Coarse-convex-compactification approach to numerical solution of nonconvex variational problems

René Meziat, Tomáš Roubíček, Diego Patiño

Preprint no. 2010-014

Research Team 1
Mathematical Institute of the Charles University
Sokolovská 83, 186 75 Praha 8
http://ncmm.karlin.mff.cuni.cz/

COARSE-CONVEX-COMPACTIFICATION APPROACH TO NUMERICAL SOLUTION OF NONCONVEX VARIATIONAL PROBLEMS

René Meziat[1], Tomáš Roubíček[2,3], Diego Patiño[1,4]

[1] *Departamento de Matemáticas, Universidad de los Andes, Carrera 1 No.18A-10, Bogotá, Colombia.*
[2] *Mathematical Institute, Charles University, Sokolovská 83, CZ-186 75 Praha 8, Czech Republic.*
[3] *Institute of Thermomechanics of the ASCR, Dolejškova 5, CZ-182 00 Praha 8, Czech Republic.*
[4] *Departamento de Ingeniería Electrónica, Pontificia Universidad Javeriana, Carrera 7 No. 40 - 62, Bogotá, Colombia.*

Abstract. A numerical method for a (possibly non-convex) scalar variational problem for the functional $\Phi(u) := \int_\Omega \varphi_1(x, \nabla u(x)) + \varphi_0(x, u(x))\,\mathrm{d}x$ to be minimized for $u \in W^{1,p}(\Omega)$ and $u|_{\partial\Omega} = u_\mathrm{D}$, is proposed; $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain, $n = 1$ or 2. This method allows for computation of the Young-measure solution of the generalized relaxed version of the original problem and applies to those cases in which $\varphi_1(x, \cdot)$ is polynomial. The Young measures involved in the relaxed problem can be represented by their algebraic moments and a finite-element mesh is used to discretize Ω and thus to approximate both u and the Young measure (in the momentum representation). Eventually, thus obtained convex semidefinite program can be solved by efficient specialized mathematical-programming solvers. This method is justified by convergence analysis and eventually tested on a 2-dimensional benchmark numerical example. It serves as an example how convex compatification can efficiently be used numerically if enough "small", i.e. enough coarse.

Key Words. Relaxed variational problems, convex approximations, method of moments, semidefinite programming.

AMS Subject Classification: 49M05, 65K10, 65N30, 90C22.

1 INTRODUCTION

Nonconvex optimization problems often lack any solution because of fast oscillations of minimizing sequences that eventually break lower semicontinuity with respect to a weak convergence, cf. [66] and references therein for a survey on scalar variational problems which will be the concern in this paper. Therefore, a relaxation is urgent to solve such problems in a suitably generalized sense. The most general way of relaxation is certainly a suitable continuous extension, using also a suitable linear-space structure not necessarily completely coherent with the linear structure occurring in the formulation of the original problem. Thus extended, so-called *relaxed*, problems then may get a convex structure even if the original problem does not have any.

The relaxed problems can be discretized by a theory of convex approximation of the set of the generalized Young functionals developed in [61–63], see also [42, 57, 58] or a survey paper [65]. Thus the relaxed problem can directly be implemented on computer, without approximating the original non-relaxed problem; cf. [12, 30–32, 53, 63] for this approach. If the (additively coupled, cf. e.g. ($\mathscr{P}$) below) problem is linear in a lower-order term (i.e. $\varphi_0(x, \cdot)$ in ($\mathscr{P}$) is linear), such approach leads to a linear-programming problem and was shown very efficient in [3]. In the quadratic case, it naturally leads to a quadratic-programming problem, which is a considerably less efficient but still possible approach if the dimensionality is not too high, cf. [12, 32, 42]. For non-quadratic case, one can still consider various iterative schemes, see Remark 3.9 below.

All the above mentioned references use conventional Young measures and treat them numerically in various more or less sophisticated ways. However, if the particular problem involves only a finite number of nonlinearities, it suffices to consider only moments of these Young measures with respect to these nonlinearities. This is the general idea of *coarse convex compactifications* as thoroughly exposed in [63]. In general, however, it is not easy to characterize explicitly such convex compactifications. The goal of this paper is to exploit this alternative coarse-compactification approach in a particular case where the involved nonlinearities are *polynomials*. We show efficiency of this approach on a concrete problem of scalar multidimensional variational calculus with an *additively coupled* integral functional:

$$(\mathscr{P})\qquad \begin{cases} \text{Minimize} & \Phi(u) := \displaystyle\int_\Omega \varphi_1(x,\nabla u(x)) + \varphi_0(x,u(x))\,\mathrm{d}x,\\ \text{subject to} & u\in W^{1,p}(\Omega),\quad u|_{\partial\Omega}=u_{\mathrm{D}}, \end{cases}$$

where $\varphi_1(x,\cdot)$ is a (possibly non-convex) coercive multidimensional polynomial and $\Omega\subset\mathbb{R}^n$ is a bounded domain with the Lipschitz boundary $\partial\Omega$ and a given boundary condition $u_{\mathrm{D}}\in W^{1-1/p,p}(\partial\Omega)$; for more general problems see Remark 3.13 below. Thorough the whole paper we will assume that $\varphi_1:\Omega\times\mathbb{R}^n\to\mathbb{R}$ and $\varphi_0:\Omega\times\mathbb{R}\to\mathbb{R}$ are Carathéodory functions satisfying, for almost all $x\in\Omega$, all $s\in\mathbb{R}^n$, and all $u\in\mathbb{R}$,

$$c_1|s|^p \le \varphi_1(x,s)\le c_2(1+|s|^p), \tag{1.1a}$$

$$|\varphi_0(x,u)|\le a(x)+c_3|u|^q, \tag{1.1b}$$

where $p>1$, $c_1,c_2,c_3>0$, $a\in L^1(\Omega)$, and $1<q<pn/(n-p)$ if $p<n$ and $1<q<\infty$ if $p\ge n$. The exponent $p>1$ in $(\mathscr{P})$ naturally refers to (1.1a).

The outline and the main contributions of this paper read as follows: Section 2 briefly outlines the ways for relaxation and spatial discretization of $(\mathscr{P})$ in terms of conventional Young measures (and their certain classes of equivalence). Section 3 presents theory and analysis of the proposed method to solve the discretized relaxed problem numerically. In Section 4 we report on the performance of our algorithm applied to a benchmark of a 2-dimensional model problem. Eventually, for reader's convenience, a background of the theory of relaxation by convex compactification and its approximations is briefly exposed in Appendix.

2 RELAXATION OF $(\mathscr{P})$ AND ITS DISCRETIZATION

The first stage to solve $(\mathscr{P})$ is to use a relaxed formulation in terms of Young measures as follows:

$$(\mathscr{RP})\qquad \begin{cases} \text{Minimize} & \bar\Phi(u,\nu) := \displaystyle\int_\Omega\int_{\mathbb{R}^n} \varphi_1(x,s)\nu_x(\mathrm{d}s) + \varphi_0(x,u(x))\,\mathrm{d}x,\\ \text{subject to} & \displaystyle\int_{\mathbb{R}^n} s\nu_x(\mathrm{d}s) = \nabla u(x) \quad \text{for a.a. } x\in\Omega,\\ & u\in W^{1,p}(\Omega),\quad \nu\in\mathscr{Y}^p(\Omega;\mathbb{R}^n),\quad u|_{\partial\Omega}=u_{\mathrm{D}}, \end{cases}$$

where $\mathscr{Y}^p(\Omega;\mathbb{R}^n)$ denotes the set of so-called L^p-Young measures defined as

$$\mathscr{Y}^p(\Omega;\mathbb{R}^n) := \Big\{\nu: x\mapsto\nu_x:\Omega\to\mathcal{P}(\mathbb{R}^n) \text{ weakly* measurable}\\ \text{such that } x\mapsto\int_{\mathbb{R}^n}|s|^p\nu_x(\mathrm{d}s) \text{ belongs to } L^1(\Omega)\Big\} \tag{2.1}$$

where $\mathcal{P}(\mathbb{R}^n)$ is the set of all Borel regular probability measures supported on $\mathbb{R}^n$. The generalized problem in Young measures $(\mathscr{RP})$ is a relaxation of $(\mathscr{P})$ as has been shown in [27, 59]. It is well known that the set $\mathscr{Y}^p(\Omega;\mathbb{R}^n)$ is convex in the Banach space $L^\infty_{\mathrm{w}*}(\Omega;\mathcal{M}(\mathbb{R}^n)) \cong L^1(\Omega;C_0(\mathbb{R}^n))^*$ of weakly* measurable essentially bounded mapping from Ω to the set of regular Borel measures $\mathcal{M}(\mathbb{R}^n) \cong C_0(\mathbb{R}^n)^*$. Moreover, roughly speaking, up to suppressing concentrations it enjoys certain compactness properties. Together with weak* density of $L^p(\Omega;\mathbb{R}^n)$ when embedded through the mapping $v \mapsto \{\delta_{v(x)}\}_{x\in\Omega}$ into $\mathscr{Y}^p(\Omega;\mathbb{R}^n)$, this fits with the theory of so-called *convex compactifications*, cf. [63] or Appendix below.

On the other hand, the above relaxation is unnecessarily "fine", leading thus to unnecessarily large optimization problems after its discretization. In fact, it suffices to formulate it in a "coarser way" in terms of certain classes of equivalences of Young measures. Considering a set H of Carathéodory integrands $\Omega\times\mathbb{R}^n \to \mathbb{R}$, we define the factor set $\nu \overset{H}{\sim} \tilde\nu$ if $\int_{\mathbb{R}^n} h(x,s)\nu_x(\mathrm{d}s) = \int_{\mathbb{R}^n} h(x,s)\tilde\nu_x(\mathrm{d}s)$ for a.a. $x\in\Omega$ and all $h\in H$. Then we define

$$\mathscr{Y}^p_H(\Omega;\mathbb{R}^n) := \mathscr{Y}^p(\Omega;\mathbb{R}^n)/\overset{H}{\sim}\,.$$

For $\eta \in \mathscr{Y}^p_H(\Omega;\mathbb{R}^n)$ and $h\in H$, we denote by $h\bullet\eta$ the function $x\mapsto \int_{\mathbb{R}^n} h(x,s)\nu_x(\mathrm{d}s)$ with some $\nu\in\eta$; obviously, the particular choice of ν from the class η is not important. Then we consider

$$(\mathscr{RP}_H)\qquad \begin{cases} \text{Minimize} & \bar\Phi(u,\eta) := \displaystyle\int_\Omega \big[\varphi_1\bullet\eta\big](x) + \varphi_0(x,u(x))\,\mathrm{d}x,\\ \text{subject to} & [\mathrm{Id}\bullet\eta](x) = \nabla u(x) \quad \text{for a.a. } x\in\Omega,\\ & u\in W^{1,p}(\Omega),\quad \eta\in\mathscr{Y}^p_H(\Omega;\mathbb{R}^n),\quad u|_{\partial\Omega} = u_{\mathrm{D}}, \end{cases}$$

where $\mathrm{Id} : (x,s)\mapsto s$. If H contains all the integrands φ_1 and $(x,s)\mapsto s_i$, $i=1,\dots,n$, then $(\mathscr{RP}\,)$ is obviously equivalent to $(\mathscr{RP}_H)$ in the sense that any solution (u,ν) yields a solution (u,η) to $(\mathscr{RP}_H)$ just by taking $\eta\ni\nu$ and, conversely, for any solution (u,η) to $(\mathscr{RP}_H)$ there always is some $\nu\in\eta$ such that (u,ν) solves $(\mathscr{RP}\,)$.

If H is a linear subspace of $L^1(\Omega;C_0(\mathbb{R}^n))$ or of some of its extension (cf. $\mathrm{Car}^p(\Omega;\mathbb{R}^m)$ in (A.1) below), then the mentioned convexity property of $\mathscr{Y}^p(\Omega;\mathbb{R}^n)$ is inherited also by $\mathscr{Y}^p_H(\Omega;\mathbb{R}^n)$, which is what we will exploit and what is called *coarse-convex-compactification* method.

For a mesh parameter $d>0$, we further consider a triangulation $\mathcal{T}_d$ of Ω which is assumed polyhedral, and then the relaxed discretized optimization problem:

$$(\mathscr{RP}_{H,d})\qquad \begin{cases} \text{Minimize} & \bar\Phi(u,\eta) := \displaystyle\int_\Omega \big[\varphi_1\bullet\eta\big](x) + \varphi_0(x,u(x))\,\mathrm{d}x,\\ \text{subject to} & [\mathrm{Id}\bullet\eta](x) = \nabla u(x) \quad \text{for a.a. } x\in\Omega,\\ & u\in W^{1,p}(\Omega)\ \text{ element-wise affine on } \mathcal{T}_d, \qquad u|_{\partial\Omega} = u_{\mathrm{D}},\\ & \eta\in\mathscr{Y}^p_H(\Omega;\mathbb{R}^n)\ \text{element-wise constant on } \mathcal{T}_d. \end{cases}$$

In [62,63], it has essentially been proved that there exists a solution (y_d,η_d) for $(\mathscr{RP}_{H,d})$ and any (weak$\times$weak*)-cluster point of the sequence of solutions to $(\mathscr{RP}_{H,d})$ for $d\to0$ is a solution to $(\mathscr{RP})$.

