u)$ under the natural boundary condition seems to be a challenging question.

The following figure (Figure 2) gives and illustration of an example of a three well model H_1 restricted to $E \simeq \mathbb{R}^2$ in the form

$$\min\left\{ \sqrt{9(x+2)^2 + y^2 + 1},\ \sqrt{x^2 + 9(y+2)^2 + 0.5^2},\ \sqrt{(x-2)^2 + (y-2)^2} \right\}$$

with the three 'anisotropic' wells and with different heights at $(-2, 0)$, $(0, -2)$, $(2, 2)$ $(\lambda = 1/2)$.

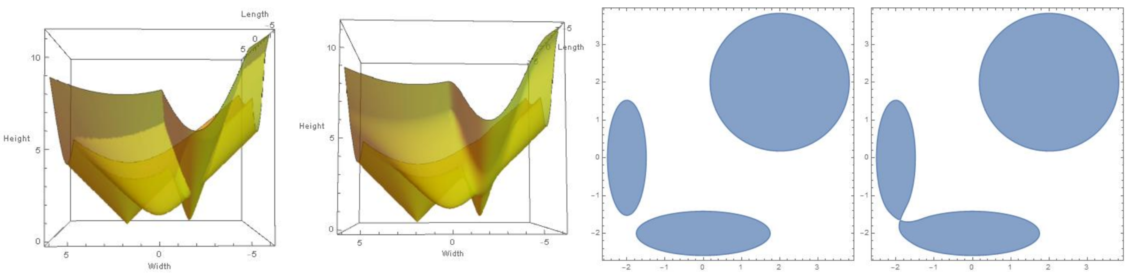

Figure 2. $H_1(x, y)$, lower bound $G_1(x, y)$ with $\lambda = 1/2$, and their sublevel set at 1.8211.

The following figure (Figure 3) gives and illustration of an example of a single non-elliptic well model H_1 restricted to $E \simeq \mathbb{R}^2$ in the form

$$\min\left\{ \sqrt{x^2/100 + y^2},\quad \sqrt{(x\cos(\pi/3) + \sin(\pi/3)^2/100 + (x\sin(\pi/3) - y\cos(p\pi/3))^2},\right.$$

$$\left.\sqrt{(x\cos(2\pi/3) + \sin(2\pi/3)^2/100 + (x\sin(2\pi/3) - y\cos(2\pi/3))^2}\right\}$$

with the three 'anisotropic' wells and with different heights at $(-2, 0)$, $(0, -2)$, $(2, 2)$ $(\lambda = 1/2)$.

4. A generalised H_1 with affine wells

Suppose $\dim(E) = s$. Using an orthonormal basis of E given by $B_1, B_2, \ldots, B_s$ and let $x_i = P_E X \cdot B_i$ and $x = (x_1, x_2, \ldots, x_s) \in \mathbb{R}^s$, we define $L_i(x) = A_i x - b_i$ for $y \in \mathbb{R}^s$ with $A_i \in M^{m_i\times s}$ and $b_i \in \mathbb{R}^{m_i}$ with $1 \le m_i < s$ and $i = 1, 2 \ldots, k$.

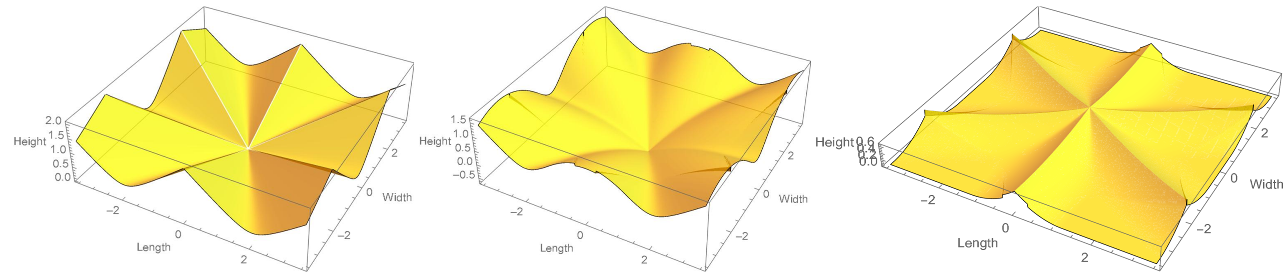

Figure 3. $H_1(x, y)$, lower bound $G_1(x, y)$ with $\lambda = 0.25$, and their difference $H_1(x, y) - G_1(x, y)$.

