

Occupation measure relaxations in variational problems: the role of convexity

Didier Henrion^{1,2}, Milan Korda^{1,2}, Martin Kružík^{3,4}, and Rodolfo Ríos-Zertuche⁵

¹LAAS-CNRS, Université de Toulouse, France

²Faculty of Electrical Engineering, Czech Technical University in Prague, Czechia

³Czech Academy of Sciences, Institute of Information Theory and Automation, Prague, Czechia

⁴Faculty of Civil Engineering, Czech Technical University in Prague, Czechia

⁵Department of Mathematics and Statistics, UiT The Arctic University of Norway, Tromsø, Norway

March 1, 2024

Abstract

This work addresses the occupation measure relaxation of calculus of variations problems, which is an infinite-dimensional linear programming reformulation amenable to numerical approximation by a hierarchy of semidefinite optimization problems. We address the problem of equivalence of this relaxation to the original problem. Our main result provides sufficient conditions for this equivalence. These conditions, revolving around the convexity of the data, are simple and apply in very general settings that may be of arbitrary dimensions and may include pointwise and integral constraints, thereby considerably strengthening the existing results. Our conditions are also extended to optimal control problems. In addition, we demonstrate how these results can be applied in non-convex settings, showing that the occupation measure relaxation is at least as strong as the convexification using the convex envelope; in doing so, we prove that a certain weakening of the occupation measure relaxation is equivalent to the convex envelope. This opens the way to application of the occupation measure relaxation in situations where the convex envelope relaxation is known to be equivalent to the original problem, which includes problems in magnetism and elasticity.

1 Introduction

The moment-sum-of-squares (SOS) hierarchy is a mathematical technology that consists of two steps: 1) formulating a nonlinear problem (of calculus of variations, optimization or control) as a linear convex optimization problem on the cone of nonnegative measures; 2) solving approximately the infinite-dimensional linear problem on measures by a hierarchy of finite-dimensional convex (typically semidefinite) optimization problems of increasing size. This hierarchy builds on the duality between the cone of moments and the cone of nonnegative polynomials, and its convergence relies on SOS representations of nonnegative polynomials – see [11] for a recent overview. For calculus of variations and optimal control problems, the linear problem on measures in step 1 is referred to as the *occupation measure relaxation*.

Whereas the convergence analysis of the hierarchy of finite-dimensional convex optimization problems to the solution of the infinite-dimensional linear problem on measures (in step 2) is now well understood, an essential remaining difficulty consists of ensuring that there is *no relaxation gap* when formulating (in step 1) the original nonlinear problem as a linear problem on measures. Indeed, if there is a relaxation gap, it is currently not understood in full generality how the approximate solutions obtained by the hierarchy are related to the solutions to the original problem, which is undesirable. For optimal control or regions of attraction of nonlinear ordinary differential equations, the absence of a relaxation gap was ensured under convexity

assumptions [26], see [17, Theorem 2.3] and [9, Assumption I], or also [10, Assumption 2]. For scalar hyperbolic conservation laws, the absence of a relaxation gap was ensured by introducing appropriate entropy inequalities [18]. See [15, Remark 6] for a discussion on the question of the relaxation gap in the general case of polynomial partial differential equations (PDEs).

Any attempt to apply the moment-SOS hierarchy to solve nonlinear calculus of variations (calculus of variations) problems [5] is invariably faced with this key relaxation gap issue. In [4], no relaxation gap was proved for three very specific calculus of variations problems with an integrand which is quadratic convex both in the function and its gradient. Similarly, there is no relaxation gap when the integrand is separably convex in the function and the gradient [8]. In [16] the absence of a relaxation gap was proved for calculus of variations problems with PDE constraints under convexity assumption in the gradient when the dimension of the codomain is 1, with the help of techniques from geometric measure theory. This reference also provides a counterexample with a positive relaxation gap when the codomain has dimension greater than one, despite the Lagrangian density being convex in the gradient.

The main result of our paper, Theorem 3.1, provides a simple proof of the absence of a relaxation gap for calculus of variations in a very general dimension-independent setting for problems with the Lagrangian and constraints that are jointly convex in the function and its gradient, thereby considerably strengthening the existing results. Concretely, on the one hand, the result eliminates the need for a separability of the Lagrangian in [8] and allows for a very broad class of constraints to be considered. On the other hand, the result eliminates the dimensionality dependence of [16], at the price of stronger convexity assumptions.

Similarly to [8], our proof technique works for a large class of measures that satisfy a Liouville condition for test functions that are affine in codomain variables. We analyze this weaker relaxation with affine rather than nonlinear test functions further and prove in Theorem 4.2 that when it is applied to a nonconvex problem its infimum is equal to the infimum of the original problem convexified by taking the classical convex envelope. Therefore, for nonconvex problems, the value of the occupation measure relaxation with nonlinear test functions is sandwiched between the infimum of the original problem and the infimum of the convexification using the convex envelope. Consequently, the occupation measure relaxation is at least as strong as the convexification using the convex envelope. In particular, whenever one can prove the absence of a relaxation gap for the latter, it holds for the former.

The convexity assumptions of our result hold for some problems in continuum physics, namely for static micromagnetics [6], which we explore in detail in Section 6, and for some models of elasticity [14, 23].

This opens the way to application of the occupation measure relaxation in situations where the convex envelope relaxation is known to be equivalent to the original problem. For these problems, the use of Young measures (also called parametrized measures) is classical, see e.g. [27, 24, 20]. As already explained in [12], occupation measures can be seen as an extension of Young measures allowing the application of numerical methods based on the moment-SOS hierarchy.

The paper is organized as follows. In Section 2 we state the calculus of variations problem to be solved, and we introduce the occupation measures we are using to reformulate it as a linear optimization problem. Section 3 contains our main result, Theorem 3.1, about the absence of a relaxation gap between the calculus of variations problem and its linear formulation, under convexity assumptions. Non-convex problems are the subject of Section 4 where we show how our linear formulation relates to the classical convex envelope relaxation. Section 5 explains how our calculus of variations results can be extended to optimal control problems. Finally, calculus of variations applications in micromagnetics are described in Section 6.

Notations. We will denote the n -dimensional Euclidean space by \mathbb{R}^n , its standard inner product by $\langle \cdot, \cdot \rangle$ and its Euclidean norm by $\| \cdot \|$. Let us also denote the transpose of a matrix A by A^\top .

For a set $\Omega \subset \mathbb{R}^n$, we denote its topological closure by $\bar{\Omega}$ and its boundary by $\partial\Omega$. We will denote by $C(\Omega)$ the set of all continuous functions on Ω and by $C^k(\Omega)$ the set of all k -times continuously differentiable functions on Ω . The set $C^\infty(\Omega)$ denotes the set of infinitely-

differentiable functions on Ω . Similarly $C_c^\infty(\Omega)$ and $C_c^k(\Omega)$ denote the subsets of $C^\infty(\Omega)$ and $C^k(\Omega)$, respectively, consisting of compactly-supported functions. Additionally, let $C_0^\infty(\Omega)$ and $C_0^k(\Omega)$ denote the subsets of $C^\infty(\Omega)$ and $C^k(\Omega)$, respectively, consisting of functions vanishing on the boundary $\partial\Omega$. Unless otherwise specified, the codomain of all functions involved is \mathbb{R} . If we want to specify that the codomain is a set $Y \subset \mathbb{R}^m$, we write, e.g., $C(\Omega; Y)$. The set $W^{k,p}(\Omega; Y)$ denotes the Sobolev space of measurable functions from $\Omega \subset \mathbb{R}^n$ to $Y \subset \mathbb{R}^m$ that are weakly differentiable k times with each of those derivatives in $L^p(\Omega)$. Given a differentiable function $f: \mathbb{R}^n \rightarrow \mathbb{R}$, ∇f denotes the vector of partial derivatives $[\partial f/\partial x_1, \dots, \partial f/\partial x_n]^\top$. A set-valued map ϕ associating to each point in a set X a subset of a set Y will be denoted $\phi: X \rightrightarrows Y$.

2 Calculus of variations and linear formulations with occupation measures

Let $\Omega \subseteq \mathbb{R}^n$ be a connected, bounded domain with piecewise C^1 boundary $\partial\Omega$, let $Y \subset \mathbb{R}^m$ and $Z \subset \mathbb{R}^{m \times n}$. Denote by σ the $(n-1)$ -dimensional Hausdorff measure on the boundary $\partial\Omega$.

Consider the calculus of variations problem

$$M := \inf_{y \in W^{1,p}(\Omega; Y)} \int_{\Omega} L(x, y(x), Dy(x)) dx \quad (1)$$

$$\text{subject to } A_i(x, y(x), Dy(x)) = 0, \quad B_i(x, y(x), Dy(x)) \leq 0 \quad \text{a.e. } x \in \Omega, \quad i \in I, \quad (2)$$

$$A_{\partial,i}(x, y(x)) = 0, \quad B_{\partial,i}(x, y(x)) \leq 0 \quad \sigma\text{-a.e. } x \in \partial\Omega, \quad i \in I, \quad (3)$$

$$\int_{\Omega} C_i(x, y(x), Dy(x)) dx \leq 0, \quad \int_{\partial\Omega} C_{\partial,i}(x, y(x)) d\sigma(x) \leq 0, \quad i \in I, \quad (4)$$

where I is a possibly uncountable index set and all functions involved are assumed to be measurable. Considering constraints indexed by a possibly uncountable index set provides a significant modeling freedom, allowing one, for example, to model higher-order PDE constraints in the weak form as shown in Example 2.3.

Example 2.1 (Plateau's problem). Let Ω be the unit disc in \mathbb{R}^2 , $Y = \mathbb{R}^3$, and $Z = \mathbb{R}^{2 \times 3}$. Let $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^3$ be a continuous, non-self-intersecting loop, $\gamma(0) = \gamma(2\pi)$. Let $L(x, y, Dy) = \left\| \frac{\partial y}{\partial x_1} \times \frac{\partial y}{\partial x_2} \right\|$, that is, L is the area of the parallelogram spanned by the columns of Dy , known as the element of surface area. Let $A_{\partial,1}(x, y)$ be the distance from y to the image $\gamma([0, 2\pi])$. Let all other constraint functions $A_i, B_i, C_i, A_{\partial,i}, B_{\partial,i}, C_{\partial,i}$ be zero. In this setting, M is the area of the minimal surface with boundary $\gamma([0, 2\pi])$.

Example 2.2 (Geodesics of a differential inclusion). Let Ω be the interval $[0, 1]$ and let V be a connected, open subset of $Y = \mathbb{R}^n$, and $Z = \mathbb{R}^n$. Fix two points y_0 and y_1 in V , as well as a Riemannian metric $g: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$. Let $B_1: Y \times Z \rightarrow \mathbb{R}$ be a continuous function such that the sets $X(y) = \{z \in Z : B(y, z) \leq 0\}$ are nonempty for each $y \in Y$. Let $L(x, y, z) = \sqrt{g(z, z)}$, let $A_1(x, y, z)$ be the distance from y to V , and $A_{\partial,1}(x, y, z)$ be the distance from y to the set $\{y_0, y_1\}$. Then M is the minimal g -distance from y_0 to y_1 for curves joining these points and respecting the constraint $\gamma(t) \in V$ and the differential inclusion $\gamma'(t) \in X(\gamma(t))$, and the minimizers are the geodesics joining these two points and respecting the differential inclusion. When $X(y)$ is a linear subspace of Z for each y , this setting is known as sub-Riemannian geometry.