Even more, $(\mathscr{RP}_{H,d})$ can be equivalent to $(\mathscr{RP}_{H^d})$ provided $H^d\subset H$ is suitably composed from element-wise constant integrands. In this paper, we focus on taking H^d finite-dimensional so that such approximate problems can directly be implemented on computers, and thus solve $(\mathscr{RP}_{H,d})$ when $\varphi_1(x,\cdot)$ is a coercive multidimensional nonconvex polynomial by representing any $\eta\in\mathscr{Y}^p_H(\Omega;\mathbb{R}^n)$ by a finite set of its algebraic multidimensional moments. Indeed, this paper formalizes ideas introduced in [19, 45–48] where the authors use projections of Young measures on finite dimensional convex bodies.

3 POLYNOMIAL NONLINEARITIES, METHOD OF MOMENTS

Given a positive integer k, we define H_k as

$$H_k := \sum_{l=1}^{2k-1} L^{p/(p-l)}(\Omega) \otimes \Pi_l(\mathbb{R}^n) + G_0 \otimes \Pi_{2k}(\mathbb{R}^n) \tag{3.1}$$

where $p = 2k$ and $k \in \mathbb{N}$ and $\Pi_l(\mathbb{R}^n)$ is the linear space composed of multidimensional polynomials with degree less or equal than l, and where G_0 is given by

$$G_0 := \bigcup_{d>0} \big\{ g \in L^\infty(\Omega);\ \ \forall S \in \mathcal{T}_d :\ \ g|_S \in C(\bar{S}) \big\}. \tag{3.2}$$

Let us just remark that H_k is a linear separable subspace of the natural linear space $\mathrm{Car}^p(\Omega;\mathbb{R}^n)$ defined (A.1)–(A.2) in Appendix.

In this special case, η is a (class of equivalence of) Young measure(s) and can be represented by affine sets of algebraic moments, i.e.

$$m_\iota = \eta \bullet (1 \otimes s_1^{\iota_1} \cdots s_n^{\iota_n}) \tag{3.3}$$

where $\iota = (\iota_1, \cdots, \iota_n)$ is the multi-index of non-negative integers such that $|\iota| := \iota_1 + \cdots + \iota_n \leq 2k$. Namely, for any $h \in H_k$ with H_k from (3.1), i.e. $h = \sum_{|\iota| \leq 2k} g_\iota(x) s_1^{\iota_1} \cdots s_n^{\iota_n}$ with $g_\iota \in L^{p/(p-|\iota|)}(\Omega)$, it holds that

$$\big\langle \eta, h \big\rangle = \sum_{|\iota| \leq 2k} \big\langle \eta, g_\iota \otimes s_1^{\iota_1} \cdots s_n^{\iota_n} \big\rangle = \sum_{|\iota| \leq 2k} \int_\Omega g_\iota(x) m_\iota(x)\,\mathrm{d}x. \tag{3.4}$$

Further, denoting $m = (m_\iota)_{|\iota| \leq 2k}$, we define a matrix $\mathbb{H}_k(m)$ as

$$\mathbb{H}_k(m) := \Big(m_{\iota_1+\iota_1', \cdots, \iota_n+\iota_n'} \Big)_{\substack{0 \leq \iota_1 + \cdots + \iota_n \leq k \\ 0 \leq \iota_1' + \cdots + \iota_n' \leq k}} \tag{3.5}$$

and following the current literature on global optimization of polynomials with moments [21, 25, 33–36, 38–40, 54–56] we define the *localizing* matrix $\mathbb{L}_k(m)$ as

$$\Big(\varrho_d^2 m_{\iota_1+\iota_1', \cdots, \iota_n+\iota_n'} - m_{\iota_1+\iota_1'+2, \cdots, \iota_n+\iota_n'} - \cdots - m_{\iota_1+\iota_1', \cdots, \iota_n+\iota_n'+2} \Big)_{\substack{0 \leq \iota_1 + \cdots + \iota_n \leq k-1 \\ 0 \leq \iota_1' + \cdots + \iota_n' \leq k-1}} \tag{3.6}$$

It is well known that a probability measure μ on $\mathbb{R}^n$ induces a moments sequence $m_\iota = \int_{\mathbb{R}^n} s_1^{\iota_1} \cdots s_n^{\iota_n} \mu(\mathrm{d}s)$ which always makes the matrix $\mathbb{H}_k(m)$ positive semidefinite. The converse, i.e. existence of a probability measure μ inducing a prescribed sequence of moments $(m_\iota)_{|\iota| \leq 2k}$ with $\mathbb{H}_k(m) \succeq 0$ and $m_0(x) = 1$, is unfortunately not true even if $n = 1$ or even when $\mathbb{H}_k(m) \succ 0$ and $n > 1$. The role of the localizing matrix $\mathbb{L}_k(m)$ is revealed when we focus on the family of measures supported on the n-dimensional ball $B_{\varrho_d} := \{ s \in \mathbb{R}^n;\ s_1^2 + \cdots + s_n^2 \leq \varrho_d^2 \}$. Thus, a probability measure μ on B_{ϱ_d} induces a moments sequence $m_\iota = \int_{B_{\varrho_d}} s_1^{\iota_1} \cdots s_n^{\iota_n} \mu(\mathrm{d}s)$ which makes the localizing matrix $\mathbb{L}_k(m)$ positive semidefinite. Even considering the localizing matrix $\mathbb{L}_k(m)$, the converse statement is no longer true again, however something useful can be done by applying recent characterizations of positive polynomials on compact semialgebraic sets like the ball B_{ϱ_d}. We will back on this point below when we analyze the multidimensional case in Section 3.2.

3.1 The one-dimensional case

In the one-dimensional case of (3.5), the matrix $\mathbb{H}_k(m) = (m_{\iota+\iota'})_{0\le\iota,\ \iota'\le k}$ takes the form of a *Hankel* matrix $[m_{\iota+\iota'}]_{\iota,\iota'=1}^{k}$. This one-dimensional case is particularly simple because the closure of the cone of moments of positive measures on the real line, i.e.

$$M = \left\{ m\in\mathbb{R}^{2k+1} : \ m = \int_{\mathbb{R}} (1, t, \ldots, t^{2k})\, \mu(\mathrm{d}t) \ \text{ for a positive measure } \mu \text{ on } \mathbb{R} \right\} \tag{3.7}$$

is precisely the cone of those vectors $m \in \mathbb{R}^{2k+1}$ which make $\mathbb{H}_k(m)$ positive semidefinite. According to the usual convention, we use t instead of s in this 1-dimensional case. Although not every vector $m \in \mathbb{R}^{2k+1}$ satisfying this condition is a vector of moments, in the one dimensional case the coercivity of φ avoids any difficulty. The following lemma from [19, 45, 46] clarifies this point.

Lemma 3.1. *Let $\varphi(t) = \sum_{\iota=0}^{2k} c_\iota t^\iota$ be a one dimensional, coercive polynomial (i.e. $c_{2k} > 0$). Then, any solution m^* of the semidefinite program:*

$$(\mathscr{SDP}) \qquad \left\{ \begin{array}{ll} \textit{Minimize} & c\cdot m := \displaystyle\sum_{\iota=0}^{2k} c_\iota m_\iota \\ \textit{subject to} & \mathbb{H}_k(m) \succeq 0 \textit{ with } \ m_0 = 1 \textit{ and } \ m_1 = a, \end{array} \right.$$

is composed of the algebraic moments of a measure μ^* solving the following abstract optimization problem defined in measures:

$$\left. \begin{array}{ll} \textit{Minimize} & \langle \varphi, \mu \rangle := \displaystyle\int_{\mathbb{R}} \varphi(t)\, \mu(\mathrm{d}t) \\ \textit{subject to} & \displaystyle\int_{\mathbb{R}} t\, \mu(\mathrm{d}t) = a, \qquad \mu \in \mathcal{P}(\mathbb{R}) \end{array} \right\} \tag{3.8}$$

where $\mathcal{P}(\mathbb{R})$ stands for the family of all probability measures supported on the real line $\mathbb{R}$. The converse is also true, i.e. when μ^* solves (3.8), then its algebraic moments solve $(\mathscr{SDP})$.

On one hand, this fact certainly allows us to determine exact relaxations in the one dimensional case. On the other hand, it also has an important geometrical meaning in convex analysis. Since the polynomial φ is coercive on $\mathbb{R}$, every point of the graph of its convex envelope φ_c can be expressed as a convex combination of points on the graph of φ itself. By applying the classical Carathéodory's theorem of convex analysis we obtain the following formula:

$$(a, \varphi_c(a)) = \lambda_1(a_1, \varphi(a_1)) + \lambda_2(a_2, \varphi(a_2)) \tag{3.9}$$

with $\lambda_1 + \lambda_2 = 1$, $\lambda_1 \ge 0$, $\lambda_2 \ge 0$. It is remarkable that every optimal measure for (3.8) comes from the geometrical representation in (3.9). Thus, a probability measure $\overline{\mu}$ solves (3.8) if and only if it satisfies the equation:

$$(a, \varphi_c(a)) = \int_{\mathbb{R}} (t, \varphi(t))\, \overline{\mu}(\mathrm{d}t), \tag{3.10}$$

see [44, 48]. From this observation we can see that

$$\overline{\mu} = \lambda_1 \delta_{a_1} + \lambda_2 \delta_{a_2} \tag{3.11}$$

is a solution of (3.8), whenever the coefficients λ_i and the points a_i satisfy (3.9). Thus, we can use a set of optimal values m^* from the semidefinite program $(\mathscr{SDP})$ to determine the support and the weights of an optimal measure $\overline{\mu}$ solving (3.8). These facts can be used to prove the following result; later they will also be useful when applied into the multidimensional setting.

Proposition 3.2. *Assume that $n = 1$ and $\varphi_1(x,t) = \sum_{\iota=0}^{2k} g_\iota(x)t^\iota$ with $g_{2k}(\cdot) \geq \varepsilon$ for some $\varepsilon > 0$. Then*

$$\left.\begin{array}{ll} \textit{Minimize} & \Phi(u,m) = \displaystyle\int_0^1 \sum_{\iota=0}^{2k} g_\iota(x)m_\iota(x) + \varphi_0(u,x)\,\mathrm{d}x \\ \textit{subject to} & m_0(x) = 1, \quad u'(x) = m_1(x) \\ & \mathbb{H}_k(m(x)) \succeq 0 \textit{ for every } x \in [0,1] \\ & u(0) = u_{\mathrm{D}}(0), \quad u(1) = u_{\mathrm{D}}(1) \end{array}\right\} \tag{3.12}$$

has a solution and the minimum of (3.12) *is just* $\min(\mathscr{OCP}) = \min(\mathscr{RP})$. *Moreover, every solution m^* of* (3.12) *can be traced back to a particular optimal η^* for the corresponding one dimensional case in the formulation* ($\mathscr{RP}$), *in the sense that m^* is the vector function of the algebraic moments of the Young measure η^*, i.e. $m^*_\iota = \eta^* \bullet (1 \otimes s^\iota)$ for $\iota = 0, \ldots, 2k$.*

This result has been proved and exploited constructively in [8,48]. Notice that coercivity of φ_1 implies a finite support of the optimal measures of ($\mathscr{RP}$). We would like to remark here that in the one-dimensional case the relaxation of ($\mathscr{RP}$) takes the form of a convex optimal control problem like (3.12) which has a minimizer under coercivity assumptions. See [14,20,49–51,70].