We may consider the following more general model

$$H_1(X) = \alpha |P_{E^\perp} X|^2 + V_0(P_E X),$$

where

$$V_0(P_E(X) = \min\{|A_i x - b_i| : \ i = 1, 2, \ldots, k\}, \quad x \in \mathbb{R}^d,$$

or more generally,

$$V_\gamma(x) = \min\{ \sqrt{|A_i x - b_i|^2 + \gamma_i^2} : \ i = 1, 2, \ldots, k\}, \quad x \in \mathbb{R}^d,$$

where $|A_i x - b_i|$ is the Euclidean norm in $\mathbb{R}^{m_i}$ and $\gamma_i \geq 0$. For the function V_0, the zero set of can be the union of finitely many planes.

Let $F_k(x) = (|A_1 x - b_1|, \ldots, |A_k x - b_k|)$. We can approximate $H_1(x)$ from below by $G_1(x) = \alpha |P_{E^\perp} X|^2 - C^u_\lambda(f_k)(-F_k(x))$, with $x = (P_E X \cdot B_1, .(P_E X \cdot B_1, \ldots, (P_E X \cdot B_k)$.

We can use G_1 to define quasiconvex functions as before. However due to the special feature of $-C^u_\lambda(f_k)(-x)$, at intersections of the planes defined by the zero set of $V_0(x)$, we see that $-C^u_\lambda(f_k)(-F_k(x)) < 0$ at points of intersections. As we commented earlier, this is due to the fact that if $x_1 = x_2 \leq x_3 \leq \cdots \leq x_m$, then by Lemma 12 in Section 5 and the fact that $C^l_\lambda(g_2(x_1, x_2)) = -C^u_\lambda(f_2(-x_1, -x_2))$ in $\mathbb{R}^2$, we have

$$C^l_\lambda(g_m)(x) = -C^u_\lambda(f_m(-x)) \leq -C^u_\lambda(f_2(-x_1, -x_2)) = x_1 - \frac{1}{8\lambda} = g_m(x) - \frac{1}{8\lambda}.$$

If $x_1 = x_2 = 0$, we see that

$$-\frac{1}{4\lambda} \leq -C^u_\lambda(f_m(-x)) \leq -\frac{1}{8\lambda} < 0.$$

Example 11. Let $H_1(x, y) = 2\min\{|x|, |y|, \sqrt{(x-1)^2 + (y-1)^2}\}$ and let $G_1(x, y) = -C^u_\lambda(f_3)(-F_3(x, y))$

with $F_3(x, y) = (|x|, |y|, \sqrt{(x-1)^2 + (y-1)^2}) \in \mathbb{R}^3$. The graphs of $H_1(x, y)$ and $G_1(x, y)$ with $\lambda = 1/2$ are shown in Figure 4.

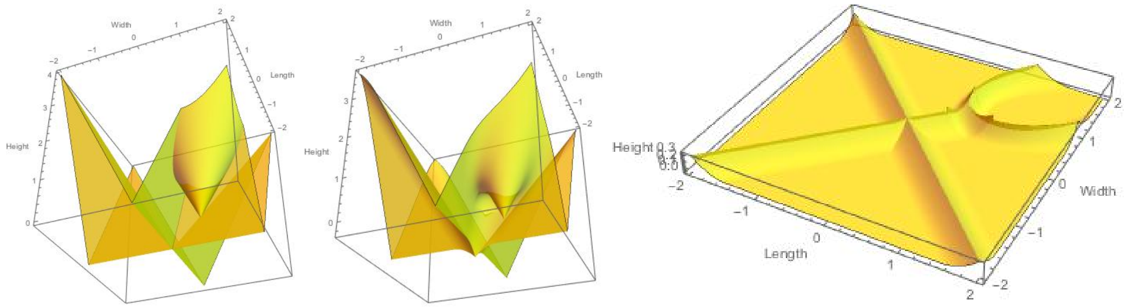

Figure 4. Views of H_1, G_1 and the difference $H_1 - G_1$ in the 2d subspace E ($\lambda = 0.5$).