Example 2.3 (Higher-order PDEs). Considering an uncountable index set in (1) permits higher-order PDEs to be considered as constraints in (1) through their weak reformulation. We demonstrate this on the Laplace equation. In the $m = 1$ case, to model a Laplace equation in the weak form, we can use the constraint

$$\int_{\Omega} \nabla \phi(x)^\top \nabla y(x) dx = 0, \quad \phi \in C_0^1(\Omega).$$

Observe that if y were a smooth function, integrating by parts we would obtain $\int \phi(x) \Delta y(x) dx = 0$ for all $\phi \in C_0^1(\Omega)$, whence the above conditions indeed give a weak analogue of the Laplace equation $\Delta y = 0$. In order to model it as in (4), we set I to be $C_0^1(\Omega)$, and for each $\phi \in I = C_0^1(\Omega)$ we set $C_\phi(x, y, z) = \nabla \phi(x)^\top z$. Observe that the equality constraint results from the inequalities for ϕ and $-\phi$, which are both contained in $C_0^1(\Omega)$.

To reformulate calculus of variations problem (1) as a linear optimization problem, we now introduce two classes of measures. The first ones, the *affinely relaxed occupation measures*, form a larger class that contains the second ones, the *relaxed occupation measures*. Our main no-gap results will hold for the first class of measures (see Theorems 3.1 and 5.2), and hence will also hold for the second one (see Corollary 3.2 and Remark 5.1). The second class is more commonly found in the literature, and we chose to introduce the first class in order to present our results in greater generality.

Definition 2.4 (Affinely relaxed occupation measures). Let $p \in [1, +\infty]$. Let $\mathcal{O}_p^{\text{aff}}$ be the set of pairs (μ, μ_∂) consisting of positive Borel measures on $\bar{\Omega} \times Y \times Z$ respectively $\partial\Omega \times Y$ satisfying:

- The total mass of μ equals the Lebesgue measure of Ω , that is,

$$\mu(\bar{\Omega} \times Y \times Z) = |\Omega|. \quad (5)$$

- The sets $\text{supp } \mu$ and $\text{supp } \mu_\partial$ are compact if $p = +\infty$, or

$$\int_{\Omega \times Y \times Z} \|y\|^p + \sum_i \|z_i\|^p d\mu(x, y, z) < +\infty \quad \text{and} \quad \int_{\partial\Omega \times Y} \|y\|^p d\mu_\partial(x, y) < +\infty \quad (6)$$

if $p \in [1, +\infty)$. Here, $z = (z_1, \dots, z_n)$, $z_i \in \mathbb{R}^m$.

- For all $\phi_1 \in C^\infty(\Omega; \mathbb{R})$ and all $\phi_2 \in C^\infty(\Omega; \mathbb{R}^m)$ it holds

$$\begin{aligned} \int_{\Omega \times Y \times Z} \left[\frac{\partial \phi_1}{\partial x}(x) + y^\top \frac{\partial \phi_2}{\partial x}(x) + \phi_2(x)^\top z \right] d\mu(x, y, z) \\ = \int_{\partial\Omega \times Y} (\phi_1(x) + y^\top \phi_2(x)) \mathbf{n}(x) d\mu_\partial(x, y). \end{aligned} \quad (7)$$

Here \mathbf{n} denotes the exterior unit (row) vector normal to the boundary $\partial\Omega$.

Remark 2.5. Observe that (7) amounts to the Liouville equation

$$\int_{\Omega \times Y \times Z} \frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial y} z d\mu = \int_{\partial\Omega \times Y} \phi \mathbf{n} d\mu_\partial$$

for a test function $\phi \in C^\infty(\Omega \times Y; \mathbb{R})$ that is affine in y , i.e., of the form $\phi(x, y) = \phi_1(x) + y^\top \phi_2(x)$.

Remark 2.6. The assumption of Ω being connected can be dropped at the price of replacing (5) by the slightly stronger requirement that its x -marginal is equal to the Lebesgue measure on Ω . When Ω is connected and its mass is equal to the volume of Ω (condition (5)), its x -marginal is automatically equal to the Lebesgue measure on Ω due to (7) and Lemma A.1 in the appendix.

Definition 2.7 (Relaxed occupation measures). Let $p \in [1, +\infty]$. Let \mathcal{O}_p be the set of pairs $(\mu, \mu_\partial) \in \mathcal{O}_p^{\text{aff}}$ that additionally satisfy:

- For all $\phi \in \mathbf{F}_p$, it holds

$$\int_{\Omega \times Y \times Z} \frac{\partial \phi}{\partial x}(x, y) + \frac{\partial \phi}{\partial y}(x, y) z d\mu(x, y, z) = \int_{\partial\Omega \times Y} \phi(x, y) \mathbf{n}(x) d\mu_\partial(x, y), \quad (8)$$

where \mathbf{n} denotes the exterior unit vector normal to the boundary $\partial\Omega$,

$$\begin{aligned} \mathbf{F}_p = \left\{ \phi \in C^\infty(\Omega \times Y) : \exists c > 0 \text{ such that } \left\| \frac{\partial \phi}{\partial x}(x, y) \right\| \leq c(1 + \|y\|^p), \right. \\ \left. \left\| \frac{\partial \phi}{\partial y}(x, y) \right\| \leq c(1 + \|y\|^{p-1}), |\phi(x, y)| \leq c(1 + \|y\|^p) \right\}, \quad p \in [1, \infty) \end{aligned}$$

and $\mathbf{F}_\infty = C^\infty(\Omega \times Y)$. By Lemma B.1, the integrals in (8) are well defined.

Consider the relaxations with affinely relaxed occupation measures (Definition 2.4),

$$\begin{aligned}
M_r^{\text{aff}} := & \inf_{(\mu, \mu_\partial) \in \mathcal{O}_p^{\text{aff}}} \int_{\Omega \times Y \times Z} L(x, y, z) d\mu(x, y, z) & (9) \\
\text{subject to} & \text{supp } \mu \subseteq \{(x, y, z) \in \Omega \times Y \times Z \mid A_i(x, y, z) = 0, B_i(x, y, z) \leq 0, i \in I\}, \\
& \text{supp } \mu_\partial \subseteq \{(x, y) \in \partial\Omega \times Y \mid A_{\partial, i}(x, y) = 0, B_{\partial, i}(x, y) \leq 0, i \in I\}, \\
& \int_{\Omega \times Y \times Z} C_i(x, y, z) d\mu(x, y, z) \leq 0, \quad i \in I, \\
& \int_{\partial\Omega \times Y} C_{\partial, i}(x, y) d\mu_\partial(x, y) \leq 0, \quad i \in I,
\end{aligned}$$

and with relaxed occupation measures (Definition 2.7),

$$\begin{aligned}
M_r := & \inf_{(\mu, \mu_\partial) \in \mathcal{O}_p} \int_{\Omega \times Y \times Z} L(x, y, z) d\mu(x, y, z) & (10) \\
\text{subject to} & \text{ [same constraints as in (9)].}
\end{aligned}$$

Remark that the sole difference between (9) and (10) is the set over which the infimum is taken, namely, $\mathcal{O}_p^{\text{aff}}$ and \mathcal{O}_p , respectively.

Observe also that problems (9) and (10) are *linear* optimization problems in the unknown occupation measures (μ, μ_∂) , whereas the original calculus of variations problem (1) can be *nonlinear* in the unknown function y .

3 Convex case – main result

Our main result states that under convexity assumption there is no relaxation gap between the original calculus of variations problem and its linear reformulation with affinely relaxed occupation measure.

Theorem 3.1. *Suppose that for every $i \in I$ and $x \in \Omega$ the following holds: L , B_i and C_i are convex in (y, z) , and A_i is affine in (y, z) . Suppose also that for every $i \in I$ and $x \in \partial\Omega$ the following holds: $A_{\partial, i}$ is affine in y , and $B_{\partial, i}$ and $C_{\partial, i}$ are convex in y . Then, if Y and Z are closed and convex, we have*

$$M_r^{\text{aff}} = M.$$

Recall that these are defined in (1) and (9).

Proof. Clearly we have $M_r^{\text{aff}} \leq M$ because every function y induces measures $(\mu, \mu_\partial) \in \mathcal{O}_p^{\text{aff}}$ by letting

$$\begin{aligned}
\int_{\Omega \times Y \times Z} \phi(x, y, z) d\mu(x, y, z) &= \int_{\Omega} \phi(x, y(x), Dy(x)) dx, \\
\int_{\partial\Omega \times Y} \phi(x, y) d\mu_\partial(x, y) &= \int_{\partial\Omega} \phi(x, y(x)) d\sigma(x),
\end{aligned}$$

for all bounded measurable functions ϕ . Equivalently, μ is the Borel measure that, for every measurable set $A \subset \Omega \times Y \times Z$, verifies $\mu(A) = \int_{\Omega} \chi_A(x, y(x), Dy(x)) dx$, where χ_A is the indicator function of A that equals 1 on A and vanishes elsewhere; and analogously for μ_∂ . We obviously have $\mu(\bar{\Omega} \times Y \times Z) = |\Omega|$, which is (5); let us verify also (6) and (7). Since every function y viable in (1) is in $W^{1, p}$, we have that y is weakly differentiable and both y and Dy are in L^p , thus

$$\begin{aligned}
\int_{\Omega \times Y \times Z} \|y\|^p + \sum_i \|z_i\|^p d\mu(x, y, z) &= \int_{\Omega} \left(\|y(x)\|^p + \sum_i \left\| \frac{\partial y}{\partial x_i}(x) \right\|^p \right) dx \\
&= \|y\|_p^p + \|Dy\|_p^p < +\infty,
\end{aligned}$$

which shows that (6) holds. Let $\phi_1 \in C^\infty(\Omega; \mathbb{R})$ and $\phi_2 \in C^\infty(\Omega; \mathbb{R}^m)$; then by the Stokes theorem we have, letting $\phi(x, y) = \phi_1(x) + y^\top \phi_2(x)$ (cf. Remark 2.5),

$$\begin{aligned}
& \int_{\Omega \times Y \times Z} \left[\frac{\partial \phi_1}{\partial x}(x) + y^\top \frac{\partial \phi_2}{\partial x}(x) + \phi_2(x)^\top z \right] d\mu(x, y, z) \\
&= \int_{\Omega} \left[\frac{\partial \phi_1}{\partial x}(x) + y(x)^\top \frac{\partial \phi_2}{\partial x}(x) + \phi_2(x)^\top Dy(x) \right] dx \\
&= \int_{\Omega} \frac{\partial \phi}{\partial x}(x, y(x)) + \frac{\partial \phi}{\partial y}(x, y(x)) Dy(x) dx \\
&= \int_{\partial \Omega} \phi(x, y(x)) \mathbf{n}(x) d\sigma(x) \\
&= \int_{\partial \Omega \times Y} \phi(x, y) \mathbf{n}(x) d\mu_{\partial}(x, y) \\
&= \int_{\partial \Omega \times Y} (\phi_1(x) + y^\top \phi_2(x)) \mathbf{n}(x) d\mu_{\partial}(x, y).
\end{aligned}$$

This shows that (7) holds as well.