An analogous assertion holds for element-wise constant η's. This fact suggests to formulate ($\mathscr{RP}_{H_k,d}$) (i.e. in ($\mathscr{RP}_{H,d}$) with $H = H_k$ from (3.1)) terms of moments. Let $\mathcal{T}_d$ be an equidistant partition of the interval $[0,1]$ with $d > 0$ a mesh size. Then the *approximate problem in terms of moments* looks as:

$$(\mathscr{MP}_d) \qquad \left\{\begin{array}{ll} \text{Minimize} & \hat{\Phi}(u,m) := \displaystyle\sum_{i=1}^{1/d} \sum_{\iota=0}^{2k} g_{i,\iota} m_{i,\iota} + \int_0^1 \varphi_0(x,u(x))\,\mathrm{d}x, \\ \text{subject to} & m_{i,0} = 1, \quad m_{i,1} = u'(x) \quad \text{for } x \in \big((i-1)d, id\big), \\ & \mathbb{H}_k(m_{i,0}, \cdots, m_{i,2k}) \succeq 0 \quad \text{for all } \ i = 1, \cdots, 1/d, \\ & u \in W^{1,p}(\Omega) \text{ element-wise affine on } \mathcal{T}_d, \\ & u(0) = u_{\mathrm{D}}(0), \quad u(1) = u_{\mathrm{D}}(1), \end{array}\right.$$

where the coefficients $g_{i,\iota}$ come from the expansion of the element-wise constant integrand $P_d\varphi_1$, i.e.

$$\big[P_d\varphi_1\big](x,s) = \sum_{\iota=0}^{2k} g_{i,\iota} s^\iota \qquad \text{for } x \in \big((i-1)d, id\big). \tag{3.13}$$

Thus we transform the problem ($\mathscr{RP}_{H_k,d}$) into a *semidefinite programming* problem. Depending whether $\varphi_0(x,\cdot)$ is linear, convex quadratic, or more general, more or less efficient computer codes are available for solving ($\mathscr{MP}_d$), e.g. the primal-dual interior point algorithm, generalized augmented-Lagrangian method, or a log-barrier method, respectively. See [6,7,28,45,48,69].

From the analysis of the one-dimensional case exposed above, the following equivalence clearly follows:

Proposition 3.3. *If $n = 1$ and* (3.13) *holds, then the problem* ($\mathscr{RP}_{H_k,d}$) *with H_k from* (3.1) *is equivalent to* ($\mathscr{MP}_d$) *in the sense that:*

(i) $\min(\mathscr{RP}_{H_k,d}) = \min(\mathscr{MP}_d)$, *and*

(ii) (u^*,η^*) *solves* ($\mathscr{RP}_{H_k,d}$) *if and only if* (u^*,m^*) *solves* ($\mathscr{MP}_d$), *where optimal η^* is related to optimal m^* through* (3.3), *which means here $m^*_\iota = \eta^* \bullet (1 \otimes s^\iota)$ for $\iota = 0, \cdots, 2k$.*

It is worth mentioning here that we can use algebraic tools for constructing finite supported measures from a finite set of their moments, so we can use the optimal vectors m^*_d from ($\mathscr{MP}_d$) to construct some η^*_d as a minimizer of ($\mathscr{RP}_{H_k,d}$). See [43,44].

3.2 The multi-dimensional case

The Multidimensional Moment Problem is still an open problem in pure mathematics, see [1, 5, 13, 26, 29, 68]. Nonetheless, important progress has been made in recent years as algebraists have been able to characterize positive polynomials defined in compact semi-algebraic sets [15–17, 22, 24, 39, 67]. This result has been applied to global optimization of polynomials and non-convex situations in optimization theory, see [19, 43–46, 48, 52]. We use this methodology here to transform $(\mathscr{RP}_{H_k,d})$ into a convex optimization problem in which the Young-measures ν_d are represented as moments-like vectorial functions m_d within a proper convex mathematical program. In particular we will follow J. B. Lasserre's approach on global optimization of polynomials developed in [25, 33–36, 38] to describe the convex envelope of n-dimensional coercive polynomials by semidefinite programming.

The case $n > 1$ is more complicated because the characterization of elements η from $\mathring{Y}^p_H(\Omega;\mathbb{R}^m)$ only can be attained in a limit sense, provided that the supports of the parameterized measures in the Young-measure representation of η lie in a compact semi-algebraic set. Thus, we assume that every parameterized measure ν_x is supported on the n-dimensional ball: $B_{\varrho_d} := \{s \in \mathbb{R}^n;\ s_1^2 + \cdots + s_n^2 \le \varrho_d^2\}$ where ϱ_d is chosen big enough. Due to the uniform coercivity of φ_1, the Weierstrass maximum principle (see equation (A.11) from Appendix) can be achieved inside a ball B_{ϱ_d} for ϱ_d sufficiently large. By proceeding in this way, we can apply recent results on the characterization of multidimensional moments without changing the original formulation of the problem.

It is enlightening for our approach to focus on the convex envelope of the integrand $\varphi_1(x, \nabla u(x))$ with respect to the gradient variables. Indeed, it has been observed by other authors, see [18, 23, 59, 63], that this kind of convexification allows us to obtain an exact convex relaxation of the original non-convex problem $(\mathscr{P})$. Thus, when we denote by $\mathrm{C}\varphi_1(x, \cdot)$ the *convex envelope* of $\varphi_1(x, \cdot)$ and consider the convex formulation

$$(\mathscr{CP}) \qquad \begin{cases} \text{Minimize} & \Phi_c(u) := \displaystyle\int_\Omega \mathrm{C}\varphi_1(x, \nabla u(x)) + \varphi_0(x, u(x))\,\mathrm{d}x, \\ \text{subject to} & u \in W^{1,p}(\Omega), \quad u|_{\partial\Omega} = u_\mathrm{D}, \end{cases}$$

we obtain an exact relaxation of $(\mathscr{P})$.

Proposition 3.4. (See [59].) *By assuming the coercivity of φ_1, we have the following results:*

(i) *Let u^* be a solution for $(\mathscr{CP})$. The couple (u^*, η^*) with $\eta^* \in Y^p_H(\Omega;\mathbb{R}^n)$ is a solution of $(\mathscr{RP})$ provided that $\mathrm{C}\varphi_1(x, \nabla u^*(x)) = [\varphi_1 \bullet \eta^*](x)$ and $[\mathrm{Id} \bullet \eta^*](x) = \nabla u^*(x)$ for almost every $x \in \Omega$.*

(ii) *Reciprocally, if (u^*, η^*) solves $(\mathscr{RP})$, the function u^* is a solution for $(\mathscr{CP})$ and (u^*, η^*) itself solves the equations $\mathrm{C}\varphi_1(x, \nabla u^*(x)) = [\varphi_1 \bullet \eta^*](x)$ and $[\mathrm{Id} \bullet \eta^*](x) = \nabla u^*(x)$ for almost every $x \in \Omega$.*

Thus, we are confronted with the discretized convexified problem:

$$(\mathscr{CP}_d) \qquad \begin{cases} \text{Minimize} & \displaystyle\int_\Omega \mathrm{C}[P_d\varphi_1](x, \nabla u(x)) + \varphi_0(x, u(x))\,\mathrm{d}x, \\ \text{subject to} & u \in W^{1,p}(\Omega) \text{ element-wise affine on } \mathcal{T}_d, \quad u|_{\partial\Omega} = u_\mathrm{D}. \end{cases}$$

Let us emphasize that only one-sided inequality $\mathrm{C}[P_d\varphi_1] \ge P_d(\mathrm{C}\varphi_1)$ is at disposal and, using Φ_c from $(\mathscr{CP})$ "blindly" also for $(\mathscr{CP}_d)$, might lead to under-relaxation. We will prove next that $(\mathscr{CP}_d)$ is equivalent to $(\mathscr{RP}_{H_k,d})$ under proper coercivity assumptions. For this, we need the following:

Lemma 3.5. (See [18].) *Given an n-dimensional polynomial $\varphi : \mathbb{R}^n \to \mathbb{R}$ satisfying $c_1|s|^p \le \varphi(s) \le c_2(1 + |s|^p)$ for every $s \in \mathbb{R}^n$, with $p > 1$ and positive constants c_1 and c_2,* *we can determine its convex envelope at a fixed point $a \in \mathbb{R}^n$ as:*

$$\mathrm{C}\varphi(a) = \begin{cases} \textit{Minimize} & \langle \varphi, \mu \rangle := \int_{\mathbb{R}^n} \varphi(s)\, \mu(\mathrm{d}s) \\ \textit{subject to} & \int_{\mathbb{R}^n} s\, \mu(\mathrm{d}s) = a, \qquad \mu \in \mathcal{P}(\mathbb{R}^n), \end{cases} \tag{3.14}$$

where, as above, $\mathcal{P}(\mathbb{R}^n)$ is the set of all Borel regular probability measures supported on $\mathbb{R}^n$.

Remark 3.6. At the boundary point $(a, \mathrm{C}\varphi(a))$ there exists a supporting hyperplane for the convex set $\mathrm{Epi}(\mathrm{C}\varphi)$. Such hyperplane can be defined by a linear-affine function $L_a : \mathbb{R}^n \to \mathbb{R}$ which satisfies $L_a \le \mathrm{C}\varphi \le \varphi$ and $L_a(a) = \varphi(a)$. Thus, we easily characterize the optimal measures μ^* for (3.14) as the set of probability measures supported on $\mathbb{F}_a = \{s \in \mathbb{R}^n : L_a(s) = \varphi(s)\}$ satisfying $a = \int_{\mathbb{R}^n} s\, \mu^*(\mathrm{d}s)$. Hence, we find that a necessary condition for μ^* to be optimal in (3.14) is that μ^* be supported in $\{s \in \mathbb{R}^n : \mathrm{C}\varphi(s) = \varphi(s)\}$. Let us emphasize that this condition is not sufficient.

Proposition 3.7. *By assuming the coercivity requirements on φ_1, the discrete problem $(\mathscr{CP}_d)$ is equivalent to $(\mathscr{RP}_{H_k,d})$ in the following sense:*

(i) *Let u_d^* be a solution for $(\mathscr{CP}_d)$. The couple (u_d^*, η_d^*) with $\eta_d^* \in P_d^* Y_H^p(\Omega; \mathbb{R}^n)$ is a solution of $(\mathscr{RP}_{H_k,d})$ provided that*

$$\mathrm{C}[P_d\varphi_1](x, \nabla u_d^*(x)) = \big[\varphi_1 \bullet \eta_d^*\big](x) \quad \textit{and} \quad [\mathrm{Id} \bullet \eta_d^*](x) = \nabla u_d^*(x) \tag{3.15}$$

for a.a. $x \in \Omega$; in fact, the functions involved in (3.15) are constant inside every triangle $S \in \mathcal{T}_d$.

(ii) *Conversely, if (u_d^*, η_d^*) solves $(\mathscr{RP}_{H_k,d})$, the function u_d^* is a solution for $(\mathscr{CP}_d)$ and (u_d^*, η_d^*) itself solves (3.15) for a.a. $x \in \Omega$.*

Proof. Let (u_d, η_d) be a solution for the relaxed problem $(\mathscr{RP}_{H_k,d})$, then $[\mathrm{Id} \bullet \eta_d](x) = \nabla u_d(x)$ for every $S \in \mathcal{T}_d$. Since u_d^* is optimal for $(\mathscr{CP}_d)$, we have:

$$\int_\Omega \mathrm{C}[P_d\varphi_1](x, \nabla u_d(x)) + \varphi_0(x, u_d(x))\, \mathrm{d}x \ge \int_\Omega \mathrm{C}[P_d\varphi_1](x, \nabla u_d^*(x)) + \varphi_0(x, u_d^*(x))\, \mathrm{d}x.$$

By using Lemma 3.5, we can see that

$$\int_\Omega [\varphi_1 \bullet \eta_d](x) + \varphi_0(x, u_d(x))\, \mathrm{d}x \ge \int_\Omega \mathrm{C}[P_d\varphi_1](x, \nabla u_d(x)) + \varphi_0(x, u_d(x))\, \mathrm{d}x \tag{3.16}$$

and finally we have

$$\int_\Omega [\varphi_1 \bullet \eta_d](x) + \varphi_0(x, u_d(x))\ \mathrm{d}x \ge \int_\Omega [\varphi_1 \bullet \eta_d^*](x) + \varphi_0(x, u_d^*(x))\, \mathrm{d}x \tag{3.17}$$

because of the assumptions in (i) on u_d^* and η_d^*. Thus, we have shown that (u_d^*, η_d^*) is the solution to the problem $(\mathscr{RP}_{H_k,d})$, and the point (i) has thus been proved.