4.1. Proof of the tight approximation of the upper trasnform of the maximum function

In this section we prove the tight approximation of the upper compensated convex transform of the maximum function (Proposition 2 (ii)) . Recall from [1] that

$$C^u_\lambda(f_m)(x) = \lambda|x|^2 - \lambda \mathrm{dist}^2\left(x,\ \mathrm{co}\left(\frac{\Delta_m}{2\lambda}\right)\right) + \frac{1}{4\lambda}$$

for $x \in \mathbb{R}^m$ and $\lambda > 0$. Items i) and iv) of Proposition 2 were established in [1] with the maximum function as a special case. The fact that The function $C^u_\lambda(f_m)$ is a $C^{1,1}$ approximation of f_m and satisfies

$$|DC^u_\lambda(f_m)(x) - DC^u_\lambda(f_m)(y)| \le 8\lambda|x - y|,$$

for $x,\ y \in \mathbb{R}^m$ is a consequence of a general result in [1] (Theorem 4.1). The error estimate in iv) that

$$f_m(x) \le C^u_\lambda(f_m)(x) \le f_m(x) + \frac{1}{2\lambda}$$

for $x \in \mathbb{R}^m$ was established in [1] (Theorem 5.1) which also contains the estimate $|DC^u_\lambda(f_m)(x)| \le 1$ for $x \in \mathbb{R}^m$. Thus the statement in i) that $C^u_\lambda(f_m)$ is of linear growth is a direct consequence of this gradient estimate.

Before we prove Proposition 2 (ii) we state the following simple lemma which can be verified through a direct calculation using the definition of the upper compensated convex transform.

Lemma 12. *Let* $f_2(x) = \max\{x_1, x_2\}$ *for* $x = (x_1, x_2) \in \mathbb{R}^2$. *We have*

$$C^u_\lambda(f_2)(x) = \lambda|x|^2 - \lambda\mathrm{dist}^2(x,\ \Delta_2/(2\lambda)) + \frac{1}{4\lambda}$$

$$= \begin{cases} \dfrac{x_1 + x_2}{2} + \dfrac{\lambda}{2}(x_1 - x_2)^2 + \dfrac{1}{8\lambda}, & |x_1 - x_2| \le \dfrac{1}{2\lambda}, \\ f_2(x), & |x_1 - x_2| \ge \dfrac{1}{2\lambda}. \end{cases} \tag{4.1}$$

Next we prove ii), the 'Tight approximation' property: Assume $x = (x_1, x_2, \dots, x_m)$ is 'sorted' in the increasing order: $x_1 \ge x_2 \ge \cdots \ge x_m$, then $f_m(x) = C^u_\lambda(f_m)(x) = x_1$ if and only if $x_1 - x_2 \ge \frac{1}{2\lambda}$. Note that Item iii) for the lower transform of the minimum function can be proved by using the identity $C^l_\lambda(g_m)(x) = -C^u_\lambda(f_m(-x))$.

Proof of Proposition 2 Item ii) If $x^{(0)}_1 - x^{(0)}_2 \ge 1/(2\lambda)$, we see that f_m is differentiable at $x^{(0)}$. By the translation invariance property of compensated convex transforms [7], we have,

$$C^u_\lambda(f_m)(x^{(0)}) = -\operatorname{co}(\lambda|\cdot - x^{(0)}|^2 - f_m(\cdot))(x^{(0)}).$$

The conclusion in this case follows if we can show that $\operatorname{co}(\lambda|\cdot - x^{(0)}|^2 - f_m(\cdot))(x^{(0)}) = -f_m(x^{(0)})$. To prove this we only need to show that the tangent plane for $-f_m$ defined by $\ell(x) = -f_m(x^{(0)}) - \nabla f_m(x^{(0)})\cdot(x - x^{(0)})$ lies below the graph of the function $\lambda|x - x^{(0)}|^2 - f_m(x)$ for all $x \in \mathbb{R}^m$, that is,

$$-f_m(x^{(0)}) - \nabla f_m(x^{(0)}) \cdot (x - x^{(0)}) \le \lambda|x - x^{(0)}|^2 - f_m(x), \qquad x \in \mathbb{R}^m. \tag{4.2}$$