The harder part is to prove that $M_{\Gamma}^{\text{aff}} \geq M$. To this end, let (μ, μ_{∂}) be feasible in (9). By Lemma A.1, the projection of μ onto Ω is Lebesgue measure dx . Disintegrate μ such that

$$\mu(x, y, z) = \nu(y, z | x) dx, \quad (11)$$

where ν is the conditional distribution of (y, z) given x ; in particular, for every fixed x , $\nu(\cdot, \cdot | x)$ is a probability measure. Define

$$y(x) := \int_{Y \times Z} y d\nu(y, z | x), \quad z(x) := \int_{Y \times Z} z d\nu(y, z | x) \quad (12)$$

to be the conditional expectation (or centroid) of μ in the y and z variables given x . We note that $y(x) \in \mathbb{R}^m$ and $z(x) \in \mathbb{R}^{m \times n}$. By Jensen's inequality we have

$$\int_{\Omega \times Y \times Z} L d\mu = \int_{\Omega} \int_{Y \times Z} L(x, y, z) d\nu(y, z | x) dx \geq \int_{\Omega} L(x, y(x), z(x)) dx.$$

Now we shall prove that $y(\cdot)$ is weakly differentiable and $Dy(x) = z(x)$. To prove weak differentiability of $y(\cdot)$, take $\phi_1 = 0$ and $\phi_2 = \psi \in C_0^\infty(\Omega; \mathbb{R}^m)$, and use it as a test function in (7). This gives

$$\int_{\Omega \times Y \times Z} y^\top \frac{\partial \psi}{\partial x} + \psi(x)^\top z d\mu(x, y, z) = 0.$$

Using the disintegration of μ we obtain

$$\int_{\Omega} \int_{Y \times Z} y^\top \frac{\partial \psi}{\partial x} + \psi(x)^\top z d\nu(y, z | x) dx = 0.$$

Since the integrand is linear in (y, z) , we obtain, using the definitions of $y(\cdot)$ and $z(\cdot)$,

$$\int_{\Omega} y(x)^\top \frac{\partial \psi}{\partial x} + \psi(x)^\top z(x) dx = 0, \quad \forall \psi \in C_0^\infty(\Omega).$$

This is nothing but the definition of weak differentiability of $y(\cdot)$ with weak derivative $z(\cdot)$. If $p = +\infty$, $\text{supp } \mu$ is bounded, and then $z(\cdot)$ is L^∞ , so it follows that $y \in W^{1, \infty}$. If $p \in [1, +\infty)$, then $z(\cdot)$ is in L^p ; indeed, by Jensen's inequality applied to the convex function $\|\cdot\|^p$, we have, writing $z = (z_1, \dots, z_m)$,

$$\begin{aligned}
& \int_{\Omega \times Y \times Z} \|z_i(x)\|^p dx = \int_{\Omega} \left\| \int_{Y \times Z} z_i d\nu(y, z | x) \right\|^p dx \\
& \leq \int_{\Omega} \int_{Y \times Z} \|z_i\|^p d\nu(y, z | x) dx = \int_{\Omega \times Y \times Z} \|z_i\|^p d\mu(x, y, z) < +\infty,
\end{aligned}$$

by (6). A similar calculation shows that $y(\cdot)$ is in L^p as well. It follows that y is in $W^{1,p}$.

We also need to verify the constraints (2)–(4). This follows from the support constraints on μ and μ_∂ and the fact that A_i and $A_{\partial,i}$ are affine and B_i , $B_{\partial,i}$, C_i and $C_{\partial,i}$ are convex in (y, z) . Indeed, the support constraints imply

$$\int_{\Omega \times Y \times Z} \psi(x) A_i(x, y, z) d\mu(x, y, z) = 0 \quad \forall \psi \in C(\Omega).$$

Using the disintegration (11), and definition of $y(\cdot)$, $z(\cdot)$ and the fact that A_i is affine, we obtain

$$\int_{\Omega} \psi(x) A_i(x, y(x), z(x)) dx = 0 \quad \forall \psi \in C(\Omega)$$

and hence the constraint $A_i(x, y(x), z(x)) = 0$ is satisfied a.e. in Ω , $i \in I$. For the inequality constraint $B_i(x, y(x), Dy(x)) \leq 0$, we observe that for any nonnegative $\psi \in C(\Omega)$, we have

$$\begin{aligned} 0 &\geq \int_{\Omega \times Y \times Z} \psi(x) B_i(x, y, z) d\mu(x, y, z) \\ &= \int_{\Omega} \int_{Y \times Z} \psi(x) B_i(x, y, z) d\nu(y, z | x) dx \geq \int_{\Omega} \psi(x) B_i(x, y(x), z(x)) dx, \end{aligned}$$

where the first inequality follows from the support constraints on μ and the last one from Jensen's inequality. Since $z(x) = Dy(x)$, this implies that $B_i(x, y(x), Dy(x)) \leq 0$ a.e. in Ω . Another application of Jensen's inequality gives, for all $i \in I$,

$$\int_{\Omega} C_i(x, y(x), z(x)) dx \leq \int_{\Omega} \int_{Y \times Z} C_i(x, y, z) d\nu(y, z | x) dx = \int_{\Omega \times Y \times Z} C_i(x, y, z) d\mu(x, y, z) \leq 0.$$

Let us now verify the boundary condition. To this end, disintegrate μ_∂ as

$$\mu_\partial = \nu_\partial(y | x) d\sigma(x) \tag{13}$$

with σ being the Hausdorff measure on $\partial\Omega$. Define

$$y_\partial(x) = \int_Y y d\nu_\partial(y | x).$$

We shall show that $y_\partial(\cdot)$ is a Sobolev trace of $y(\cdot)$ (see [7, Definition, 4.5]). To this end, take $\phi_1 = 0$ and $\phi_2 = \psi \in C^\infty(\Omega; \mathbb{R}^m)$, so that plugging in (7) and using the fact that $z(x) = Dy(x)$ almost everywhere in Ω gives

$$\int_{\Omega} y(x)^\top \frac{\partial \psi}{\partial x} + \psi(x) Dy(x) dx = \int_{\partial\Omega} \psi(x)^\top y_\partial(x) \mathbf{n}(x) d\sigma(x), \quad \forall \psi \in C^\infty(\mathbb{R}^n).$$

This is the Stokes theorem for the Sobolev functions which shows that $y_\partial(\cdot)$ is indeed a Sobolev trace of $y(\cdot)$ [7, Theorem 4.6]. Hence, in order to show that $A_{\partial,i}(x, y(x)) = 0$ on $\partial\Omega$ it suffices to prove it for y_∂ . The second support constraint of (9) implies that

$$\int_{\partial\Omega \times Y} \phi(x) A_{\partial,i}(x, y) d\mu_\partial(x, y) = 0 \quad \forall \phi \in C(\partial\Omega).$$

Using the disintegration (13), the definition of y_∂ and the fact that A_∂ is affine in y , it follows that

$$\int_{\partial\Omega \times Y} \phi(x) A_{\partial,i}(x, y_\partial(x)) dx = 0 \quad \forall \phi \in C(\partial\Omega).$$

This implies that $y_\partial(\cdot)$ and hence $y(\cdot)$ satisfy the boundary condition σ a.e. in $\partial\Omega$. For $p = +\infty$, $y(\cdot)$ is Lipschitz and hence $y(\cdot)$ satisfies it in fact everywhere in $\partial\Omega$. The boundary inequality constraint $B_{\partial,i}(x, y(x)) \leq 0$ follows by Jensen's inequality the same way as for $B_i(x, y(x), Dy(x)) \leq 0$, and similarly for the integral boundary constraint (4).

It remains to verify that $y(x) \in Y$ for all $x \in \Omega$. This, however, follows by the fact that Y is closed and convex and the last support constraint of (9). \square

Corollary 3.2. *Under the assumptions of Theorem 3.1, we have $M_r = M_r^{\text{aff}} = M$.*

Proof. Since $\mathcal{O}_p \subseteq \mathcal{O}_p^{\text{aff}}$ and \mathcal{O}_p contains the pairs of measures induced by functions in $W^{1,p}$, we have that $M_r^{\text{aff}} \leq M_r \leq M$. Moreover, by Theorem 3.1 we know that $M = M_r^{\text{aff}}$, so M_r must be equal to them as well. \square

4 Nonconvex case

We now turn to the case when the problem is not convex and show that the affinely relaxed occupation measures provide a lower bound at least as good as the the convex envelope relaxation and, in the absence of integral constraints, the two lower bounds coincide.

4.1 Convex underestimators

First, we make an observation comparing the value of the occupation measure relaxation to the value of the original problem with data replaced by its convex underestimators. Concretely, suppose that there exist convex sets $\hat{Y} \supset Y$ and $\hat{Z} \supset Z$, and functions $\hat{L}, \hat{B}_i, \hat{C}_i: \Omega \times \hat{Y} \times \hat{Z} \rightarrow \mathbb{R}$ such that for all $(x, y, z) \in \Omega \times Y \times Z$ satisfying $A_i(x, y, z) = 0$, $B_i(x, y, z) \leq 0$, it holds $\hat{L}(x, y, z) \leq L(x, y, z)$, $\hat{B}_i(x, y, z) \leq B_i(x, y, z)$, $\hat{C}_i(x, y, z) \leq C_i(x, y, z)$. Suppose also that there exist functions $\hat{L}_\partial, \hat{B}_{\partial,i}, \hat{C}_{\partial,i}: \Omega \times \hat{Y} \rightarrow \mathbb{R}$ such that for all $(x, y) \in \partial\Omega \times Y$ satisfying $A_{\partial,i}(x, y) = 0$, $B_{\partial,i}(x, y) \leq 0$, it holds $\hat{L}_\partial(x, y) \leq L_\partial(x, y)$, $\hat{B}_{\partial,i}(x, y) \leq B_{\partial,i}(x, y)$ and $\hat{C}_{\partial,i}(x, y) \leq C_{\partial,i}(x, y)$. Consider the problem

$$\begin{aligned} M_{\text{under}} = \inf_{y \in W^{1,p}(\Omega, \hat{Y})} & \int_{\Omega} \hat{L}(x, y(x), Dy(x)) dx + \int_{\partial\Omega} \hat{L}_\partial(x, y(x)) d\sigma(x) & (14) \\ \text{subject to} & A_i(x, y(x), Dy(x)) = 0, \quad \hat{B}_i(x, y(x), Dy(x)) \leq 0, \quad x \in \Omega, \quad i \in I, \\ & A_{\partial,i}(x, y(x)) = 0, \quad \hat{B}_{\partial,i}(x, y(x)) \leq 0, \quad x \in \partial\Omega, \quad i \in I, \\ & \int_{\Omega} \hat{C}_i(x, y(x), Dy(x)) dx \leq 0, \quad \int_{\partial\Omega} \hat{C}_{\partial,i}(x, y(x)) dx \leq 0, \quad i \in I. \end{aligned}$$

We have the following simple corollary, where we for simplicity assume that A_i and $A_{\partial,i}$ are affine, i.e., they do not need further convexification. This could be relaxed by replacing A_i and $A_{\partial,i}$ by A_i^2 and $A_{\partial,i}^2$ and considering convex underestimators thereof.

Corollary 4.1. *Suppose that A_i is affine in (y, z) for every $x \in \Omega$ and $A_{\partial,i}$ is affine in y for every $x \in \partial\Omega$. It holds $M_{\text{under}} \leq M_r^{\text{aff}} \leq M_r \leq M$ (defined, respectively, in (14), (9), (10), and (1)). In particular, if $M_{\text{under}} = M$, then $M_r^{\text{aff}} = M_r = M$.*

Proof. Let \hat{M}_r denote the infimum of the occupation measure relaxation applied to the convexified problem (14). Applying Corollary 3.2, we have $\hat{M}_r = M_{\text{under}}$. Given a pair of measures (μ, μ_∂) feasible in (9), we observe that these two measures are feasible in the occupation measure relaxation of the convexified problem since the feasible set of the latter is larger than the feasible set of the original problem. Since $\hat{L} \leq L$, it follows that $\hat{M}_r \leq M_r^{\text{aff}}$ and as a result $M_{\text{under}} \leq M_r^{\text{aff}} \leq M_r \leq M$ since $M_r^{\text{aff}} \leq M_r \leq M$ is always satisfied. \square

Note that when $\hat{M} = M = M_r^{\text{aff}} = M_r$, one has two choices for a relaxation to be solved numerically: either the convexification (14) or the occupation measure relaxation (9). Both are convex optimization problems and it depends on particular situation whether one or the other is more advantageous to solve. For example, the occupation measure relaxation may be preferable when no explicit description of the nonnegative convex underestimators used in (14) is available.