As to the point (ii), let ν_d^* be the (element-wise homogeneous) Young measure induced by the convex envelope of $P_d\varphi_1$ according to Lemma 3.5, i.e.

$$\mathrm{C}[P_d\varphi_1](x, \nabla u_d^*(x)) = \int_{\mathbb{R}^n} [P_d\varphi_1](x, s) \big[\nu_d^*\big]_x(\mathrm{d}s) \quad \text{and} \quad \int_{\mathbb{R}^n} s \big[\nu_d^*\big]_x(\mathrm{d}s) = \nabla u_d^*(x)$$

for a.a. $x \in \Omega$. As (u_d^*, ν_d^*) is admissible and (u_d^*, η_d^*) is optimal for the problem $(\mathscr{RP}_{H_k,d})$, we have

$$\int_\Omega \int_{\mathbb{R}^n} [P_d\varphi_1](x, s) \big[\nu_d^*\big]_x(\mathrm{d}s) + \varphi_0(x, u_d^*(x))\, \mathrm{d}x \ge \int_\Omega [\varphi_1 \bullet \eta_d^*](x) + \varphi_0(x, u_d^*(x))\, \mathrm{d}x$$

and then

$$\int_\Omega \mathrm{C}[P_d\varphi_1](x,\nabla u_d^*(x))\,\mathrm{d}x = \int_\Omega\int_{\mathbb{R}^n}[P_d\varphi_1](x,s)\big[\nu_d^*\big]_x(\mathrm{d}s)\,\mathrm{d}x$$
$$\geq \int_\Omega [\varphi_1\bullet\eta_d^*](x)\,\mathrm{d}x \geq \int_\Omega \mathrm{C}[P_d\varphi_1](x,\nabla u_d^*(x))\,\mathrm{d}x. \qquad (3.18)$$

Hence,

$$\int_\Omega [\varphi_1\bullet\eta_d^*](x)\,\mathrm{d}x = \int_\Omega \mathrm{C}[P_d\varphi_1](x,\nabla u_d^*(x))\,\mathrm{d}x. \qquad (3.19)$$

By applying Lemma 3.5 again, we can claim that $\mathrm{C}[P_d\varphi_1](x,\nabla u_d^*(x)) \leq [\varphi_1\bullet\eta_d^*](x)$ for every $x\in S\in\mathcal{T}_d$. Since the integrals in (3.19) can be expressed as a finite sum on the members of $\mathcal{T}_d$, we have: $\sum_{S\in\mathcal{T}_d}\int_S[\varphi_1\bullet\eta_d^*](x)\mathrm{d}x = \sum_{S\in\mathcal{T}_d}\int_S \mathrm{C}[P_d\varphi_1](x,\nabla u_d^*(x))\,\mathrm{d}x$. Therefore, we can conclude that:

$$[\varphi_1\bullet\eta_d^*](x) = \mathrm{C}[P_d\varphi_1](x,\nabla u_d^*(x)) = \int_{\mathbb{R}^n}[P_d\varphi_1](x,s)\big[\nu_d^*\big]_x(\mathrm{d}s) \qquad (3.20)$$

for every $x\in S\in\mathcal{T}_d$. In this way, we can see that ν_d^* corresponds to η_d^* in the sense that (3.15)) holds for every $x\in S\in\mathcal{T}_d$. To see that u_d^* solves the convexified problem $(\mathscr{CP}_d)$, we take an admissible u_d for $(\mathscr{CP}_d)$ and we take ν_d as the Young measure generating the convex envelope of $P_d\varphi_1$ according to Lemma 3.5. Therefore, we have:

$$\int_\Omega \mathrm{C}[P_d\varphi_1](x,\nabla u_d(x)) + \varphi_0(x,u_d(x))\,\mathrm{d}x = \int_\Omega [\varphi_1\bullet\eta_d](x) + \varphi_0(x,u_d(x))\,\mathrm{d}x$$
$$\geq \int_\Omega [\varphi_1\bullet\eta_d^*](x) + \varphi_0(x,u_d^*(x))\,\mathrm{d}x$$
$$\geq \int_\Omega [[P_d\varphi_1]\bullet\eta_d^*](x) + \varphi_0(x,u_d^*(x))\,\mathrm{d}x = \int_\Omega \mathrm{C}[P_d\varphi_1](x,\nabla u_d^*(x)) + \varphi_0(x,u_d^*(x))\,\mathrm{d}x,$$

having employed that $\varphi_1\bullet\eta_d^* = \varphi_1\bullet[P_d^*\eta_d^*] = [P_d\varphi_1]\bullet\eta_d^*$ since $P_dP_d = P_d$. In this way we conclude that u_d^* is optimal for $(\mathscr{CP}_d)$. □

From Proposition 3.7, the importance of the convex envelope of the polynomial φ_1 into the relaxed formulation $(\mathscr{RP}_{H_k,d})$ becomes clear. From Lemma 3.5, it is also clear that the convex envelope of a multidimensional polynomial φ_1 can be evaluated by a particular optimization problem defined in probability measures. Thus, we can apply the moments technique to the relaxed formulation $(\mathscr{RP}_{H_k,d})$, taking into account that we must obtain at last the convex envelope of the polynomial φ_1 to be certain that we get an exact relaxation of the original problem. See [47]. The following result supports this approach.

Proposition 3.8. *Consider the convex, relaxed and discretized problem:*

$$(\mathscr{MP}_{d,\kappa})\qquad \begin{cases} \textit{Minimize} & \hat{\Phi}(u,m) := \displaystyle\sum_{S\in\mathcal{T}_d}\sum_{0\leq|\iota|\leq 2k}\phi_{S,\iota}m_{S,\iota} + \int_\Omega \varphi_0(x,u(x))\,\mathrm{d}x,\\ \textit{subject to} & m_{S,e_i} = \dfrac{\partial u}{\partial x_i} \ \textit{ on } S\in\mathcal{T}_d,\quad i=1,\cdots,n,\\ & \left.\begin{array}{l} m_{S,0,\cdots,0} = 1,\\ \mathbb{H}_\kappa\big(\{m_{S,\iota}\}_{|\iota|\leq 2\kappa}\big)\succeq 0,\\ \mathbb{L}_\kappa\big(\{m_{S,\iota}\}_{|\iota|\leq 2(\kappa-1)}\big)\succeq 0,\end{array}\right\}\ \textit{for all}\ \ S\in\mathcal{T}_d,\\ & u\in W^{1,p}(\Omega)\ \textit{element-wise affine of}\ \mathcal{T}_d,\quad u|_{\partial\Omega} = u_{\mathrm{D}}, \end{cases}$$

where $e_i = (0, \cdots, 0, 1, 0, \cdots, 0) \in \mathbb{R}^n$ is the vector with 1 on the i-th position and where, similarly as in $(\mathscr{MP}_{d,\kappa})$, the coefficients $\phi_{S,\iota}$ come from the expansion of the element-wise constant integrand $P_d\varphi_1$, i.e. $[P_d\varphi_1](x,s) = \sum_{0\le|\iota|\le 2k} \phi_{S,\iota} s_1^{\iota_1}\cdots s_n^{\iota_n}$ for $x \in S \in \mathcal{T}_d$. Then:

(i) *$(\mathscr{MP}_{d,\kappa})$ has a solution for every $d > 0$, k and κ with $\kappa \ge k$.*

(ii) *The solution of $(\mathscr{MP}_{d,\kappa})$ provides a lower bound for $(\mathscr{RP}_{H_k,d})$ and $\min(\mathscr{MP}_{d,\kappa}) \nearrow \inf(\mathscr{RP}_{H_k,d})$ when $\kappa \to \infty$.*

(iii) *When (u^*, m^*) solves $(\mathscr{MP}_{d,\kappa})$ and there exists $\eta^* \in P_d^* Y_H^p(\Omega;\mathbb{R}^n)$ such that $m_\iota^* = \eta^* \bullet (1 \otimes s_1^{\iota_1}\cdots s_n^{\iota_n})$ for $|\iota| := \iota_1 + \cdots + \iota_n \le 2\kappa$, then η^* solves $(\mathscr{RP}_{H_k,d})$.*

Proof. We use again some tools from convex optimization and convex analysis to specify the convex envelope of a coercive polynomial $\varphi : \mathbb{R}^n \to \mathbb{R}$ at the point $a = (a_1, \ldots, a_n)$. The coefficients of the polynomial $\varphi(x) = \sum_{2\le|\iota|\le 2k} c_{\iota_1,\ldots\iota_n} x_1^{\iota_1}\ldots x_n^{\iota_n}$ allow us to define the following semidefinite program:

$$(\mathscr{SDP}_\kappa) \qquad \left\{ \begin{array}{ll} \text{Minimize} & c\cdot m := \displaystyle\sum_{2\le|\iota|\le 2k} c_{\iota_1,\ldots\iota_n} m_{\iota_1,\ldots\iota_n} \\ \text{subject to} & \mathbb{H}_\kappa(m) \succeq 0 \ \text{ and } \mathbb{L}_\kappa(m) \succeq 0 \\ & \text{with } m_{0,\ldots,0} = 1 \text{ and } m_{e_1} = a_1, \ldots, m_{e_n} = a_n. \end{array} \right.$$

Because of $\kappa \ge k$, the problems $(\mathscr{SDP}_\kappa)$ make a sequence of convex programs indexed by κ. Notice that we omit the linear part of φ as it does not affect the analysis of the convex envelope of the polynomial φ. Since $\mathbb{H}_\kappa(m) \succeq 0$ and $\mathbb{L}_\kappa(m) \succeq 0$ are necessary conditions for m to be a valid vector of multidimensional moments of a multidimensional measure supported on the n-dimensional ball B_{ϱ_d}, the value of $(\mathscr{SDP}_\kappa)$ is a lower bound for the value $\mathrm{C}\varphi(a)$ if $k \to \infty$. We have used here the right-hand side of (3.14). This program $(\mathscr{SDP}_\kappa)$ has a finite optimal value, and this optimal value of $(\mathscr{SDP}_\kappa)$ defines a nondecreasing sequence of lower bounds of $\mathrm{C}\varphi(a)$. We will show that the optimal value of $(\mathscr{SDP}_\kappa)$ converges to $\mathrm{C}\varphi(a)$ when $\kappa \to \infty$ by following J. B. Lasserre's proposal for global optimization of polynomials stated in his seminal paper [33].

Let L_a be the affine function defining the supporting hyperplane of the convex set $\mathrm{Epi}(\mathrm{C}\varphi) \subset \mathbb{R}^{n+1}$ at the point $(a, \mathrm{C}\varphi(a))$. See Remark 3.6. Given $\varepsilon > 0$, we have $\varphi(x) - L_a(x) + \varepsilon > 0$ for every $x \in B_{\varrho_d}$. Since the ball B_{ϱ_d} is a semialgebraic compact set, we can express the positive polynomial $\varphi(x) - La(x) + \varepsilon$ in B_{ϱ_d} as:

$$\varphi(x) - L_a(x) + \varepsilon = \sum_{j=1}^{J} q_j^2(x) + (\varrho_d^2 - x_1^2 - \cdots - x_n^2) \sum_{j'=1}^{J'} q_{j'}^2 \tag{3.21}$$

where q_j and $q_{j'}$ are n-dimensional polynomials whose degrees can not be determined in advance, i.e. they depend on ε. See [60]. If we take 2κ as the degree of the polynomial at the right-hand side of (3.21), then the quadratic representation of (3.21) gives a feasible solution of the dual of the semidefinite program $(\mathscr{SDP}_\kappa)$. See [33, 47].

The dual form of the semidefinite program $(\mathscr{SDP}_\kappa)$ is

$$\left. \begin{array}{ll} \text{Maximize} & -\gamma_{0,\ldots,0} - 2a_1\gamma_{e_1} - \cdots - 2a_n\gamma_{e_n} \\ & \qquad\qquad -\varrho_d^2\lambda_{0,\ldots,0} - 2a_1\varrho_d^2\lambda_{e_1} - \cdots - 2a_n\varrho_d^2\lambda_{e_n} \\ \text{subject to} & \langle \Gamma, A_{\iota_1,\ldots\iota_n}\rangle_\kappa + \langle \Lambda, \widetilde{A}_{\iota_1,\ldots\iota_n}\rangle_{\kappa-1} = c_{\iota_1,\ldots\iota_n} \quad \forall \iota : 2 \le |\iota| \le 2k, \\ & \Gamma \succeq 0 \text{ and } \Lambda \succeq 0, \end{array} \right\} \tag{3.22}$$

cf. e.g. [4], [33]. Matrices Γ and Λ have the forms $\mathbb{H}_\kappa(\gamma)$ and $\mathbb{H}_{\kappa-1}(\lambda)$ respectively, $\langle\cdot,\cdot\rangle$ stands for the Frobenius product, and the matrix $A_{\iota_1\ldots\iota_n}$ is full of zeros but its $\iota_1\ldots\iota_n$ entries which are ones. To make explicit the form of every matrix $A_{\iota_1\ldots\iota_n}$ we take the quadratic form $ZZ^\top$ where $Z = (x_1^{\iota_1}\ldots x_n^{\iota_n})_{0\le\iota_i+\cdots+\iota_n\le\kappa}$. The $\iota_1\ldots\iota_n$ entries in $A_{\iota_1\ldots\iota_n}$ are the entries of the function $x_1^{\iota_1}\dots x_n^{\iota_n}$ in $ZZ^\top$, $A_{\iota_1\dots\iota_n}$ is the same size as $ZZ^\top$. The form of the matrix $\widetilde{A}_{\iota_1\dots\iota_n}$ comes from the quadratic form $\left(\varrho_d^2 - x_1^2 - \cdots - x_n^2\right) WW^\top$ where $W = (x_1^{\iota_1}\dots x_n^{\iota_n})_{0\le\iota_1+\cdots+\iota_n\le(\kappa-1)}$. Each entry in $\widetilde{A}_{\iota_1\dots\iota_n}$ has the coefficient of $x_1^{\iota_1}\dots x_n^{\iota_n}$ in the respective position in $\left(\varrho_d^2 - x_1^2 - \cdots - x_n^2\right) WW^\top$.