Since $f_m(x) = x_1$ for x near $x^{(0)}$, we have $f_m(x^{(0)}) = x^{(0)}_1$ and $\nabla f_m(x^{(0)}) = e_1$, hence (4.2) is equivalent to

$$-x^{(0)}_1 - (x_1 - x^{(0)}_1) \le \lambda|x - x^{(0)}|^2 - x_k, \tag{4.3}$$

where $x_k = f_m(x)$ for some $k \in \{1, 2, \dots, m\}$ by definition. If $k = 1$, then inequality (4.3) clearly holds. If $k \ne 1$, we have, as (4.3) is equivalent to

$$A := \lambda|x - x^{(0)}|^2 - (x_k - x^{(0)}_k) + (x_1 - x^{(0)}_1) + (x^{(0)}_1 - x^{(0)}_k) \ge 0. \tag{4.4}$$

If we complete squares in A defined in (4.4) above, we obtain

$$A = \sum_{j\ne k,\,1}^{m} \lambda(x_j - x^{(0)}_j)^2 + \lambda\left(x_k - x^{(0)}_k - \frac{1}{2\lambda}\right)^2 + \lambda\left(x_1 - x^{(0)}_1 + \frac{1}{2\lambda}\right)^2 + (x^{(0)}_1 - x^{(0)}_k) - \frac{1}{2\lambda}.$$

Therefore $A \ge 0$ if and only if

$$x^{(0)}_1 - x^{(0)}_k - \frac{1}{2\lambda} \ge 0.$$

Since $k \ne 1$ we have

$$x^{(0)}_1 - x^{(0)}_k \ge x^{(0)}_1 - x^{(0)}_2 \ge \frac{1}{2\lambda}.$$

The conclusion follows. □

Acknowledgments

The work of the first author is partially supported by NSFC grants 11801200 and 12141107. We wish to thank the anonymous referees for their helpful comments and suggestions which have greatly improved the presentation of this paper.

Conflict of interest

The authors declare there is no conflicts of interest.

References

1. K. Zhang, Compensated convexity and its applications, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **25** (2008), 743–771. https://doi.org/10.1016/j.anihpc.2007.08.001
2. Y. Chen, X. Ye, Projection onto a simplex, *arXiv preprint*, (2011), `arXiv:1101.6081`.
3. C. W. Combettes, S. Pokutta, Complexity of linear minimization and projection on some sets, *Oper. Res. Lett.*, **49** (2021), 565–571. https://doi.org/10.1016/j.orl.2021.06.005
4. L. Condat, Fast projection onto the simplex and the l_1 ball, *Math. Program.*, **158** (2016), 575–585. https://doi.org/10.1007/s10107-015-0946-6
5. E. Y. Pee, J. O. Royset, On solving large-scale finite minimax problems using exponential smoothing, *J. Optim. Theory Appl.*, **148** (2011), 390–421. https://doi.org/10.1023/B:JOTA.0000006685.60019.3e
6. K. Zhang, Convex analysis based smooth approximations of maximum functions and squared distance functions, *J. Nonlinear Convex Anal.*, **9** (2008), 379–406.
7. K. Zhang, A. Orlando, E. Crooks, Compensated convexity and Hausdorff stable geometric singularity extractions, *Math. Models Methods Appl. Sci.*, **25** (2015), 747–801. https://doi.org/10.1142/S0218202515500189
8. K. Zhang, A. Orlando, E. Crooks, Compensated convexity and Hausdorff stable extraction of intersections for smooth manifolds, *Math. Models Methods Appl. Sci.*, **25** (2015), 839–873. https://doi.org/10.1142/S0218202515500207
9. K. Zhang, E. Crooks, A. Orlando, Compensated convexity, multiscale medial axis maps and sharp regularity of the squared distance function, *SIAM J. Math. Anal.*, **47** (2015), 4289–4331. https//doi.org/10.1137/140993223
10. K. Zhang, E. Crooks, A. Orlando, Compensated convexity methods for approximations and interpolations of sampled functions in Euclidean spaces: theoretical foundations, *SIAM J. Math. Anal.*, **48** (2016), 4126–4154. https://doi.org/10.1137/15M1045673
11. K. Zhang, E. Crooks, A. Orlando, Compensated convexity methods for approximations and interpolations of sample functions in Euclidean spaces: Applications to sparse data, contour lines and inpainting, *SIAM J. Imaging Sci.*, **11** (2018), 2368–2428. https://doi.org/10.1137/17M116152X
12. C. B. Morrey Jr, *Multiple Integrals in the Calculus of Variations*, Springer-Verlag, New York, 1966.https://doi.org/10.1007/978-3-540-69952-1
13. B. Dacorogna, *Direct Methods in the Calculus of Variations*, 2^{nd} edition, Springer-Verlag, New York, 1989. https://doi.org/10.1007/978-0-387-55249-1
14. J. M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rational Mech. Anal.*, **63** (1977), 337–403. https://doi.org/10.1007/BF00279992