4.2 Equivalence with convex envelope

Now we significantly refine the previous observation and prove that in fact M_r^{aff} is equal to the infimum of a convexification of the problem based on the convex envelope of the Lagrangian and

the constraints. As a result, under the assumptions introduced below, we obtain the following inequalities

$$\text{Convex envelope} = \text{Affinely relaxed OM} \leq \text{Relaxed OM} \leq \text{Original},$$

where OM stands for ‘‘occupation measures.’’ To set the stage for proving this, we let $\text{conv } X$ denote the closed convex hull of a set $X \subset \mathbb{R}^n$. Let $p = +\infty$, and let Y and Z be compact sets. For $x \in \Omega$, let $K(x)$ denote the set of points $(y, z) \in Y \times Z$ such that the constraints $A_i(x, y, z) = 0$ and $B_i(x, y, z) \leq 0$ are satisfied, that is,

$$K(x) = \{(y, z) \in Y \times Z : A_i(x, y, z) = 0, B_i(x, y, z) \leq 0, i \in I\}.$$

We set $C_i = 0$, $C_{i,\partial} = 0$, i.e., we work without integral constraints in this section, except for Examples 4.4 and 4.5, that will show that we would not be able to obtain the same results in the presence of such constraints.

Let $\hat{L}: \Omega \times \text{conv } Y \times \text{conv } Z \rightarrow \mathbb{R} \cup \{+\infty\}$ be the function whose restriction to $\{x\} \times \text{conv } K(x)$ is the convex envelope of $L|_{\{x\} \times K(x)}$; in other words, for $x \in \Omega$,

$$\begin{aligned} \hat{L}(x, y, z) = \sup\{\varphi(y, z) \mid \varphi: \text{conv } Y \times \text{conv } Z \rightarrow \mathbb{R} \text{ convex,} \\ \varphi(y', z') \leq L(x, y', z') \forall (y', z') \in K(x)\}. \end{aligned}$$

Analogously, let $K_\partial(x) \subseteq Y$ be, for each $x \in \partial\Omega$, the set of points $y \in Y$ such that $A_{\partial,i}(x, y) = 0$ and $B_{\partial,i}(x, y) \leq 0$. Let $\hat{A}_i(x) = \text{conv}(A_i|_{\{x\} \times Y \times Z})^{-1}(0)$ and $\hat{A}_i = \bigcup_{x \in \Omega} \hat{A}_i(x)$; in other words, \hat{A}_i is the convexification in the $Y \times Z$ directions of the set $A_i^{-1}(0)$. Similarly, let $\hat{A}_{\partial,i}(x) = \text{conv}(A_{\partial,i}|_{\{x\} \times Y})^{-1}(0)$ and $\hat{A}_{\partial,i} = \bigcup_{x \in \partial\Omega} \hat{A}_{\partial,i}(x)$.

Let

$$\hat{M} = \inf_{y \in W^{1,\infty}(\Omega, \text{conv } Y)} \int_{\Omega} \hat{L}(x, y(x), Dy(x)) dx \quad (15)$$

$$\begin{aligned} \text{subject to} \quad & (y(x), Dy(x)) \in \text{conv } K(x), \quad \text{a.e. } x \in \Omega, \quad (16) \\ & y(x) \in \text{conv } K_\partial(x), \quad \text{a.e. } x \in \partial\Omega. \end{aligned}$$

The following result shows that the convexified problem (15) has the same infimum as the occupation measure relaxation (9) with affine test functions applied to the original problem with nonconvex data.

Theorem 4.2. *Assume that $K(x)$ is closed for each $x \in \Omega$ and that the set-valued map K is continuous in the topology induced by the Hausdorff metric. If L is continuous in (y, z) for every $x \in \Omega$, then it holds that*

$$\hat{M} = M_r^{\text{aff}}.$$

Proof. Let us first show that $M_r^{\text{aff}} \leq \hat{M}$. Let $y \in W^{1,p}(\Omega, \text{conv } Y)$ be viable for the optimization problem \hat{M} . Form the set

$$\tilde{K}(x) = \{(y, z, L(x, y, z)) \in Y \times Z \times \mathbb{R} : (y, z) \in K(x)\}.$$

Then the convex hull $\text{conv } \tilde{K}(x)$ contains the point $(y(x), Dy(x), \hat{L}(x, y, Dy(x)))$ for almost every $x \in \Omega$, because of (16) and by the definition of \hat{L} . In these context, Choquet’s Theorem [21, p. 14] can be applied because the set $\tilde{K}(x)$ is a metric space (with the metric induced by the norm, $\text{dist}(x, y) = \|x - y\|$) and, being a closed subset of a compact set,

$$\tilde{K}(x) \subseteq Y \times Z \times \left[\min_{(y,z) \in K(x)} L(x, y, z), \max_{(y,z) \in K(x)} L(x, y, z) \right],$$

$\tilde{K}(x)$ is compact, and it is a subset of \mathbb{R}^{m+mn+1} that is a locally convex space by virtue of being normed. By Choquet’s Theorem, we know that, for each $x \in \Omega$ and each $p \in \text{conv } \tilde{K}(x)$, there is a probability measure $\bar{\nu}_x$ in the set of extreme points of $\tilde{K}(x)$ such that, for every affine function $f: \mathbb{R} \times Y \times Z \rightarrow \mathbb{R}$,

$$f(p) = \int_{\tilde{K}(x)} f(y, z, w) d\bar{\nu}_x(y, z, w).$$

Applying this to $p = (y(x), Dy(x), \hat{L}(x, y, Dy(x)))$ and to each of the integrand $f(y, z, w) = w$, we get

$$\hat{L}(x, y(x), Dy(x)) = f(p) = \int_{\tilde{K}(x)} w d\bar{\nu}_x(w, y, z).$$

Observe that the set of extreme points of $\text{conv } \tilde{K}(x)$ is contained in $\tilde{K}(x)$, so the latter integral gives

$$\hat{L}(x, y(x), Dy(x)) = \int_{\tilde{K}(x)} w d\bar{\nu}_x(y, z, w) = \int_{K(x)} L(x, y, z) d\nu_x(y, z), \quad (17)$$

where $\nu_x = \pi_* \bar{\nu}_x$ is the pushforward by the projection $\pi(y, z, w) = (y, z)$.

Let us show that the mapping $x \mapsto \nu_x$ can be chosen so that it is measurable. Let \mathcal{M}_x be the set of Borel probability measures ν on $K(x)$ such that $\int_{K(x)} (y, z) d\nu(y, z) = (y(x), Dy(x))$. Endow \mathcal{M}_x with the weak* topology with respect to $C(Y \times Z; \mathbb{R})$. Let

$$I_L(\nu, x) = \int_{\{x\} \times Y \times Z} L d\nu \quad \text{and} \quad \phi(x) = \min_{\nu \in \mathcal{M}_x} I_L(\nu, x) = \hat{L}(x, y(x), Dy(x)).$$

With these definitions, I_L is continuous and ϕ is measurable. By the Prokhorov theorem [22], \mathcal{M}_x is compact and metrizable; observe that this application of the Prokhorov theorem requires us to prove that \mathcal{M}_x is tight, which is obvious because $K(x)$ is compact. The inverse map $\Phi_x = I_L(\cdot, x)^{-1}$, that is, $\Phi_x(r) = \{\nu \in \mathcal{M}_x : I_L(\nu, x) = r\}$, $r \in \mathbb{R}$, is weakly measurable as well, since every open set $U \subset \mathcal{M}_x$ can be written as a countable union of compact sets $U = \bigcup_i K_i$, and $\Phi_x^{-1}(U) = \bigcup_i \Phi_x^{-1}(K_i) = \bigcup_i I_L(K_i, x)$ is a countable union of continuous images of compact sets, each of which is compact, so the union is measurable. Then the set-valued map

$$H(x) = I_L(\cdot, x)^{-1}(\phi(x)) = \arg \min_{\nu \in \mathcal{M}_x} \int_{\{x\} \times Y \times Z} L d\nu$$

is measurable, as it is the composition of two measurable maps. The image $H(x)$ is closed, so we may apply the Kuratowski–Ryll–Nardzewski Selection Theorem [3, Th. 18.13] to get an appropriate measurable function $x \mapsto \nu_x \in H(x)$.

Letting $\mu = \nu_x(y, z) dx$ we obtain a measure that is viable for M_r^{aff} ; to see that it satisfies (7) note that since $\nu_x \in \mathcal{M}_x$, for each x the averages (12) on Y and Z coincide almost everywhere with $y(x)$ and $z(x) = Dy(x)$, respectively, and (7) is linear in these variables, so the fact that the Liouville equation for affine test functions,

$$\int_{\Omega} \left[\frac{\partial \phi_1}{\partial x}(x) + y(x)^\top \frac{\partial \phi_2}{\partial x}(x) + \phi_2(x)^\top Dy(x) \right] dx = \int_{\partial \Omega} (\phi_1(x) + y(x)^\top \phi_2(x)) \mathbf{n}(x) d\sigma(x),$$

is verified for $y(\cdot)$ implies that (7) is verified for μ . The measure μ also satisfies

$$\int_{\Omega \times Y \times Z} L(x, y, z) d\mu(x, y, z) = \int_{\Omega} \int_{Y \times Z} L(x, y, z) d\nu_x(y, z) dx = \int_{\Omega} \hat{L}(x, y(x), Dy(x)) dx$$

by (17). Arguing analogously in the boundary $\partial \Omega$ to obtain a measure μ_{∂} supported in $\bigcup_{x \in \partial \Omega} \{x\} \times K_{\partial}(x)$ from $y(\cdot)|_{\partial \Omega}$, this implies that $M_r^{\text{aff}} \leq \hat{M}$.

Now we turn to showing that $\hat{M} \leq M_r^{\text{aff}}$. Let (μ, μ_{∂}) be viable for M_r^{aff} . Define $y(x)$ and $z(x)$ for $x \in \Omega$ as in (12); then the same arguments as in the proof of Theorem 3.1 show that $y(\cdot)$ is a $W^{1,p}$ function on Ω whose weak derivative is $Dy(x) = z(x)$ for almost every $x \in \Omega$, as well as the fact that the average of y with respect to μ_{∂} coincides with the trace $y(\cdot)|_{\partial \Omega}$. Moreover, $(y(x), Dy(x)) \in \text{conv } K(x)$ for almost every $x \in \Omega$, since it is an average of points in $(\text{supp } \mu) \cap (\{x\} \times Y \times Z) \subseteq K(x)$, and similarly $y(x) \in \text{conv } K_{\partial}(x)$ for a.e. $x \in \partial \Omega$. By the Jensen inequality, we have

$$\int_{\Omega \times Y \times Z} L(x, y, z) d\mu(x, y, z) \geq \int_{\Omega \times Y \times Z} \hat{L}(x, y, z) d\mu(x, y, z) \geq \int_{\Omega} \hat{L}(x, y(x), Dy(x)) dx.$$

This shows that $\hat{M} \leq M_r^{\text{aff}}$. □

Remark 4.3 (Comparison with the relaxation of [4]). In [4], the consideration of the so-called pointwise dual relaxation \mathcal{L}_{pdr} is suggested; roughly speaking¹, this corresponds to replacing the minimization of the integral of L by

$$\mathcal{L}_{\text{pdr}} \approx \sup_{\varphi \in C^2(\Omega \times Y; Z)} \int_{\Omega} \inf_{\substack{y \in Y \\ z \in Z}} \left[L(x, y, z) + \frac{\partial \varphi}{\partial x} + \sum_j \frac{\partial \varphi}{\partial y_j} z_j \right] dx - \int_{\partial \Omega} \varphi(x, y) \cdot \eta(x) d\sigma(x).$$

In [4] it is proven that this provides a lower bound $\mathcal{L}_{\text{pdr}} \leq \hat{M}$, i.e., the occupation measure relaxation is always at least as strong as \mathcal{L}_{pdr} . The two are related by duality theory and the situation of when strong duality occurs (i.e., $\mathcal{L}_{\text{pdr}} = \hat{M}$) is analyzed in [8].