The primal semidefinite program $(\mathscr{SDP}_\kappa)$ is strictly feasible as we can always find a set of moments m induced by a continuously distributed probability measure in B_{ϱ_d} whose first marginal moments are the values $a_1,\dots,a_n$. See [17]. In this way the moments m define positive definite matrices $\mathbb{H}_\kappa(m)$ and $\mathbb{L}_\kappa(m)$. As usual in convex optimization, feasible solutions of the dual program (3.22) provide a lower bound for the primal program $(\mathscr{SDP}_\kappa)$. So, as proposed in [33], we take a couple of dual variables Γ and Λ from the coefficients of the polynomials q_j and $q_{j'}$ in $(\mathscr{QF})$. This is, $\Gamma = \sum_{j=1}^{J} q_j \cdot q_j^t$ and $\Lambda = \sum_{j'=1}^{J'} q_{j'} \cdot q_{j'}^t$ which implies that matrices Γ and Λ must be positive semidefinite. From the representation (3.21) of the positive polynomial $\varphi(x) - L_a(x) + \varepsilon$ we can see that:

$$-\gamma_{0,\dots,0} - 2a_1\gamma_{e_1} - \cdots - 2a_n\gamma_{e_n} - \varrho_d^2\lambda_{0,\dots,0} - 2a_1\varrho_d^2\lambda_{e_1} - \cdots - 2a_n\varrho_d^2\lambda_{e_n} = L_a(a) - \varepsilon.$$

Therefore, the dual variables Γ and Λ determine the following lower bound for $(\mathscr{SDP}_\kappa)$:

$$-\gamma_{0,\dots,0} - 2a_1\gamma_{e_1} - \cdots - 2a_n\gamma_{e_n} - \varrho_d^2\lambda_{0,\dots,0} - 2a_1\varrho_d^2\lambda_{e_1} - \cdots - 2a_n\varrho_d^2\lambda_{e_n} = \mathrm{C}\varphi(a) - \varepsilon.$$

As the optimal value in $(\mathscr{SDP}_\kappa)$ is a lower bound for $\mathrm{C}\varphi(a)$, we have:

$$\forall\varepsilon>0\ \exists\kappa_0\in\mathbb{N}\ \forall\kappa\ge\kappa_0: \qquad \mathrm{C}\varphi(a) \ge \min(\mathscr{SDP}_\kappa) \ge \mathrm{C}\varphi(a) - \varepsilon$$

where the optimal value $\min(\mathscr{SDP}_\kappa)$ is finite because the primal program $(\mathscr{SDP}_\kappa)$ is strictly feasible.

The present analysis of the convex envelope of a n-dimensional polynomial proves items (i)-(iii) of Proposition 3.8 when applied on the coercive polynomial φ_1 given in $(\mathscr{RP}_{H_k,d})$ provided that φ_0 is convex in u. $\square$

3.3 Remarks

Let us end this section with some remarks that outline various advanced numerical strategies and various generalizations.

Remark 3.9. One can think about a linearization of the φ_0-term in the relaxed problem $(\mathscr{RP}_{H_k,d})$ and then, starting from some u^0, consider the iterative process based for $k = 1,2,\dots$ on the solution $(u_d^{(k)},\eta_d^{(k)})$ to the problem

$$\begin{aligned}
&\text{Minimize} && \Psi(u_d^{(k-1)};u,\eta) := \int_\Omega \big[\varphi_1 \bullet \eta\big](x) + [\varphi_0]'_u(x,u_d^{(k-1)})(u-u_d^{(k-1)})\,\mathrm{d}x,\\
&\text{subject to} && [\mathrm{Id}\bullet\eta](x) = \nabla u(x) \quad \text{for a.a. } x\in\Omega,\\
& && u\in W^{1,p}(\Omega) \ \text{ element-wise affine on } \mathcal{T}_d, \qquad u|_{\partial\Omega} = u_{\mathrm{D}},\\
& && \eta\in\mathscr{Y}^p_{H_k}(\Omega;\mathbb{R}^n) \text{ element-wise constant on } \mathcal{T}_d\,,
\end{aligned}$$

where $u_d^{(k-1)}$ is known from the previous iteration. This SLP (=*sequential linear programming*) strategy was proposed in [3]. In some qualified cases (e.g. in the benchmark problem in Section 4.1 below), it converges by the Banach fixed-point argument, see [3] for details. Here it may open a possibility, after applying the methods of moments from Section 3, to use efficient semidefinite programming algorithms with linear cost functional even if $\varphi_1(x,\cdot)$ is nonlinear and nonconvex. For an SQP (=*sequential quadratic programming*) strategy, which allows for usage of still relatively efficient semidefinite programming algorithms with quadratic cost functional even if $\varphi_0(x,\cdot)$ is non-quadratic, see [41].

Remark 3.10. By solving relaxation ($\mathscr{MP}_{d,\kappa}$) with increasing values of κ, we eventually attain an exact solution or at least a good approximation of the exact moments values. See [33]. By doing so, we obtain a set of parameterized multidimensional moments $m^*(S)$ for every $S \in \mathcal{T}_d$. With these optimal moments we can calculate the optimal Young measure on every $S \in \mathcal{T}_d$. This procedure exploits the marginal algebraic moments and the convex hull properties of the non-convex potential φ_1. See [47] for further details of this implementation.

Remark 3.11. Some semidefinite programming algorithms yield the Lagrange multipliers $\lambda_{d,\kappa}$ to the first constraint in ($\mathscr{RP}_{d,\kappa}$), i.e. $(m_{S,\iota})_{|\iota|=1} = \nabla u$. This is now called an active-set strategy algorithm, developed originally in [12] and latter used e.g. in [3, 31, 32, 41]. In this way, we would get an upper estimate for $\min(\mathscr{RP}_{H_k,d})$ which would complete the previously obtained lower estimate $\min(\mathscr{MP}_{d,\kappa})$.

Remark 3.12. Considering $\kappa \geq k$ and only the part of this solution, namely the moments $m_{S,\iota}$ with $|\iota| \leq 2k$, we can define $\eta_\kappa \in H^*$ by (3.4) where $m_\iota(x) = m_{S,\iota}$ for $x \in S \in \mathcal{T}_d$. Unfortunately, except special cases as in Sect. 4.1, such η_κ need not belong to $\mathcal{Y}^p_{H_k}(\Omega;\mathbb{R}^n)$ even in a limit for $\kappa \to \infty$, contrary to the one-dimensional case.

Remark 3.13. In principle, in contrast to the additively coupled problem ($\mathscr{P}$), we could consider more general problems involving the generally coupled functional $\Phi(u) := \int_\Omega \varphi(x, u(x), \nabla u(x))\,\mathrm{d}x$. Our results allow relatively easily for an extension to the case $\varphi(x, u, s) = \sum_{|\iota|\leq 2k} g_\iota(x, u(x)) s_1^{\iota_1} \cdots s_n^{\iota_n}$. Then the φ_0-term in ($\mathscr{MP}_{d,\kappa}$) would be out but the coefficients $\phi_{i,\iota}$ and $\phi_{S,\iota}$ would depend on u, which would turn ($\mathscr{MP}_{d,\kappa}$) into general nonconvex positive-semidefinite mathematical programs.

4 ILLUSTRATIVE EXAMPLE AND NUMERICAL RESULTS

4.1 Benchmark problem: two-dimensional broken extremal

We illustrate the proposed approach on a benchmark problem used in [3, 41], namely the Tartar's broken-extremal problem [53] modified for the two-dimensional case. See [9, 10] or [2, Sect.3.5] or [11, Sect.8]. Let us consider the square $\Omega := (0, K)^n$ with $n = 2$ and size $K > 0$. For $x \in \Omega$, $s \in \mathbb{R}^2$, and $u \in \mathbb{R}$, we define

$$\varphi_1(x, s) := |s - a|^2 |s + a|^2, \tag{4.1a}$$

$$\varphi_0(x, u) := \big(u - g(a \cdot x)\big)^2 \quad \text{with} \tag{4.1b}$$

$$g(\xi) := -\frac{3}{128}\left(\xi - \frac{1}{2}\right)^5 - \frac{1}{3}\left(\xi - \frac{1}{2}\right)^3, \tag{4.1c}$$

where $a = (\sqrt{3}/2, 1/2)$. Note that (4.1a) is coercive for $p = 4$ and that $\varphi_1(x, \cdot) \in \Pi_4(\mathbb{R}^2)$. Hence, Section 3 applies with $k = 2 = n$ and note that:

$$\varphi_1(x, s_1, s_2) := s_1^4 + s_2^4 + 2s_1^2 s_2^2 - s_1^2 + s_2^2 - 2\sqrt{3}s_1 s_2 + 1. \tag{4.2}$$

Then, according to [53], the relaxed problem ($\mathscr{RP}$) has the unique solution:

$$u(x) = \begin{cases} g(a \cdot x) & \text{for } a \cdot x \in (0, 1/2), \\ \dfrac{(a \cdot x - 1/2)^3}{24} + (a \cdot x - 1/2) & \text{for } a \cdot x \in (1/2, \sqrt{2}), \end{cases} \tag{4.3a}$$

$$\nu_x = \begin{cases} \dfrac{1 - a \cdot \nabla u(x)}{2}\,\delta_{-a} + \dfrac{1 + a \cdot \nabla u(x)}{2}\,\delta_a & \text{for } a \cdot x \in (0, 1/2), \\ \delta_{\nabla u(x)} & \text{for } a \cdot x \in (1/2, \sqrt{2}), \end{cases} \tag{4.3b}$$

provided the boundary data are chosen as $u_{\rm D} := u|_{\partial\Omega}$ with u just from (4.3a).

We use the regular triangulation $\mathcal{T}_d$ shown in Figure 1. For this problem, we have implemented the convex mathematical program corresponding to the optimization problem $(\mathscr{MP}_{d,\kappa})$ in function of (u, m) variables. Here, we represent the admissible function u by using a finite element linear interpolation basis, defined by the finite element mesh shown in Figure 1. With the optimal vectors m^* so obtained, we first check whether they actually represent the algebraic moments of a two dimensional probability measure. If some of them do not represent a probability measure, we increase κ and try again. Usually, for coercive polynomials, we obtain a valid set of moments in a few steps, just taking κ a bit greater than the degree of φ_1. In our case, we have stopped in $\kappa = 3$, cf. also the experiments below summarized in Table 1 below.

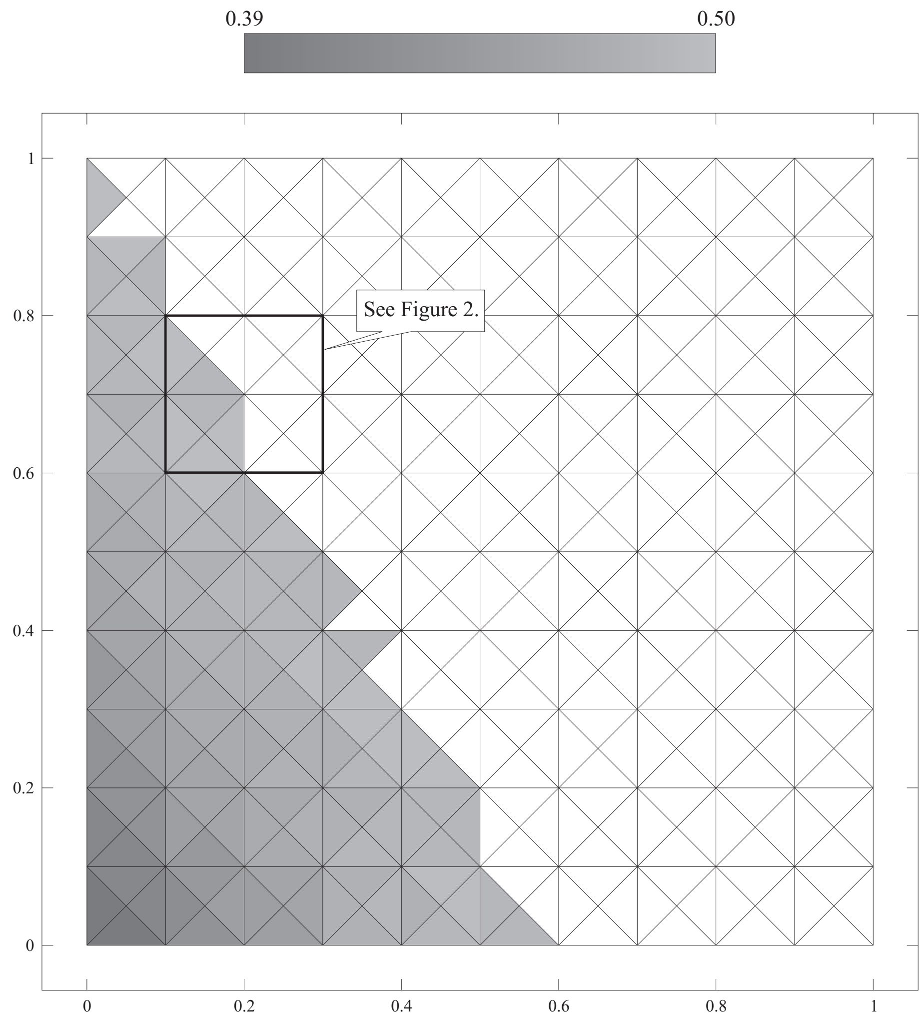

Figure 1. *A (10×10)-square finite-element mesh (composed from 400 elements) and a grey-scaled image of numerical values of* λ_1 *(with* $\lambda_1 = 1$ *on the white region).*

Afterwards, we obtain a set of bi-dimensional moments, and we use the marginal moments on the axis x and y to construct the probability measure that they represent. Let us notice that this procedure is not possible in general cases, and we succeed here because we have obtained moments which represent the probability measures defining the convex hull of the nonconvex potential φ_1. This is a nontrivial application of Carathéodory's theorem. See [47] for a detailed description of the algorithms that we use to calculate the optimal Young measure from the optimal moments involved in $(\mathscr{MP}_{d,\kappa})$.