15. E. Acerbi, N. Fusco, Semicontinuity problems in the calculus of variations, *Arch. Rational Mech. Anal.*, **86** (1984), 125–145. https://doi.org/10.1007/BF00275731

16. J. M. Ball, R. D. James, Fine phase mixtures as minimizers of energy, *Arch. Rational Mech. Anal.*, **100** (1987), 13–52. https://doi.org/10.1007/BF00281246

17. J. M. Ball, R. D. James, Proposed experimental tests of a theory of fine microstructures and the two-well problem, *Philos. Trans. R. Soc. A*, **338** (1992), 389–450. https://doi.org/10.1098/rsta.1992.0013

18. V. Šverák, Quasiconvex functions with subquadratic growth, *Proc. Royal Soc. London Ser. A*, **433** (1991), 723–F725. https://doi.org/10.1098/rspa.1991.0073

19. K. Zhang, A construction of quasiconvex functions with linear growth at infinity, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, **19** (1992), 313–326.

20. R. V. Kohn, The relaxation of a double-well energy, *Cont. Mech. Therm.*, **3** (1991), 981–1000. https://doi.org/10.1007/BF01135336

21. N. B. Firoozye, Optimal use of the translation method and relaxations of variational problems, *Comm. Pure Appl. Math.*, **44** (1991), 643–678. https://doi.org/10.1002/cpa.3160440603

22. K. Bhattacharya, N. B. Firoozye, R. D. James, R.V. Kohn, Restrictions on microstructures, *Proc. Royal Soc. Edinb. - A*, **124** (1994), 843–878. https://doi.org/10.1017/S0308210500022381

23. M. Giaquinta, Introduction to regularity theory for nonlinear elliptic systems, *Lectures in Mathematics ETH Zürich*, Birkhüser Verlag, Basel, 1993. https://doi.org/10.1007/978-88-7642-443-4

24. A. A. Ahmadi, A. Olshevsky, P. Parrilo, J. N. Tsitsiklis, NP-hardness of deciding convexity of quartic polynomials and related problems, *Math. Program.*, **137** (2013), 453–476. https://doi.org/10.1007/s10107-011-0499-2

25. G. Dolzmann, Numerical computation of rank-one convex envelopes, *SIAM J. Numer. Anal.*, **36** (1999), 1621–1635. https://doi.org/10.1137/S0036142997325581

26. K. Zhang, E. Crooks, A. Orlando, Compensated convexity on bounded domains, mixed Moreau envelopes and computational methods, *Appl. Math. Model.*, **94** (2021), 688–720. https://doi.org/10.1016/j.apm.2021.01.040

27. K. Zhang, Mountain pass solutions for a double-well energy, *J. Differ. Equ.*, **182** (2002), 490–510. https://doi.org/10.1006/jdeq.2001.4113

28. K. Zhang, An elementary derivation of the generalized Kohn-Strang relaxation formulae, *J. Convex Anal.*, **9** (2002), 269–285.