Example 4.4 (With integral constraints we may have $\hat{M} < M_{\text{r}}^{\text{aff}}$; nonconvex Y). Let $Y \subset \mathbb{R}^2$ be the set containing only the four points (y_1, y_2) with $y_i \in \{-1, 1\}$. Let $\Omega = (0, 1) \subset \mathbb{R}$ and $Z = \overline{B(0, 1)} \subset \mathbb{R}^2$ the closed unit disc. Thus $K(x) = \{(\pm 1, \pm 1)\} \times \overline{B(0, 1)}$ for all $x \in (0, 1)$. Let, for $(x, y, z) \in \Omega \times Y \times Z$ and writing $y = (y_1, y_2)$,

$$L(x, y, z) = y_1 y_2 + \|z\|^2, \quad C(x, y, z) = -y_1 y_2.$$

Observe that $y_1 y_2$ only takes the values ± 1 on Y . The convexifications in (y, z) are equal to $\text{conv } Y = [-1, 1]^2$, $\text{conv } Z = Z = \overline{B(0, 1)}$, $\text{conv } K(x) = [-1, 1]^2 \times \overline{B(0, 1)}$, and, for $x \in (0, 1)$ and $(y, z) \in \text{conv } K(x)$,

$$\hat{L}(x, y, z) = |y_1 + y_2| - 1 + \|z\|^2, \quad \hat{C}(x, y, z) = |y_2 - y_1| - 1.$$

Consider the problem

$$\begin{aligned} \hat{M} = & \inf_{y \in W^{1, \infty}(\Omega; \text{conv } Y)} \int_{\Omega} \hat{L}(x, y(x), Dy(x)) dx & (18) \\ \text{subject to} & (y(x), Dy(x)) \in \text{conv } K(x), \quad \text{a.e. } x \in \Omega, \\ & \int_{\Omega} \hat{C}(x, y(x), Dy(x)) dx \leq 0. \end{aligned}$$

This problem is analogous to (15) in the bulk, with the addition of the integral constraint. Let us compare it with the problem (9) of finding $M_{\text{r}}^{\text{aff}}$ (with $L_{\partial} = A_i = A_{\partial, i} = B_i = B_{\partial, i} = C_{\partial, i} = 0$). If $y: (0, 1) \rightarrow \mathbb{R}^2$ is a constant curve whose image is a point in the segment joining $(-1, 1)$ to $(1, -1)$, then the integral of \hat{L} takes its minimal value of -1 , and since the points in that segment that are close to the origin satisfy $\hat{C}(x, y, z) < 0$, we conclude that $\hat{M} = -1$. On the other hand, for a measure μ viable in (9) we have

$$\begin{aligned} \int_{\Omega \times Y \times Z} L(x, y, z) d\mu(x, y, z) &= \int_{\Omega \times Y \times Z} y_1 y_2 + \|z\|^2 d\mu(x, y, z) \\ &= \int_{\Omega \times Y \times Z} -C(x, y, z) + \|z\|^2 d\mu(x, y, z). \end{aligned}$$

Thus in order to satisfy the constraint that the integral of C be ≤ 0 , the integral of L must be ≥ 0 . The integral of L achieves its minimal value of 0 by letting μ be uniformly distributed on $\Omega \times Y \times \{0\}$. This means that $M_{\text{r}}^{\text{aff}} = 0 > -1 = \hat{M}$.

Example 4.5 (With integral constraints we may have $\hat{M} < M_{\text{r}}^{\text{aff}}$; nonconvex Z). We now modify Example 4.4 slightly to show that a similar situation can occur when the nonconvexity appears in the set of allowable velocities Z .

Let $Z \subset \mathbb{R}^2$ be the set containing only the four points (z_1, z_2) with $z_i \in \{-1, 1\}$. Let $\Omega = (0, 1) \subset \mathbb{R}$ and $Y = \overline{B(0, 1)} \subset \mathbb{R}^2$ the closed unit disc. Thus $K(x) = \overline{B(0, 1)} \times \{(\pm 1, \pm 1)\}$ for all $x \in (0, 1)$. Let, for $(x, y, z) \in \Omega \times Y \times Z$ and writing $z = (z_1, z_2)$,

$$L(x, y, z) = z_1 z_2, \quad C(x, y, z) = -z_1 z_2.$$

¹For the full definition of the pointwise dual relaxation, we refer the reader to [4, §2].

Observe that $z_1 z_2$ only takes the values ± 1 on Z . The convexifications in (y, z) are equal to $\text{conv } Y = Y = \overline{B(0, 1)}$, $\text{conv } Z = [-1, 1]^2$, $\text{conv } K(x) = \overline{B(0, 1)} \times [-1, 1]^2$, and, for $x \in (0, 1)$ and $(y, z) \in \text{conv } K(x)$,

$$\hat{L}(x, y, z) = |z_1 + z_2| - 1, \quad \hat{C}(x, y, z) = |z_2 - z_1| - 1.$$

Consider the problem (18). Let us compare it with the problem (9) of finding M_r^{aff} (with $L_\partial = A_i = A_{\partial,i} = B_i = B_{\partial,i} = C_{\partial,i} = 0$). If $y: (0, 1) \rightarrow \mathbb{R}^2$ is a curve whose derivatives $Dy(x)$ are, for almost every $x \in (0, 1)$, contained in the segment joining $(-1, 1)$ to $(1, -1)$, then the integral of \hat{L} takes its minimal value of -1 , and since the points in that segment that are close to the origin satisfy $\hat{C}(x, y, z) < 0$, we conclude that $\hat{M} = -1$; for an explicit example, one may take the curve

$$y(x) = \begin{cases} (x, -x), & x \in (0, \frac{1}{2}), \\ (\frac{1}{2} - x, -\frac{1}{2} + x), & x \in [\frac{1}{2}, 1). \end{cases}$$

On the other hand, for a measure μ viable in (9) we have

$$\int_{\Omega \times Y \times Z} L(x, y, z) d\mu(x, y, z) = \int_{\Omega \times Y \times Z} z_1 z_2 d\mu(x, y, z) = - \int_{\Omega \times Y \times Z} C(x, y, z) d\mu(x, y, z).$$

Thus as in the previous example, we find that in order to satisfy the constraint that the integral of C be ≤ 0 , the integral of L must be ≥ 0 . Reasoning similarly as before, the integral of L achieves its minimal value of 0 by letting μ be uniformly distributed on $\Omega \times \{0\} \times Z$, and this gives $M_r^{\text{aff}} = 0 > -1 = \hat{M}$.

5 Optimal control

This section extends our results to optimal control. We first formulate the results under the most general assumptions possible and then present their simplified versions. We shall denote the control inside Ω by $u: \Omega \rightarrow U$ and on the boundary by $u: \partial\Omega \rightarrow U_\partial$. We shall assume that U and U_∂ are complete separable metric spaces (e.g., closed subsets of a Euclidean space). The function $\pi_{\Omega \times Y \times Z}: \Omega \times Y \times Z \times U \rightarrow \Omega \times Y \times Z$ will denote the projection on $\Omega \times Y \times Z$, i.e., $\pi_{\Omega \times Y \times Z}(x, y, z, u) = (x, y, z)$. Similarly $\pi_{\partial\Omega \times Y}: \partial\Omega \times Y \times U_\partial \rightarrow \partial\Omega \times Y$ is defined by $\pi_{\partial\Omega \times Y}(x, y, u) = (x, y)$.

Let $A_i, B_i: \Omega \times Y \times Z \times U \rightarrow \mathbb{R}$ and $C_i: \Omega \times Y \times Z \rightarrow \mathbb{R}$ be Borel measurable functions for all i in a (possibly uncountably-infinite) index set I . Similarly, let $A_{\partial,i}, B_{\partial,i}: \partial\Omega \times Y \times U_\partial \rightarrow \mathbb{R}$ and $C_{\partial,i}: \partial\Omega \times Y \rightarrow \mathbb{R}$ be Borel measurable functions for all $i \in I$. Let also $L: \Omega \times Y \times Z \times U \rightarrow \mathbb{R}$ and $L_\partial: \partial\Omega \times Y \times U_\partial \rightarrow \mathbb{R}$ be measurable and locally bounded functions.

5.1 Problem statement

For $p \in [1, +\infty]$, consider the optimal control problem

$$\begin{aligned} M^{\text{oc}} = & \inf_{\substack{y: \Omega \rightarrow Y \\ u: \Omega \rightarrow U \\ u_\partial: \partial\Omega \rightarrow U_\partial}} \int_{\Omega} L(x, y(x), Dy(x), u(x)) dx + \int_{\partial\Omega} L_\partial(x, y(x), u_\partial(x)) d\sigma(x) \\ & \text{subject to } y \in W^{1,p}(\Omega; Y), \quad u, u_\partial \text{ are measurable,} \\ & A_i(x, y(x), Dy(x), u(x)) = 0, \quad B_i(x, y(x), Dy(x), u(x)) \leq 0, \quad x \in \Omega, \quad i \in I, \\ & A_{\partial,i}(x, y(x), u_\partial(x)) = 0, \quad B_{\partial,i}(x, y(x), u_\partial(x)) \leq 0, \quad x \in \partial\Omega, \quad i \in I, \\ & \int_{\Omega} C_i(x, y(x), Dy(x)) dx \leq 0, \quad \int_{\partial\Omega} C_{\partial,i}(x, y(x)) dx \leq 0, \quad i \in I. \end{aligned}$$

Consider also its relaxation, for $p \in [1, +\infty)$:

$$\begin{aligned}
M_{\Gamma}^{\text{oc}} = & \inf_{\substack{\mu \in \mathcal{M}(\Omega \times Y \times Z \times U) \\ \mu_{\partial} \in \mathcal{M}(\partial\Omega \times Y \times U_{\partial})}} \int_{\Omega \times Y \times Z \times U} L(x, y, z, u) d\mu(x, y, z, u) + \int_{\partial\Omega \times Y \times U_{\partial}} L_{\partial}(x, y, u) d\mu_{\partial}(x, y, u) \\
& \text{subject to } \text{supp } \mu \subseteq \bigcap_{i \in I} A_i^{-1}(0) \cap B_i^{-1}((-\infty, 0]), \\
& \text{supp } \mu_{\partial} \subseteq \bigcap_{i \in I} A_{\partial, i}^{-1}(0) \cap B_{\partial, i}^{-1}((-\infty, 0]), \\
& \int_{\Omega \times Y \times U} C_i(x, y, z) d\mu(x, y, z, u) \leq 0, \\
& \int_{\partial\Omega \times Y \times U_{\partial}} C_{\partial, i}(x, y) d\mu_{\partial}(x, y, u) \leq 0, \quad i \in I, \\
& \mu(\Omega \times Y \times Z \times U) = |\Omega|, \\
& \int_{\Omega \times Y \times Z \times U} \|y\|^p + \|z\|^p d\mu(x, y, z, u) < +\infty, \\
& \int_{\partial\Omega \times Y \times U_{\partial}} \|y\|^p d\mu_{\partial}(x, y, u) < +\infty, \\
& \int_{\Omega \times Y \times Z \times U} \frac{\partial \phi_1}{\partial x} + y^{\top} \frac{\partial \phi_2}{\partial x} + \phi_2^{\top} z d\mu = \int_{\partial\Omega \times Y \times U_{\partial}} (\phi_1 + y^{\top} \phi_2) \mathbf{n} d\mu_{\partial} \\
& \quad \forall \phi_1 \in C^{\infty}(\Omega; \mathbb{R}), \phi_2 \in C^{\infty}(\Omega; \mathbb{R}^m).
\end{aligned}$$

For $p = +\infty$, the constraints involving integrals of $\|\cdot\|^p$ are replaced by the assumption that the supports of μ and μ_{∂} be compact.

Remark 5.1. The measures involved in M_{Γ}^{oc} satisfy a Liouville equation for affine test functions only, that is, for functions $\phi(x, y) = \phi_1(x) + \phi_2(x, y)$, and are thus analogous to the affinely relaxed occupation measures $\mathcal{O}_p^{\text{aff}}$; cf. Remark 2.5. A corollary to Theorem 5.2 analogous to Corollary 3.2 holds, although for brevity we do not state it explicitly.