The Young measure is constant on each element and, here, on each element, we have

6 parameters describing the corresponding probability measure

$$\lambda_1\delta_{a_1} + \lambda_2\delta_{a_2} \qquad \text{with} \qquad \lambda_1+\lambda_2 = 1, \quad \lambda_1, \lambda_2 \geq 0, \quad a_1, a_2 \in \mathbb{R}^2. \tag{4.4}$$

Figure 1 shows the finite element mesh and a grey scaled image of the numerical value of λ_1 for each element (see Figure 2 for a detail of the 6 values for the elements in the highlighted square of the mesh).

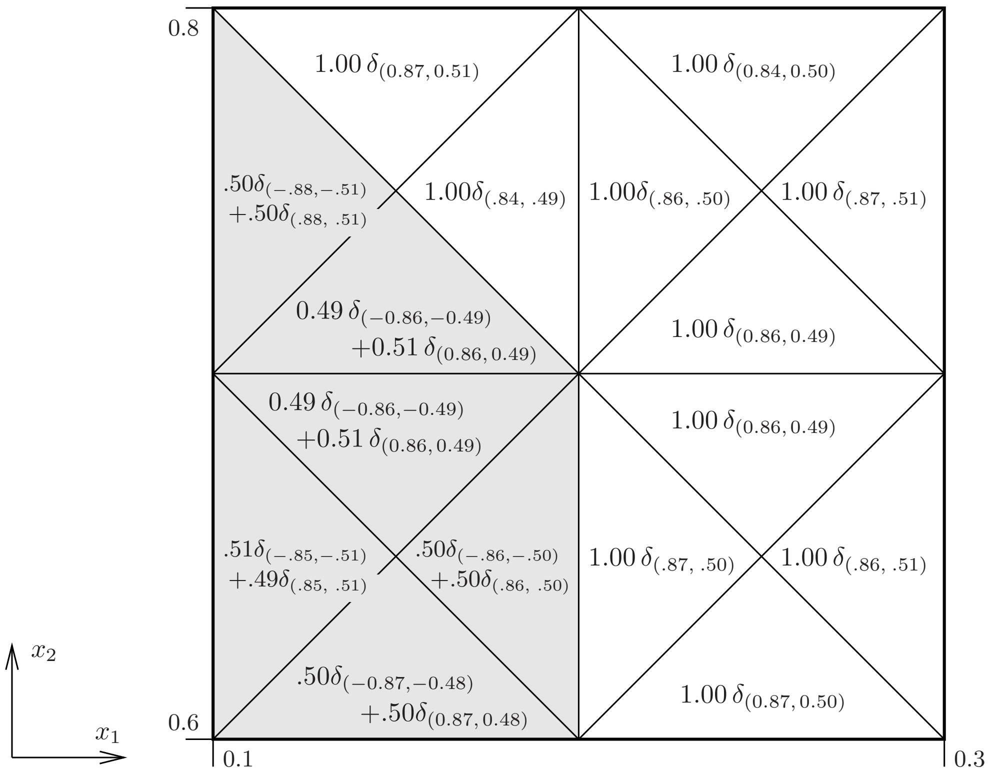

Figure 2. *Detail of a square selected as depicted on Figure 1 and of the reconstructed optimal parameterized measure ν_d^*, composed either from one or from two Diracs, presented with 2-digit accuracy.*

The optimal surface (i.e. the graph of u_d^*) shown in Figure 3, has been calculated by using the corresponding coefficients of a first-order spline basis, defined over the finite-element mesh shown in Figure 1.

Calculations have been done by implementing the model in a Matlab code, where optimization routines have been linked to an Ampl interface, see [47] for a deeper description of all the minor details of this implementation.

4.2 Comments on the role of κ in the relaxation

It is important that convergence in κ in the formulation $(\mathscr{MP}_{d,\kappa})$ is usually very fast and is attained after a few steps in κ. We must emphasize that every increase in κ amounts to an increase of the number of design variables in $(\mathscr{MP}_{d,\kappa})$ by twice. Recently, other authors have determined conditions that guarantee the final convergence in κ by using the rank of the moments matrices in $(\mathscr{MP}_{d,\kappa})$, see [37].

Example 4.1. Here we solve the non-convex variational problem in $(\mathscr{P})$ but taking φ_0

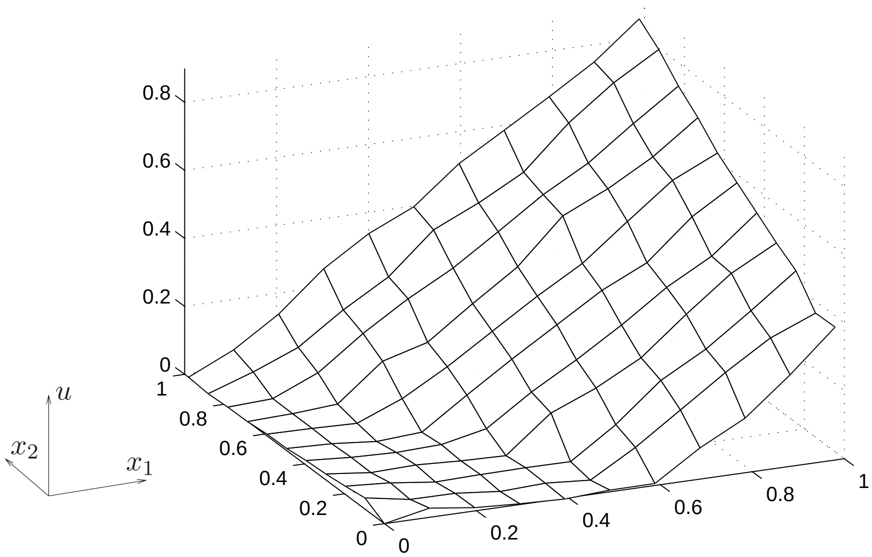

Figure 3. *Optimal u_d^* calculated on the finite-element mesh from Figure 1 and displayed here by a spline interpolation.*

and φ_1 as follows:

$$\varphi_1(x, s_1, s_2) := 4 + s_2^4 + s_1^4 + 2s_1^2 s_2^2 - 8s_1 s_2 \quad (4.5a)$$

$$\varphi_0(x, u) := (u - (x_1 + x_2))^2 \quad (4.5b)$$

$$u_{\mathrm{D}} := u|_{\partial\Omega} = 0. \quad (4.5c)$$

Example 4.2. Here we use the same function φ_1 and the same boundary condition as in (4.5a,c) but φ_0 changes into

$$\varphi_0(x, u) := u^2. \quad (4.6)$$

Example 4.3. Finally we use the same functions φ_1 and φ_0 as in (4.5a,b) but we take a different boundary condition

$$u_{\mathrm{D}} := u|_{\partial\Omega} = 1. \quad (4.7)$$

The optimal value of the functional against κ for a coarse (3×3)-square mesh is shown on Table 1. Small deviation from the monotone convergence claimed in Proposition 3.8, which occurs in the case of Example 4.2, is due to numerical errors. On the other hand, we can observe that the value of $\hat{\Phi}(\cdot, \cdot)$ decreases for a bigger value of κ.

All these examples show the good performance of the method and the faster convergence of the formulation shown in this article with the appropriate formulation by the method of moments.

	$\hat{\Phi}(m^*, \kappa)$		
κ	Example 4.1	Example 4.2	Example 4.3
2	−6.52017	0.00075	−0.36578
3	−6.52587	0.00096	−0.36654
4	−6.53562	0.00063	−0.36658
5	−6.54718	0.00061	−0.37760

Table 1. *Optimal value of $(\mathscr{MP}_{d,\kappa})$ in dependence on κ.*

APPENDIX: RELAXATION BY CONVEX COMPACTIFICATION

Here, to justify the previous (occasionally only a bit vaguely used) concepts/results, we explain in detail the (classes of) Young measures and their usage for relaxation and approximation of variational problems used in the foregoing sections in the context of the general theory of convex compactifications of Lebesgue spaces, based essentially on [61–64].

A.1 Convex local compactifications of L^p-spaces

Following [63, 64], we will briefly present a fairly universal construction of locally compact envelopes of the Lebesgue L^p-spaces that are also convex in a natural linear space and allow for a continuous and affine extension of Nemytskiĭ mappings. We assumed $\Omega \subset \mathbb{R}^n$ a bounded domain and consider the Lebesgue space $L^p(\Omega;\mathbb{R}^m) = \{u : \Omega \to \mathbb{R}^m$ measurable; $\int_\Omega |u(x)|^p \mathrm{d}x < +\infty\}$. We define a normed linear space

$$\mathrm{Car}^p(\Omega;\mathbb{R}^m) :=\Big\{h : \Omega\times\mathbb{R}^m \to \mathbb{R};\ h(\cdot,s) \text{ measurable},\ h(x,\cdot) \text{ continuous}, \\ \exists a \in L^1(\Omega),\ b\in\mathbb{R}:\quad \big|h(x,s)\big| \le a(x) + b\big|s\big|^p\Big\} \tag{A.1}$$

of Carathéodory's "test integrands", and endow it with the norm

$$\big\|h\big\|_{\mathrm{Car}^p(\Omega;\mathbb{R}^m)} := \inf_{|h(x,s)|\le a(x)+b|s|^p} \big\|a\big\|_{L^1(\Omega)} + b. \tag{A.2}$$

The essential trick is to consider a sufficiently large (but preferably still separable) linear subspace $H \subset \mathrm{Car}^p(\Omega;\mathbb{R}^m)$, to define the embedding

$$i_H : L^p(\Omega;\mathbb{R}^m) \to H^* : u \mapsto \Big(h \mapsto \int_\Omega h(x,u(x))\,\mathrm{d}x\Big), \tag{A.3}$$

and eventually to put

$$Y^p_H(\Omega;\mathbb{R}^m) := \text{ the weak* closure of } i_H(L^p(\Omega;\mathbb{R}^m)). \tag{A.4}$$

One can show that $Y^p_H(\Omega;\mathbb{R}^m)$ is always convex in H^*. Assuming, rather for simplicity, that H contains at least one coercive integrand, i.e. $H \ni h_0$ with $h_0(x,s) \ge |s|^p$, then $Y^p_H(\Omega;\mathbb{R}^m)$ is a *convex locally compact hull* of $L^p(\Omega;\mathbb{R}^m)$ and $L^p(\Omega;\mathbb{R}^m)$ itself is embedded into it (norm,weak*)-continuously. Moreover, if H is rich enough (cf. [63, 64] for details), then this embedding i_H is even homeomorphic. If H is separable, then $Y^p_H(\Omega;\mathbb{R}^m)$ is locally sequentially compact. Thus $Y^p_H(\Omega;\mathbb{R}^m)$ may be considered as indeed a very natural envelope of $L^p(\Omega;\mathbb{R}^m)$, imitating a lot of geometrical/topological properties of (convex subsets in) finite-dimensional spaces.

Moreover, let us define $h\bullet\eta$ as a Borel measure on $\bar\Omega$ by

$$\int_{\bar\Omega} g(x)[h\bullet\eta](\mathrm{d}x) = \big\langle h\bullet\eta, g\big\rangle = \big\langle \eta, gh\big\rangle \tag{A.5}$$

where $[gh](x,s) = g(x)h(x,s)$ and $g \in C(\bar\Omega)$, where $\bar\Omega$ denotes the closure of Ω. Here we need that H is so-called $C(\bar\Omega)$-*invariant* in the sense that $gh \in H$ whenever $g \in C(\bar\Omega)$ and $h \in H$.

Further, we say that $\eta \in Y^p_H(\Omega;\mathbb{R}^m)$ is *p-nonconcentrating* if there is a sequence $\{u_k\}_{k\in\mathbb{N}}$ such that $\eta = \text{w*-}\lim_{k\to\infty} i_H(u_k)$ and $\{|u_k|^p;\ k\in\mathbb{N}\}$ is weakly relatively compact in $L^1(\Omega)$. Let us denote the set of all such η's by $\mathring{Y}^p_H(\Omega;\mathbb{R}^m)$.

If H is separable, any $\eta \in \mathring{Y}^p_H(\Omega;\mathbb{R}^m)$ has a L^p-*Young measure representation* $\nu \in \mathscr{Y}^p(\Omega;\mathbb{R}^m)$, cf. (2.1), in the sense that

$$\forall h \in H: \qquad \big\langle \eta, h\big\rangle = \int_\Omega\int_{\mathbb{R}^m} h(x,s)\,\nu_x(\mathrm{d}s)\,\mathrm{d}x, \tag{A.6}$$

see [63, Proposition 3.4.15]. It holds that $[h\bullet\eta](x)=\int_{\mathbb{R}^m}h(x,s)\,\nu_x(\mathrm{d}s)$ for a.a. $x\in\Omega$. If H is rich enough, e.g. if $H\supset L^1(\Omega;C_0(\mathbb{R}^m))$, then this representation is determined uniquely by η and, in fact, $\mathscr{Y}^p(\Omega;\mathbb{R}^m)\cong\mathring{Y}^p_H(\Omega;\mathbb{R}^m)$. In the general case, $\mathscr{Y}^p_H(\Omega;\mathbb{R}^m)\cong\mathring{Y}^p_H(\Omega;\mathbb{R}^m)$.

A.2 Relaxation of $(\mathscr{P})$

We use the construction from Section A.1 for $m=n$. Again we assume (1.1). Then we will consider the already announced relaxed problem in the form:

$$(RP_H)\qquad\left\{\begin{array}{ll}\text{Minimize} & \bar\Phi(u,\eta):=\displaystyle\int_\Omega\big[\varphi_1\bullet\eta\big](x)+\varphi_0(x,u(x))\,\mathrm{d}x,\\ \text{subject to} & [\mathrm{Id}\bullet\eta](x)=\nabla u(x)\quad\text{for a.a. }x\in\Omega,\\ & u\in W^{1,p}(\Omega),\quad\eta\in Y^p_H(\Omega;\mathbb{R}^n),\quad u|_{\partial\Omega}=u_{\mathrm{D}},\end{array}\right.$$

where $\mathrm{Id}(x,s):=s$; here we have to assume $H^n\ni\mathrm{Id}$. The following assertion, which claims that (RP_H) is indeed a proper relaxation of $(\mathscr{P})$, is based essentially on results by Kinderlehrer and Pedregal [27, 59]:

Proposition A.1. (See [63, Propositions 5.2.1 and 5.2.6].) *Let* (1.1) *hold,* $p>1$*,* H *be* $C(\bar\Omega)$*-invariant,* $\varphi_1\in H$*, and* $\mathrm{Id}\in H^n$*. Then:*
(i) (RP_H) *admits a solution.*
(ii) $\inf(\mathscr{P})=\min(RP_H)$.
(iii) *The embedding* $e_H:W^{1,p}(\Omega)\to W^{1,p}(\Omega)\times Y^p_H(\Omega;\mathbb{R}^n):v\mapsto(v,i_H(\nabla v))$ *of any minimizing sequence for* $(\mathscr{P})$ *has a weakly convergent subsequence whose (weak×weak*) limit is a solution to* (RP_H).
(iv) *Each solution to* (RP_H) *is* p*-nonconcentrating and is the (weak×weak*) limit of some minimizing sequence for* $(\mathscr{P})$ *embedded via* e_H.

The consequence of Proposition A.1(iv) is that we can replace $Y^p_H(\Omega;\mathbb{R}^n)$ by $\mathring{Y}^p_H(\Omega;\mathbb{R}^n)$ with an equal effect. Then, using the already mentioned equivalence $\mathring{Y}^p_H(\Omega;\mathbb{R}^n)\cong\mathscr{Y}^p_H(\Omega;\mathbb{R}^m)$, we can even replace (RP_H) by $(\mathscr{RP}_H)$ in Proposition A.1.

A.3 Approximation: finite-element discretization in space

As for the discretization of Ω, we suppose that Ω is a polyhedral and we also consider a discretization mesh parameter $d>0$ (ranging a countable set having 0 as its accumulation point) and a triangulation $\mathcal{T}_d$ of Ω composed from simplexes $S\in\mathcal{T}_d$ such that $\max_{S\in\mathcal{T}_d}\mathrm{diam}(S)\le d$ and $\mathcal{T}_{d_1}\subset\mathcal{T}_{d_2}$ for $d_1\ge d_2>0$, i.e. we consider nested triangulations refining everywhere on Ω when $d\searrow 0$. Then we define P_d by

$$[P_dh](x,s):=\frac{1}{\mathrm{meas}_n(S)}\int_S h(\xi,s)\mathrm{d}\xi\qquad\text{if }x\in S\in\mathcal{T}_d.\tag{A.7}$$

Requiring $P_d:H\to H$, we must consider such H which contains also discontinuous element-wise constant integrands. As we consider a sequence of triangulations $\mathcal{T}_d$, it is still possible to take H separable in the norm (A.2). In this norm, one can see that $\|P_dh\|_{\mathrm{Car}^p(\Omega;\mathbb{R}^m)}\le\|h\|_{\mathrm{Car}^p(\Omega;\mathbb{R}^m)}$ and $P_d\circ P_d=P_d$, so that $P_d:H\to H$ is a continuous projector. By [63, Proposition 3.5.9], it holds that $P_d^*Y^p_H(\Omega;\mathbb{R}^n)\subset Y^p_H(\Omega;\mathbb{R}^n)$. By [63, Proposition 3.5.2(iv)], we have $\bigcup_{d>0}P_d^*Y^p_H(\Omega;\mathbb{R}^n)$ weakly* dense in $Y^p_H(\Omega;\mathbb{R}^n)$. This suggests to approximate (RP_H) by restricting it on a convex subset $P_d^*Y^p_H(\Omega;\mathbb{R}^n)$ instead of $Y^p_H(\Omega;\mathbb{R}^n)$. Since any $\eta\in P_d^*Y^p_H(\Omega;\mathbb{R}^n)$ is element-wise constant, holding the constraint $\mathrm{Id}\bullet\eta = \nabla u$, the underlying u will then be automatically element-wise affine. By this way we come to the following *approximate relaxed problem*:

$$(RP_{H,d})\qquad \begin{cases} \text{Minimize} & \bar\Phi(u,\eta) := \displaystyle\int_\Omega \big[\varphi_1\bullet\eta\big](x) + \varphi_0(x,u(x))\,\mathrm{d}x, \\ \text{subject to} & [\mathrm{Id}\bullet\eta](x) = \nabla u(x) \quad \text{for a.a. } x\in\Omega, \\ & u\in W^{1,p}(\Omega) \text{ element-wise affine on } \mathcal{T}_d, \\ & \eta \in P_d^* Y_H^p(\Omega;\mathbb{R}^n), \qquad u|_{\partial\Omega} = u_\mathrm{D}. \end{cases}$$

Let us denote

$$G_0 := \bigcup_{d>0} \big\{ g\in L^\infty(\Omega);\ \ \forall S\in\mathcal{T}_d:\ \ g|_S \in C(\bar S)\big\} \tag{A.8}$$

where $g|_S\in C(\bar S)$ means existence of a continuous extension on $\bar S$ of the restriction $g|_S$.

Proposition A.2. (See [63, Proposition 5.5.1].) *Let H be separable, G-invariant, and satisfy $G\otimes V\subset H\subset \mathrm{cl}(G\otimes V)$ for some linear space $G\subset L^\infty(\Omega)$ containing G_0 and for some linear space V of continuous functions on $\mathbb{R}^n$ of at most p-growth, containing $s\mapsto s_i$, $i=1,\dots,n$, so that $\mathrm{Id}\in H^n$, $\varphi\in H$, where the closure "cl" refers to the norm* (A.2) *and "$\otimes$" means the usual tensorial products, i.e. for functions $[g\otimes v](x,s) = g(x)v(s)$ and for spaces $G\otimes V$ is the linear span of $\{g\otimes v;\ g\in G,\ v\in V\}$. Then:*
(i) *A solution (u_d,η_d) to $(RP_{H,d})$ always exists.*
(ii) *Moreover,* $\lim_{d\to0}\min(RP_{H,d}) = \min(RP_H)$ *and there always exists a subsequence of $d_i\to0$ such that (u_{d_i},η_{d_i}) (weak×weak*)-converges in $W^{1,p}(\Omega)\times H^*$. Moreover, the limit of any such a subsequence solves (RP_H).*

Using that, like in Proposition A.1(iv), also here the solutions to $(RP_{H,d})$ are p-nonconcentrating, and the equivalence $P_d^*\mathring{Y}_H^p(\Omega;\mathbb{R}^n) \cong \mathring{Y}^p_{P_dH}(\Omega;\mathbb{R}^m) \cong \mathscr{Y}^p_{P_dH}(\Omega;\mathbb{R}^n) = \{\eta\in\mathscr{Y}^p_H(\Omega;\mathbb{R}^n)$ elementwise constant on $\mathcal{T}_d\}$, one eventually can replace $(RP_{H,d})$ by $(\mathscr{RP}_{H,d})$ and use Proposition A.2 for $H=H_k$, as implicitly exploited before.

Remark A.3. As P_d is a projector, it holds that

$$\begin{aligned}\int_\Omega \varphi_1\bullet\eta_d\,\mathrm{d}x &= \big\langle\eta_d,\varphi_1\big\rangle = \big\langle P_d^*\eta_d,\varphi_1\big\rangle = \big\langle P_d^*P_d^*\eta,\varphi_1\big\rangle \\ &= \big\langle P_d^*\eta, P_d\varphi_1\big\rangle = \big\langle \eta_d, P_d\varphi_1\big\rangle = \int_\Omega (P_d\varphi_1)\bullet\eta_d\,\mathrm{d}x \end{aligned}\tag{A.9}$$

for any $\eta_d\in P_d^*Y_H^p(\Omega;\mathbb{R}^n)$, i.e. $\eta_d = P_d^*\eta$ for some $\eta\in Y_H^p(\Omega;\mathbb{R}^n)$, and therefore we can equally consider φ_1 in $(RP_{H,d})$ replaced by its element-wise constant interpolant $P_d\varphi_1$. Also, by [63, Proposition 5.5.1(ii)], $P_d^*Y_H^p(\Omega;\mathbb{R}^n)$ in $(RP_{H,d})$ can be replaced by $P_d^*\mathring{Y}_H^p(\Omega;\mathbb{R}^n)$ with an equal effect.

In the following remark, we focus on the Weierstrass Maximum Principle as a necessary condition for general variational problems.

Remark A.4. Fixing $d>0$, the *necessary optimality conditions* for $(RP_{H,d})$, which any solution (u_d,η_d) to $(RP_{H,d})$ must satisfy, are the existence of a vector field $\lambda_d\in L^\infty(\Omega;\mathbb{R}^n)$ element-wise constant satisfying, roughly speaking, one *half of the Euler-Lagrange equation* $\mathrm{div}\lambda_d = [\varphi_0]'_u(x,u_d)$ after being discretized by finite elements, i.e.

$$\forall z\in W^{1,\infty}(\Omega) \text{ element-wise affine on } \mathcal{T}_d:\quad \int_\Omega \lambda_d\cdot\nabla z + \frac{\partial\varphi_0(x,u_d)}{\partial u} z\,\mathrm{d}x = 0, \tag{A.10}$$

and the *Weierstrass maximum principle* in the sense

$$\big[\lambda_d \otimes \mathrm{Id} - P_d\varphi_1\big] \bullet \eta_d = \max_{s\in\mathbb{R}^n} \Big(\lambda_d(x)\cdot s - [P_d\varphi_1](x,s)\Big) \qquad \text{for a.a. } x \in \Omega, \tag{A.11}$$

see [63, Sect.5.3] or [12]. The vector field λ_d is, in fact, the Lagrange multiplier to the constraint $\mathrm{Id}\bullet\eta = \nabla u$ in $(RP_{H,d})$. If $\varphi_0(x,\cdot)$ is convex, then these optimality conditions are also sufficient. Indeed, formula (A.11) suggests the use of a relaxation in probability measures of the global optimization problem: $\max_{s\in\mathbb{R}^n}(\lambda_d(x)\cdot s - [P_d\varphi_1](x,s))$. This is just the approach presented in [33, 43, 44] for global optimization of polynomials.