5.2 General formulation

The functions L and L_{∂} have close relatives that do not depend on U and U_{∂} and that will play a role in our assumptions: let the functions $\bar{L}: \Omega \times Y \times Z \rightarrow \mathbb{R}$ and $\bar{L}_{\partial}: \partial\Omega \times Y \rightarrow \mathbb{R}$ be defined by

$$\begin{aligned}
\bar{L}(x, y, z) &= \inf\{L(x, y, z, u) : u \in U, A_i(x, y, z, u) = 0, B_i(x, y, z, u) \leq 0, i \in I\} \\
\bar{L}_{\partial}(x, y) &= \inf\{L_{\partial}(x, y, u) : u \in U_{\partial}, A_{\partial, i}(x, y, u) = 0, B_{\partial, i}(x, y, u) \leq 0, i \in I\}.
\end{aligned}$$

Assume the following:

OC1. (Fiberwise convexity of the constraints) We assume:

- (a) $\pi_{\Omega \times Y \times Z}(\bigcap_{i \in I} A_i^{-1}(0) \cap B_i^{-1}((-\infty, 0])) \cap \{x\} \times Y \times Z$ is nonempty and convex for all $x \in \Omega$.
- (b) $\pi_{\partial\Omega \times Y}(\bigcap_{i \in I} A_{\partial, i}^{-1}(0) \cap B_{\partial, i}^{-1}((-\infty, 0])) \cap \{x\} \times Y$ is nonempty and convex for all $x \in \partial\Omega$ and all $i \in I$.
- (c) $C_i(x, y, z)$ is convex in (y, z) , and
- (d) $C_{\partial, i}(x, y)$ is convex in y , for all $x \in \Omega$ and all $i \in I$.

OC2. (Fiberwise minima are attained by the controls) The functions L and L_{∂} reach their minima in U and U_{∂} , respectively, once we fix (x, y, z) ; in other words, the sets

$$\begin{aligned}
& \arg \min_{\substack{u \in U \\ A_i(x, y, z, u) = 0 \\ B_i(x, y, z, u) \leq 0}} L(x, y, z, u) \quad \text{and} \quad \arg \min_{\substack{u \in U_{\partial} \\ A_{\partial, i}(x, y, u) = 0 \\ B_{\partial, i}(x, y, u) \leq 0}} L_{\partial}(x, y, u),
\end{aligned}$$

are nonempty and closed for all $(x, y, z) \in \Omega \times Y \times Z$ and $(x, y) \in \partial\Omega \times Y$, respectively.

OC3. (Convexity of the Lagrangian densities)

- (a) $\bar{L}(x, y, z)$ is locally bounded and convex in (y, z) for each fixed $x \in \Omega$.
- (b) $\bar{L}_{\partial}(x, y)$ is locally bounded and convex in y for each fixed $x \in \partial\Omega$.

5.3 Simplified formulation 1

Here is a simple situation that satisfies assumptions OC1–OC3 above. Let $p = +\infty$, $\Omega \subset \mathbb{R}^n$ be as above, $Y \subseteq \mathbb{R}^m$ and $Z \subseteq \mathbb{R}^{m \times n}$ be compact, convex sets with nonempty interior. Let U and U_∂ be compact and convex subsets of \mathbb{R}^d for some $d \in \mathbb{N}$. Assume that, for each $x \in \Omega$, $L(x, y, z, u)$, $B_i(x, y, z, u)$, and $C_i(x, y, z, u)$ are convex in (y, z, u) , and that $A_i(x, y, z, u)$ is affine in (y, z, u) . Then we are in the setting concerning our results below. Similar assumptions on the boundary Lagrangian density L_∂ and in the boundary constraints $A_{\partial,i}$, $B_{\partial,i}$, and $C_{\partial,i}$ are also possible within our framework.

5.4 Simplified formulation 2

Here is another situation that satisfies assumptions OC1–OC3 above. For simplicity we will ignore the boundary part of the problem. Let $p = +\infty$, $\Omega \subset \mathbb{R}^n$ be as above, $Y \subseteq \mathbb{R}^m$ and $Z \subseteq \mathbb{R}^{m \times n}$ be compact, convex sets with nonempty interior. Let U be compact subset of \mathbb{R}^n , and let $f: \Omega \times Y \times U \rightarrow Z$ be a continuous function such that, for each $x \in \Omega$, the image of the map $(x, y, u) \mapsto (y, f(x, y, u))$ is convex². Let $\ell: \Omega \times Y \times U \rightarrow \mathbb{R}$ be a continuous Lagrangian density such that

$$\bar{\ell}(x, y, z) = \inf_{\substack{u \in U \\ f(x, u) = z}} \ell(x, y, u)$$

is convex in (y, z) . Consider the optimal control problem

$$\begin{aligned} & \inf_{(y, u)} \int_{\Omega} \ell(x, y(x), u(x)) \, dx \\ \text{subject to } & Dy(x) = f(x, y(x), u(x)), \quad \text{a.e. } x \in \Omega, \\ & a_i(x, y(x), f(x, y(x), u(x))) = 0, \quad i \in I, \text{ a.e. } x \in \Omega, \\ & b_i(x, y(x), f(x, y(x), u(x))) \leq 0, \quad i \in I, \text{ a.e. } x \in \Omega, \\ & \int_{\Omega} c_i(x, y(x), f(x, y(x), u(x))) \, dx \leq 0, \quad i \in I, \end{aligned}$$

where $a_i, b_i, c_i: \Omega \times Y \times Z \rightarrow \mathbb{R}$ are continuous and satisfy $a_i|_{\{x\} \times Y \times Z}^{-1}(0)$ is convex for each $x \in \Omega$, and b_i and c_i are convex in (y, z) for fixed $x \in \Omega$.

To model this situation in our setup above, we add a constraint of the form

$$A_0(x, y, z, u) = \text{dist}(z, f(x, y, u))$$

so that the constraint $A_0 = 0$ means that the velocity z is controlled by u via $z \in f(x, u)$. We let, for $i \in I$, $i \neq 0$,

$$\begin{aligned} A_i(x, y, z, u) &= a_i(x, y, z), & B_i(x, y, z, u) &= b_i(x, y, z), \\ C_i(x, y, z, u) &= c_i(x, y, z), \end{aligned}$$

and

$$L(x, y, z, u) = \ell(x, y, u) \quad \text{and} \quad \bar{L}(x, y, z) = \bar{\ell}(x, y, z).$$

With these assumptions, one can check easily that OC1–OC3 hold.

5.5 Optimal control – main result

In the optimal control context outlined above, we have the following theorem.

Theorem 5.2. *In the setting described above, with assumptions OC1–OC3, we have $M^{\text{oc}} = M_r^{\text{oc}}$.*

²This is true for example when f depends only on (x, u) and $f(\{x\} \times U)$ is convex for each $x \in \Omega$.

Proof. To see that $M^{\text{oc}} \geq M_r^{\text{oc}}$, note that for every triad of functions (y, u, u_∂) that are minimization candidates in M^{oc} , we can produce a pair of measures (μ, μ_∂) defined by (cf. the proof of Theorem 3.1)

$$\begin{aligned} \int_{\Omega \times Y \times Z \times U} f(x, y, z, u) d\mu(x, y, z, u) &= \int_{\Omega} f(x, y(x), Dy(x), u(x)) dx, \quad f \in C(\Omega \times Y \times Z \times U), \\ \int_{\partial\Omega \times Y \times U_\partial} f(x, y, u) d\mu_\partial(x, y, u) &= \int_{\partial\Omega} f(x, y(x), u_\partial(x)) d\sigma(x), \quad f \in C(\partial\Omega \times Y \times U_\partial); \end{aligned}$$

then it is immediate that the pair (μ, μ_∂) is a minimization candidate in M_r^{oc} that satisfies all the relevant constraints, and also satisfies

$$\int_{\Omega \times Y \times Z \times U} L d\mu + \int_{\partial\Omega \times Y \times U_\partial} L_\partial d\mu_\partial = \int_{\Omega} L(x, y(x), Dy(x), u(x)) dx + \int_{\partial\Omega} L_\partial(x, y(x), u_\partial(x)) d\sigma(x).$$

Let us see that $M^{\text{oc}} \leq M_r^{\text{oc}}$. Pick a minimization candidate (μ, μ_∂) for M_r^{oc} satisfying all the constraints of that problem. As in the proof of Theorem 3.1, disintegrate

$$\mu(x, y, z, u) = \nu(y, z, u \mid x) dx, \quad \mu_\partial(x, y, u) = \nu_\partial(y, u \mid x) d\sigma(x),$$

and define

$$\begin{aligned} y(x) &= \int_{Y \times Z \times U} y d\nu(y, z, u \mid x), \quad z(x) = \int_{Y \times Z \times U} z d\nu(y, z, u \mid x), \\ y_\partial(x) &= \int_{Y \times U_\partial} y d\nu_\partial(y, u \mid x). \end{aligned}$$

Then the same arguments as in the proof of Theorem 3.1 show that $z(\cdot)$ is in L^p , that $Dy(\cdot) = z(\cdot)$ weakly, that $y(\cdot)$ is in $W^{1,p}$, and that $y_\partial(\cdot)$ is the Sobolev trace of $y(\cdot)$.

We pick measurable functions $u(\cdot): \Omega \rightarrow U$ and $u_\partial(\cdot): \partial\Omega \rightarrow U_\partial$ satisfying

$$u(x) \in \underset{\substack{u \in U \\ A_i(x, y(x), z(x), u) = 0 \\ B_i(x, y(x), z(x), u) \leq 0}}{\arg \min} L(x, y(x), z(x), u), \quad x \in \Omega, \quad (19)$$

$$u_\partial(x) \in \underset{\substack{u \in U_\partial \\ A_{\partial, i}(x, y_\partial(x), u) = 0 \\ B_{\partial, i}(x, y_\partial(x), u) \leq 0}}{\arg \min} L_\partial(x, y_\partial(x), u), \quad x \in \partial\Omega. \quad (20)$$

Let us explain why we can do this. The set-valued maps induced by the right-hand sides of (19) and (20) $\Omega \rightrightarrows U$ and $\partial\Omega \rightrightarrows U_\partial$ have nonempty and closed images by OC2 and have measurable graphs, so we may apply a Characterization Theorem [1, Th. 8.1.4] to conclude that they are weakly measurable. Then we may apply the Kuratowski–Ryll–Nardzewski Selection Theorem (see [3, Th. 18.13] or [1, Th. 8.1.3]) to get the functions $u(x)$ and $u_\partial(x)$.

The arguments used in the proof of Theorem 3.1 to prove that the constraints involving $A_i, B_i, C_i, A_{\partial, i}, B_{\partial, i}, C_{\partial, i}$ are satisfied by the points $(x, y(x), z(x))$, $x \in \Omega$, and $(x, y(x))$, $x \in \partial\Omega$, carry through to the present situation with minimal modifications for the points $(x, y(x), z(x), u(x))$, $x \in \Omega$, and $(x, y(x), u_\partial(x))$, $x \in \partial\Omega$, using OC1.

Thus with this choice, the triad $(y(\cdot), u(\cdot), u_\partial(\cdot))$ is a minimization candidate for M^{oc} , and by Jensen's inequality and OC3 we have

$$\begin{aligned} M^{\text{oc}} &\leq \int_{\Omega} L(x, y(x), z(x), u(x)) dx + \int_{\Omega_\partial} L_\partial(x, y(x), u_\partial(x)) d\sigma(x) \\ &= \int_{\Omega} \bar{L}(x, y(x), z(x)) dx + \int_{\Omega_\partial} \bar{L}_\partial(x, y(x)) d\sigma(x) \\ &\leq \int_{\Omega \times Y \times Z \times U} \bar{L}(x, y, z) d\mu(x, y, z, u) + \int_{\Omega_\partial \times Y \times U_\partial} \bar{L}_\partial(x, y) d\mu_\partial(x, y, u) \\ &\leq \int_{\Omega \times Y \times Z \times U} L(x, y, z, u) d\mu(x, y, z, u) + \int_{\Omega_\partial \times Y \times U_\partial} L_\partial(x, y, u) d\mu_\partial(x, y, u), \end{aligned}$$

which implies indeed that $M^{\text{oc}} \leq M_r^{\text{oc}}$. \square

6 Application to micromagnetics

In this section we describe how variational problems appearing in micromagnetics fit our convexity assumptions and hence can be solved with our linear formulation with occupation measures.