Let us actually emphasize that, as we have focused on the case where all integrands from H are polynomials with the order at most $2k$, $k \in \mathbb{N}$; i.e. $h(x,\cdot) \in \Pi_{2k}(\mathbb{R}^n)$, $P_d H$ is finite-dimensional and then $P_d^* Y_H^p(\Omega;\mathbb{R}^n)$ is automatically homeomorphic to a convex subset of a finite-dimensional Euclidean space; notice the constraints $m_{S,0,\cdots,0} = 1$ for $S \in \mathcal{T}_d$ in $(\mathscr{MP}_{d,\kappa})$ above, which allowed us a computer implementation without no further discretization.

Acknowledgments. R.M. gratefully acknowledges support by Grant 1810, Fundación para la Promoción de la Investigación y la Tecnología while T.R. warmly acknowledged the hospitality of Universidad de los Andes, Bogotá, and also the support from the grants A 107 5402 (GA AV ČR) and LC 06052 and MSM 21620839 (MŠMT ČR), and from the research plan AV0Z20760514 (ČR). Finally, we thank the hospitality of La Universidad de Castilla La Mancha, at Ciudad Real, where this research originated in 2003, and CIMNE, Centro Internacional de Métodos Numéricos, Barcelona.

References

[1] N.I. Akhiezer and M.G. Krein (1962). *Some Questions in the Theory of Moments.* Amer. Math. Soc., Providence, RI.

[2] S. Bartels (2004). Adaptive approximation of Young measure solutions in scalar nonconvex variational problems. *SIAM J. Numer. Anal.*, **42**, 505-529.

[3] S. Bartels and T. Roubíček (2004). Linear-programming approach to nonconvex variational problems. *Numer. Math.*, **99**, 251-287.

[4] A. Ben-Tal and A. Nemirovski (2001). *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications.* SIAM, Philadelphia, PA.

[5] C. Berg (1987). The multidimensional moment problem and semigroups. In: H. J. Landau (Ed.), *Moments in Mathematics* (pp. 110-124). Amer. Math. Soc., Providence, RI.

[6] D. Bertsimas, K. Natarajan, and C.-P. Teo (2004). Probabilistic combinatorial optimization: Moments, semidefinite programming, and asymptotic bounds. *SIAM J. Optim.* **15**, 185-209.

[7] D. Bertsimas, I. Popescu, and J. Sethuraman (2000). Moment problems and semidefinite optimization. In: H. Wolkovicz, R. Saigal, L. Vandenberghe (Eds.). *Handbook of Semidefinite Programming – Theory, Algorithms, Appl.* (pp. 469-509), Kluwer, Dordrecht.

[8] E. Bonnetier and C. Conca (1994). Approximation of Young measures by functions and application to a problem of optimal design for plates with variable thickness. *Proc. Roy. Soc. Edinburgh Sect. A* **124**, 399-422.

[9] C. Carstensen (2005). Ten remarks on nonconvex minimisation for phase transition simulations. *Comput. Methods Appl. Mech. Engrg.* **194** (2005), 169-193.

[10] C. Carstensen, K. Jochimsen (2003). Adaptive finite element methods for microstructures? Numerical experiments for a 2-well benchmark. *Computing* **71**, 175-204.

[11] C. Carstensen and P. Plecháč (1997). Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comp.* **66**, 997-1026.

[12] C. Carstensen and T. Roubíček (2000). Numerical approximation of Young measures in nonconvex variational problems. *Numer. Math.* **84**, 395-415.

[13] G. Cassier (1984). Problème des moments sur un compact de $\mathbb{R}^n$ et décomposition de polynômes a plusieurs variables. *J. Funct. Anal.* **58**, 254-266.

[14] L. Cesari (1983). *Optimization – Theory and applications.* Springer, New York.

[15] R.E. Curto and L.A. Fialkow (1991). Recursiveness, positivity, and truncated moment problems. *Houston J. Math.* **17**, 603-635.

[16] R.E. Curto and L.A. Fialkow (1996). Solution of the truncated complex moment problem for flat data. *Mem. Amer. Math. Soc.* **568**, 52 p.

[17] R.E. Curto and L.A. Fialkow (2000). The truncated complex K-moment problem. *Trans. Amer. Math. Soc.* **352**, 2825-2855.
[18] B. Dacorogna (1989). *Direct Methods in the Calculus of Variations.* Springer, Berlin.
[19] J. Egozcue, R.J. Meziat, and P. Pedregal (2003). From a non-linear non-convex formulation to a linear convex formulation of non-convex variational problems. *Appl. Math. Optim.* **47**, 27-44.
[20] H.O. Fattorini (1999). *Infinite Dimensional Optimization and Control Theory.* Cambridge Univ. Press, Cambridge.
[21] K. Gatermann, and P.A. Parrilo (2004). Symmetry groups, semidefinite programs, and sums of squares. *J. Pure Appl. Algebra* **192**, 95-128.
[22] T. Jacobi and A. Prestel (2001). Distinguished representations of strictly positive polynomials. *J. Reine Angew. Math.* **532**, 223-235.
[23] J. Jost and X. Li-Jost (1999). *Calculus of Variations.* Cambridge Univ. Press, Cambridge.
[24] D. Handelman (1988). Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific J. Math.* **132**, 35-62.
[25] D. Henrion and J.B. Lasserre (2003). GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi. *ACM Trans. Math. Software*, **29**, 165-194.
[26] S. Karlin and W.J. Studden (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics.* Interscience Publ., New York.
[27] D. Kinderlehrer and P. Pedregal (1994). Gradient Young measures generated by sequences in Sobolev spaces. *J. Geom. Anal.*, **4**, 59-90.
[28] M. Kočvara and M. Stingl (2003). PENNON – A code for convex nonlinear and semidefinite programming. *Optim. Methods Softw.*, **18**, 317-333. (See also `http://www.penopt.com`)
[29] M.G. Krein and A.A. Nudelman (1977). *The Markov Moment Problem and Extremal Problems.* Amer. Math. Soc., Providence, RI.
[30] M. Kružík (1998). Numerical approach to double well problems. *SIAM J. Numer. Anal.* **35**, 1833-1849.
[31] M. Kružík (2003). Maximum principle based algorithm for hysteresis in micromagnetics. *Adv. Math. Sci. Appl.* **13**, 461-485.
[32] M. Kružík and A. Prohl (2001). Young measure approximation in micromagnetics. *Numer. Math.* **90**, 291-307.
[33] J.B. Lasserre (2001). Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11**, 796-817.
[34] J.B. Lasserre (2002). Semidefinite programming vs. LP relaxations for polynomial programming. *Math. Oper. Res.* **27**, 347-360.
[35] J.B. Lasserre (2005). SOS approximations of polynomials nonnegative on a real algebraic set. *SIAM J. Optim.* **16**, 610-628.
[36] J.B. Lasserre (2006). A sum of squares approximation of nonnegative polynomials, *SIAM J. Optim.* **16**, 751-765.
[37] J.B. Lasserre (2008) A semidefinite programming approach to the generalized problem of moments. *Math. Programm.* **112**, 65-92.
[38] J.B. Lasserre and T. Prieto-Rumeau (2004). SDP vs. LP relaxations for the moment approach in some performance evaluation problems. *Stoch. Models* **20**, 439-456.
[39] M. Laurent (2005). Revisiting two theorems of Curto and Fialkow on moment matrices. *Proc. Amer. Math. Soc.* **133**, 2965-2976.
[40] M. Laurent and D. Jibetean (2005). Semidefinite approximations for global unconstrained polynomial optimization. *SIAM J. Optim.* **16**, 490-514.
[41] J. Mach (2002). Numerical solution of a class of nonconvex variational problems by SQP. *Numer. Funct. Anal. Optim.* **23**, 573-587.
[42] A.-M. Mataché, T. Roubíček, and C. Schwab (2003). Higher-order convex approximations of Young measures in optimal control. *Adv. Comput. Math.* **19**, 73-97.
[43] R.J. Meziat (2003). Analysis of nonconvex polynomial programs by the method of moments. In: C. A. Floudas, P. M. Pardalos (Eds.), *Frontiers in Global Optimization* (pp. 353-372). Kluwer, Dordrecht.
[44] R.J. Meziat (2003). The method of moments in global optimization. *J. Math. Sci. (N. Y.)* **116**, 3303-3324.
[45] R.J. Meziat (2005). Analysis of two dimensional nonconvex variational problems. In: L. Qi, K. Teo, X. Yang (Eds.), *Optimization and Control with Applications* (pp. 393-405). Springer, New York.
[46] R.J. Meziat, J. Egozcue, and P. Pedregal (2001). The method of moments for nonconvex variational problems. In: N. Hadjisavvas, P. M. Pardalos (Eds.), *Adv. in Convex Anal. and Global Optim.* (pp. 371-382). Kluwer, Dordrecht.
[47] R.J. Meziat and D. Patiño (2008). Exact relaxation of non-convex variational problems. *Optimization Letters* **2**, 505-519.

[48] R.J. Meziat and J. Villalobos (2006). Analysis of microstructures and phase transition phenomena in one-dimensional, non-linear elasticity by convex optimization. *Struct. Multidiscip. Optim.* **32**, 507-519.

[49] A.A. Milyutin and N.P. Osmolovskii (1998). Calculus of variations and optimal control. *Transl. Math. Monogr.*, **180**. Amer. Math. Soc., Providence, RI.

[50] B.S. Mordukhovich (1999). Existence theorems in nonconvex optimal control. In: A. Ioffe, S. Reich, I. Shafrir (Eds.), *Calculus of Variations and Optimal Control* (pp.173-197). Chapman & Hall/CRC, Boca Raton, FL.

[51] J. Muñoz and P. Pedregal (2001). A refinement on existence results in nonconvex optimal control. *Nonlinear Anal.* **46**, 381-398.

[52] Y. Nesterov (2000). Squared functional systems and optimization problems. In: H. Frenk, K. Roos, T. Terlaky, S. Zhang (Eds.). *High Performance Optimization* (pp. 405-440). Kluwer, Dordrecht.

[53] R.A. Nicolaides and N.J. Walkington (1995). Strong convergence of numerical solutions to degenerate variational problems. *Math. Comp.* **64**, 117-127.

[54] P.A. Parrilo (2003). Semidefinite programming relaxations for semialgebraic problems. *Math. Program.* **96**, 293-320.

[55] P.A. Parrilo and S. Lall (2003). Semidefinite programming relaxations and algebraic optimization in control. *Eur. J. Control* **9**, 307-321.

[56] P.A. Parrilo and B. Sturmfels (2003). Minimizing polynomial functions. In: S. Basu, L. González-Vega (Eds.), *Algorithmic and Quantitative Real Algebraic Geometry* (pp. 83-99). Amer. Math. Soc., Providence, RI

[57] P. Pedregal (1995). Numerical approximation of parametrized measures. *Numer. Funct. Anal. Optim.* **16**, 1049-1066.

[58] P. Pedregal (1996). On the numerical analysis of non-convex variational problems. *Numer. Math.* **74**, 325-336.

[59] P. Pedregal (1997). *Parametrized Measures and Variational Principles.* Birkhäuser, Boston.

[60] M. Putinar (1993). Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, **42**, 969-984.

[61] T. Roubíček (1995). Approximation theory for generalized Young measures. *Numer. Funct. Anal. Optim.* **16**, 1233-1253.

[62] T. Roubíček (1996). Numerical approximation of relaxed variational problems. *J. Convex Anal.* **3**, 329-347.

[63] T. Roubíček (1997). *Relaxation in Optimization Theory and Variational Calculus.* W. de Gruyter, Berlin.

[64] T. Roubíček (1999). Convex locally compact extensions of Lebesgue spaces and their applications. In: A. Ioffe, S. Reich, I. Shafrir (Eds.), *Calculus of Variations and Optimal Control* (pp. 237-250). Chapman & Hall/CRC, Boca Raton, FL.

[65] T. Roubíček (2006). Numerical techniques in relaxed optimization problems. In: A. J. Kurdila, P. M. Pardalos, M. Zabarankin (Eds.), *Robust Optimization-Directed Design* (pp. 145-161). New York: Springer.

[66] T. Roubíček and V. Šverák (2000). Nonexistence of solutions in nonconvex multidimensional variational problems. *J. Convex Anal.* **7**, 427-436.

[67] K. Schmudgen (1991). The K-moment problem for compact semi-algebraic sets. *Math. Ann.* **289**, 203-206.

[68] J.A. Shohat and J.D. Tamarkin (1943). The problem of moments. *Math. Surveys Monogr.*, **1**. Amer. Math. Soc., Providence, RI.

[69] S.J. Wright (1997). *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, PA.

[70] J. Zabczyk (1992). *Mathematical Control Theory: An Introduction.* Birkhäuser, Boston.