We confine ourselves to the two-dimensional situation. However, results of this section hold also in the N -dimensional case for any $N \geq 2$. See e.g. [19]. In the classical theory of rigid ferromagnetic bodies, based mainly on [2], a *magnetization* $m: \mathcal{D} \rightarrow \mathbb{R}^2$, describing the state of the body $\mathcal{D} \subset \mathbb{R}^2$, depends on a position $x \in \mathcal{D}$ and has a given temperature-dependent magnitude

$$|m(x)| = w(T) \text{ for almost all } x \in \mathcal{D},$$

where $w(T) = 0$ for $T \geq T_c$, the so-called Curie point. Let us treat the case when the temperature is fixed below the Curie point and thus we shall assume that $|m| = 1$ almost everywhere in \mathcal{D} . In the so-called no-exchange formulation, the energy of a large rigid ferromagnetic body $\mathcal{D} \subset \mathbb{R}^2$ consists of three parts and the variational principle governing equilibrium configurations can be stated (see e.g. [13] and references therein) as the problem of minimizing

$$E(m) = \int_{\mathcal{D}} f(m(x)) dx - \int_{\mathcal{D}} H(x) \cdot m(x) dx + \frac{1}{2} \int_{\mathcal{D}} \nabla u_m(x) \cdot m(x) dx \quad (21)$$

with respect to m and u_m , where $f: \mathcal{S}^1 \rightarrow [0; +\infty)$ is continuous function defined on the unit sphere in \mathbb{R}^2 centered at the origin \mathcal{S}^1 . Further, $m: \mathcal{D} \rightarrow \mathbb{R}^2$ satisfies

$$|m(x)| = 1 \quad \text{a.e. } x \in \mathcal{D}, \quad (22)$$

$H: \mathcal{D} \rightarrow \mathbb{R}^2$ is an applied external magnetic field and $u_m: \mathcal{D} \rightarrow \mathbb{R}$ is a potential of an induced magnetic field. The first term in (21) is an anisotropy energy with a density f which is supposed to be an even nonnegative function depending on material properties and exhibiting crystallographic symmetry. The second term is an interaction energy and the last term is a magnetostatic energy coupled with the magnetization field through the equation

$$\operatorname{div}(-\nabla u_m + m\chi_{\mathcal{D}}) = 0 \text{ in } \mathbb{R}^2, \quad (23)$$

where $\chi_{\mathcal{D}}: \mathbb{R}^2 \rightarrow \{0, 1\}$ is the characteristic function of \mathcal{D} . Equation (23) is interpreted in the classical weak sense, i.e.,

$$\int_{\mathbb{R}^2} \nabla \phi \cdot (-\nabla u_m + m\chi_{\mathcal{D}}) dx = 0 \quad \forall \phi \in C_0^\infty(\mathbb{R}^2). \quad (24)$$

Let us denote

$$\mathcal{A} := \{m \in L^2(\mathcal{D}; \mathbb{R}^2); |m(x)| = 1 \text{ for almost all } x \in \mathcal{D}\}.$$

Eventually, we are concerned with the problem of minimizing

$$E(m) = \int_{\mathcal{D}} f(m(x)) dx - \int_{\mathcal{D}} H(x) \cdot m(x) dx + \frac{1}{2} \int_{\mathcal{D}} \nabla u_m(x) \cdot m(x) dx \quad (25)$$

with respect to m and u_m , and subject to (24). It was shown in [13] that if $N = 1$ the infimum of (25) is not necessarily attained even if $H = 0$. The reason is that minimizing sequences of magnetizations oscillate between \tilde{m}_1 and $-\tilde{m}_1$ keeping the divergence of the magnetization small, so that the gradient of u_m is negligible. This construction leads to a minimizing sequence that weakly converges to zero in $L^2(\mathcal{D}; \mathbb{R}^2)$ and at the same time it shows that $\inf_{m \in \mathcal{A}} E(m) = 0$. However, no minimizer exists and $m = 0$ does not belong to \mathcal{A} . Therefore it is desirable to construct the largest lower semicontinuous envelope of E from (25) and there are well known techniques how to do that [5, 20]. One uses parameterized Young measures, the other one extends f by infinity outside \mathcal{S}^1 and then calculates the convex envelope of this newly defined function. Obviously, the convex envelope is finite only on the closed unit ball in \mathbb{R}^2 . The relaxed problem is convex but not linear, since the stray-field energy, i.e. the last term in (25) is quadratic.

6.1 Formulation with Young measures

As \mathcal{A} is not convex we cannot rely on direct methods [5] to prove the existence of a solution. In fact, the solution to (21) need not exist in \mathcal{A} , cf. [13] for the case $H = 0$. Therefore, it makes sense to look for a relaxation of the problem. It is known [19] that this can be done using Young measures [27]. Moreover, Young measures competing in the minimization problem are supported on \mathcal{S}^1 , i.e. on a compact set. The relaxed problem consists of minimizing

$$E_Y(\nu) := \int_{\mathcal{D}} \int_{\mathcal{S}^1} f(A) \nu_x(dA) dx - \int_{\mathcal{D}} H(x) \cdot m(x) dx + \frac{1}{2} \int_{\mathcal{D}} \nabla u_m(x) \cdot m(x) dx, \quad (26)$$

with respect to ν and u_m , subject to (24) with $m(x) = \int_{\mathcal{S}^1} A \nu_x(dA)$ for almost every $x \in \mathcal{D}$ and $\nu \in \bar{\mathcal{A}} := \mathcal{Y}^\infty(\mathcal{D}; \mathcal{S}^1)$, where $\mathcal{Y}^\infty(\mathcal{D}; \mathcal{S}^1)$ denotes the space of Young measures such that ν_x is a probability measure on \mathcal{S}^1 for a.a. $x \in \mathcal{D}$ and the dependence of ν_x on x is measurable.

On the other hand, it was shown in [6, Thm. 3.4, Remark 3.7] that if $f : \mathcal{S}^1 \rightarrow [0, +\infty)$ and $H \in L^2_{\text{loc}}(\mathcal{D}; \mathbb{R}^2)$ then

$$\inf_{m \in \mathcal{A}} E(m) = \min_{\nu \in \bar{\mathcal{A}}} E_Y(\nu) = \min_{m \in \mathcal{A}^{**}} E^{**}(m) \quad (27)$$

where

$$\mathcal{A}^{**} = \{m \in L^2(\mathcal{D}; \mathbb{R}^2); |m(x)| \leq 1 \text{ for almost every } x \in \mathcal{D}\}$$

and

$$E^{**}(m) = \int_{\mathcal{D}} \hat{f}^{**}(m(x)) dx - \int_{\mathcal{D}} H(x) \cdot m(x) dx + \frac{1}{2} \int_{\mathcal{D}} \nabla u_m(x) \cdot m(x) dx, \quad (28)$$

and $\hat{f}^{**} : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\hat{f}^{**} = \sup \{f; f \leq \hat{f}, f \text{ is convex}\}$ is the convex envelope of $\hat{f} : \mathbb{R}^2 \rightarrow \mathbb{R} \cup \{+\infty\}$,

$$\hat{f}(m) = \begin{cases} f(m) & \text{if } |m| = 1 \\ +\infty & \text{otherwise.} \end{cases}$$

Again, u_m in (28) is calculated from (24). The functional E^{**} is sequentially weakly lower semicontinuous on \mathcal{A}^{**} and it possesses a minimum on \mathcal{A}^{**} . We again emphasize that both (26) and (28) are convex but nonlinear. The formulation of these problems in terms of occupation measures leads to a linear problem that makes it attractive for numerical solution. Similarly, this applies to some problems in linear and nonlinear elasticity where the lower semicontinuous envelope of the original problem is convex; see [14, 23] for some examples.

6.2 Formulation with occupation measures

To formulate the problem (25) in terms of occupation measures following Definition 2.7, let Ω be a subset of \mathbb{R}^2 containing \mathcal{D} , and let $Y = \mathbb{R}^3$ and $Z = \mathbb{R}^{2 \times 3}$. Right away, we see that the problem (25) with constraints (22) and (24) can be modelled as

$$\min_{(\mu, \mu_\partial) \in \mathcal{O}_2} \int_{\Omega \times Y \times Z} f(y_1, y_2) - H(x) \cdot (y_1, y_2) + \frac{1}{2} (z_3 \cdot (y_1, y_2)) d\mu(x, y, z) \quad (29)$$

$$\text{subject to } \text{supp } \mu \subseteq \Omega \times (S^1 \times \mathbb{R}) \times Z \subset \Omega \times Y \times Z, \quad (30)$$

$$\int_{\Omega \times Y \times Z} \nabla \phi(x) \cdot (-z_3 + (y_1, y_2) \chi_{\mathcal{D}}(x)) d\mu(x, y, z) = 0, \quad \phi \in C_0^1(\Omega). \quad (31)$$

Observe that the solutions of (25) are of the form $y(x) = (m(x), u(x))$, where $m(x)$ is in the unit circle $S^1 \subset \mathbb{R}^2$ for all $x \in \Omega$, and $u(x) \in \mathbb{R}$. In the formulation above, (y_1, y_2) correspond to $m(x)$ and y_3 corresponds to $u(x)$, so that also z_3 corresponds to $\nabla u(x)$.

Let us now show how to put this in the form of our problem (10). To get the integral in (29), let

$$L(x, y, z) = f(y_1, y_2) - H(x) \cdot (y_1, y_2) + \frac{1}{2} (z_3 \cdot (y_1, y_2)), \quad (x, y, z) \in \Omega \times Y \times Z.$$

To model the constraint (30), let

$$A(x, (y_1, y_2, y_3), z) = y_1^2 + y_2^2 - 1.$$

To model the constraint (31), we use the trick described in Example 2.3: we use $I = C_0^1(\Omega)$ as index set and we let

$$C_\phi(x, y, z) = \nabla\phi(x) \cdot (-z_3 + (y_1, y_2)\chi_{\mathcal{D}}(x)), \quad \phi \in C_0^1(\Omega).$$

All other functions $L_\partial, A_{i,\partial}, B_i, B_{i,\partial}, C_{i,\partial}$ in (10) can be set to 0.

Observe that even if f is convex, Theorem 3.1 cannot be applied to (29)–(31) above because the constraint (30) is not convex. Instead, we can apply Theorem 4.2, which tells us that if we take the affine occupation relaxation (9), that is, if we minimize (29) over $\mathcal{O}_2^{\text{aff}}$ (rather than \mathcal{O}_2 ; see (32) below), then we get the same result \hat{M} as in the convexification (15). The latter coincides with the problem of minimizing E^{**} as in (28) over \mathcal{A}^{**} and, by (27), with (25). We have proved:

Proposition 6.1. *The problem (25) is equivalent to the (computationally tractable) problem*

$$\begin{aligned} \min_{(\mu, \mu_\partial) \in \mathcal{O}_2^{\text{aff}}} \quad & \int_{\Omega \times Y \times Z} f(y_1, y_2) - H(x) \cdot (y_1, y_2) + \frac{1}{2}(z_3 \cdot (y_1, y_2)) d\mu(x, y, z) \quad (32) \\ \text{subject to} \quad & (30) \text{ and } (31), \end{aligned}$$

that is, the minimum in (32) is equal to the infimum in (25).

Remark 6.2. Since the occupation measure relaxation (29) is sandwiched between the affinely relaxed problem (32) and the original problem (25), it follows from Proposition 6.1 that these are all equal to each other.

A Constancy lemma

The following is essentially a version of the what is known as the Constancy Theorem in Geometric Measure Theory, and it is a slight generalization of [16, Lemma 2.9]. We use it in the proof of Theorem 3.1.

Lemma A.1. *Let Ω be an bounded, connected, open subset of \mathbb{R}^n . Let μ be a positive, compactly-supported, Borel measure on Ω such that $\mu(\Omega) = |\Omega|$ and*

$$\int_{\Omega} \nabla\phi(x) d\mu(x) = 0 \quad (33)$$

for all compactly-supported $\phi \in C_0^\infty(\Omega; \mathbb{R})$. Then μ equals the n -dimensional Lebesgue measure on Ω .

Proof. Let $R \subset \Omega$ be a small parallelepiped, and let τ be a translation such that $\tau(R) \subset \Omega$. We will show that $\mu(R) = \mu(\tau(R))$, and since this will be true for all R and all τ , μ must be a positive multiple of Lebesgue on Ω [25, Thm. 2.20]. Write τ as a finite composition of translations τ_i in the directions of the axes x_1, \dots, x_n ,

$$\tau = \tau_k \circ \tau_{k-1} \circ \dots \circ \tau_1.$$

Denote $\tilde{\tau}_i = \tau_i \circ \tau_{i-1} \circ \dots \circ \tau_1$ and set $\tilde{\tau}_0$ equal to the identity. We assume τ_1, \dots, τ_k have been chosen also in a such a way that the convex hull of $\tilde{\tau}_{i-1}(R) \cup \tilde{\tau}_i(R)$ is contained in Ω for each i . For each $i = 1, \dots, k$, let j_i be such that τ_i is a translation in direction x_{j_i} . Recall that $\chi_{\tilde{\tau}_i(R)}$ is the indicator function of the translated rectangle $\tilde{\tau}_i(R)$, and let

$$\begin{aligned} \psi_i(x_1, \dots, x_n) = \int_{-\infty}^{x_{j_i}} \chi_{\tilde{\tau}_i(R)}(x_1, \dots, x_{j_i-1}, s, x_{j_i+1}, \dots, x_n) \\ - \chi_{\tilde{\tau}_{i-1}(R)}(x_1, \dots, x_{j_i-1}, s, x_{j_i+1}, \dots, x_n) ds. \end{aligned}$$

Observe that

$$\text{supp } \phi_i = \overline{\text{conv}}(\tilde{\tau}_{i-1}(R) \cup \tilde{\tau}_i(R)),$$

which is a compact set properly contained in Ω . Approximating with smooth, compactly-supported functions and using the Lebesgue dominated convergence theorem, we conclude that (33) is true for $\phi = \psi_i$, which means, for the j_i^{th} entry,

$$\mu(\tilde{\tau}_i(R)) - \mu(\tilde{\tau}_{i-1}(R)) = \int_{\Omega} \chi_{\tilde{\tau}_i(R)} - \chi_{\tilde{\tau}_{i-1}(R)} d\mu = \int_{\Omega} \frac{\partial \psi_i}{\partial x_{j_i}} d\mu = 0.$$

By induction we get

$$\mu(R) = \mu(\tilde{\tau}_0(R)) = \mu(\tilde{\tau}_k(R)) = \mu(\tau(R)).$$

Thus $\mu = c dx$ for some $c > 0$. Since $\mu(\Omega) = |\Omega|$, we conclude that $c = 1$. \square

B Technical lemma for the definition of occupation measures

Lemma B.1. *Let $p \in [1, +\infty)$, and let μ and μ_{∂} be measures satisfying the second item in Definition 2.4. Then for all $\phi \in \mathbf{F}_p$ we have that the functions*

$$\phi(x, y)\mathbf{n}(x) \quad \text{and} \quad \frac{\partial \phi}{\partial x}(x, y) + z^{\top} \frac{\partial \phi}{\partial y}(x, y)$$

are integrable with respect to μ_{∂} and μ , respectively.

Proof. This is true for $p = +\infty$ because in that case the supports of the measures involved are bounded and the integrands are continuous. For $p \in [1, +\infty)$ we have, from (6),

$$\begin{aligned} \int_{\partial\Omega \times Y} \|\phi(x, y)\mathbf{n}(x)\| d\mu(x, y) &\leq \int_{\partial\Omega \times Y} c(1 + \|y\|^p)\|\mathbf{n}(x)\| d\mu_{\partial}(x, y) \\ &= \int_{\partial\Omega \times Y} c(1 + \|y\|^p) d\mu_{\partial}(x, y) < +\infty \end{aligned}$$

and, by the Hölder inequality with $\frac{1}{p} + \frac{1}{q} = 1$, so that $p/q = p - 1$, and using (6),

$$\begin{aligned}
& \int_{\Omega \times Y \times Z} \left\| \frac{\partial \phi}{\partial x}(x, y) + z^\top \frac{\partial \phi}{\partial y}(x, y) \right\|_1 d\mu(x, y, z) \\
& \leq \int_{\Omega \times Y \times Z} \left\| \frac{\partial \phi}{\partial x}(x, y) \right\| + \sum_i \left| z_i^\top \frac{\partial \phi}{\partial y}(x, y) \right| d\mu(x, y, z) \\
& \leq \int_{\Omega \times Y \times Z} c(1 + \|y\|^p) d\mu(x, y, z) \\
& \quad + \sum_i \left(\int_{\Omega \times Y \times Z} \|z_i\|^p d\mu(x, y, z) \right)^{\frac{1}{p}} \left(\int_{\Omega \times Y \times Z} \left\| \frac{\partial \phi}{\partial y}(x, y) \right\|^q d\mu(x, y, z) \right)^{\frac{1}{q}} \\
& \leq \int_{\Omega \times Y \times Z} c(1 + \|y\|^p) d\mu(x, y, z) \\
& \quad + \sum_i \left(\int_{\Omega \times Y \times Z} \|z_i\|^p d\mu(x, y, z) \right)^{\frac{1}{p}} \left(\int_{\Omega \times Y \times Z} c^q (1 + \|y\|^{p-1})^q d\mu(x, y, z) \right)^{\frac{1}{q}} \\
& \leq \int_{\Omega \times Y \times Z} c(1 + \|y\|^p) d\mu(x, y, z) \\
& \quad + \sum_i \left(\int_{\Omega \times Y \times Z} \|z_i\|^p d\mu(x, y, z) \right)^{\frac{1}{p}} \\
& \quad \times \left(\left(\int_{\Omega \times Y \times Z} c^q d\mu(x, y, z) \right)^{\frac{1}{q}} + \left(\int_{\Omega \times Y \times Z} c^q \|y\|^{p-1})^q d\mu(x, y, z) \right)^{\frac{1}{q}} \right) \\
& \leq \int_{\Omega \times Y \times Z} c(1 + \|y\|^p) d\mu(x, y, z) \\
& \quad + \sum_i \left(\int_{\Omega \times Y \times Z} \|z_i\|^p d\mu(x, y, z) \right)^{\frac{1}{p}} \\
& \quad \times \left(c\mu(\Omega \times Y \times Z)^{\frac{1}{q}} + c \left(\int_{\Omega \times Y \times Z} \|y\|^p d\mu(x, y, z) \right)^{\frac{1}{q}} \right) \\
& < +\infty. \quad \square
\end{aligned}$$

Acknowledgment. The work of supported by the Barrande project 44702QH and the MŠMT ČR project 8J20FR019. M. Kružík was further supported by the GAČR project 19-26143X. The research of R. Ríos-Zertuche and M. Korda was partly supported by AI Interdisciplinary Institute ANITI funding, through the French “Investing for the Future PIA3” program under grant agreement n° ANR-19-PI3A-0004. This research was also part of the programme DesCartes and is supported by the National Research Foundation, Prime Minister’s Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. This work was co-funded by the European Union under the project ROBOPROX (reg. no. CZ.02.01.01/00/22.008/0004590).

References

- [1] J.P. Aubin, H. Frankowska. Set-valued analysis. Springer, 2009.
- [2] W.F. Brown Jr. Micromagnetics. John Wiley & Sons, 1963.
- [3] D. Charalambos, B. Aliprantis. Infinite dimensional analysis: a hitchhiker’s guide. Springer, 2013.
- [4] A. Chernyavsky, J. J. Bramburger, G. Fantuzzi, D. Goluskin. Convex relaxations of integral variational problems: pointwise dual relaxation and sum-of-squares optimization. SIAM J. Optim. 33:481–512, 2023.

- [5] B. Dacorogna. Introduction to the calculus of variations. Imperial College Press, 2004.
- [6] A. DeSimone. Energy minimizers for large ferromagnetic bodies. *Arch. Rat. Mech. Anal.* 125:99–143, 1993.
- [7] L. C. Evans, R. F. Gariepy. Measure theory and fine properties of functions. CRC Press, 1992.
- [8] G. Fantuzzi, I. Tobasco. Sharpness and non-sharpness of occupation measure bounds for integral variational problems. [arXiv:2207.13570](https://arxiv.org/abs/2207.13570), 2022.
- [9] V. Gaitsgory, M. Quincampoix. Linear programming analysis of deterministic infinite horizon optimal control problems with discounting. *SIAM J. Control Optim.* 48:2480–2512, 2009.
- [10] D. Henrion, M. Korda. Convex computation of the region of attraction of polynomial control systems. *IEEE Trans. Autom. Control* 59:297–312, 2014.
- [11] D. Henrion, M. Korda, J. B. Lasserre. The moment-SOS hierarchy - Lectures in probability, statistics, computational geometry, control and nonlinear PDEs. World Scientific, 2020.
- [12] D. Henrion, M. Kružík, T. Weisser. Optimal control problems with oscillations, concentrations and discontinuities. *Automatica* 103:159–165, 2019.
- [13] R. D. James, D. Kinderlehrer. Frustration in ferromagnetic materials. *Continuum Mech. Thermodyn.* 2:215–239, 1990.
- [14] R. V. Kohn. The relaxation of a double-well energy. *Continuum Mech. Thermodyn.* 3:193–236, 1991.
- [15] M. Korda, D. Henrion, J. B. Lasserre. Moments and convex optimization for analysis and control of nonlinear partial differential equations. In E. Trélat, E. Zuazua (Eds). *Numerical Control. Handbook of Numerical Analysis*, Elsevier, 2021.
- [16] M. Korda, R. Ríos-Zertuche. The gap between a variational problem and its occupation measure relaxation. [arXiv:2205.14132](https://arxiv.org/abs/2205.14132), 2022.
- [17] J. B. Lasserre, D. Henrion, C. Prieur, E. Trélat. Nonlinear optimal control via occupation measures and LMI relaxations. *SIAM J. Control Optim.* 47(4):1643-1666, 2008.
- [18] S. Marx, T. Weisser, D. Henrion, J. B. Lasserre. A moment approach for entropy solutions to nonlinear hyperbolic PDEs. *Mathematical Control and Related Fields* 10:113–140, 2020.
- [19] P. Pedregal. Relaxation in ferromagnetism: the rigid case, *J. Nonlinear Sci.* 4:105–125, 1994.
- [20] P. Pedregal. *Parametrized measures and variational principles*, Birkhäuser, 1997.
- [21] R.R. Phelps. *Lectures on Choquet’s theorem*. Springer, 2001.
- [22] Prokhorov, Yuri V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory of Probability and its Applications.* 1 (2): 157–214.
- [23] A. Raoult. Quasiconvex envelopes in nonlinear elasticity. In J. Schröder, P. Neff (Eds). *Poly-, quasi- and rank-one convexity in applied mechanics*. CISM International Centre for Mechanical Sciences, vol 516. Springer, 2010.
- [24] T. Roubíček. *Relaxation in optimization theory and variational calculus*. W. de Gruyter, 2nd ed., 2020.
- [25] W. Rudin. *Real and complex analysis*. McGraw-Hill, 1968.

- [26] R. Vinter. Convex duality and nonlinear optimal control. *SIAM J. Control Optim.* 31:518–538, 1993.
- [27] L. C. Young. *Lectures on the calculus of variations and optimal control theory*. W. B. Saunders Co., 1969.