

Numerical approximation of young measures in non-convex variational problems

Carsten Carstensen¹, Tomáš Roubíček^{2,3}

¹ Mathematisches Seminar, Christian-Albrecht-Universität zu Kiel,
Ludewig-Meyn-Strasse 4, D-24098 Kiel, Germany

² Mathematical Institute, Charles University, Sokolovská 83, CZ-186 75 Praha 8,
Czech Republic

³ Institute of Information Theory and Automation, Academy of Sciences,
Pod vodárenskou věží 4, CZ-182 08 Praha 8, Czech Republic

Received November 26, 1997 / Published online September 24, 1999

Summary. In non-convex optimisation problems, in particular in non-convex variational problems, there usually does not exist any classical solution but only generalised solutions which involve Young measures. In this paper, first a suitable relaxation and approximation theory is developed together with optimality conditions, and then an adaptive scheme is proposed for the efficient numerical treatment. The Young measures solving the approximate problems are usually composed only from a few atoms. This is the main argument our effective active-set type algorithm is based on. The support of those atoms is estimated from the Weierstrass maximum principle which involves a Hamiltonian whose good guess is obtained by a multilevel technique. Numerical experiments are performed in a one-dimensional variational problem and support efficiency of the algorithm.

Mathematics Subject Classification (1991): 65K10, 65N50, 49M40

1. Introduction

Many problems from non-convex optimisation theory (as e.g. optimal control and non-cooperative games) and variational calculus do not possess a classical solution because approximate solutions show typically fast oscillation. This phenomenon leads to a necessity of a natural extension of such problems, often constructed by means of Young measures. It is the purpose

Correspondence to: C. Carstensen

of this paper to propose an effective algorithm for computing a numerical approximation to Young measures solving such problems.

To simplify the explanation we focus on the following multidimensional scalar variational problem, although the adaptive approximation algorithm proposed in this paper works in a general context, as well.

$$(P) \quad \begin{cases} \text{Minimise } \Phi(u) := \int_{\Omega} F(x, \nabla u(x)) + G(x, u(x)) \, dx, \\ \text{subject to } u \in W^{1,p}(\Omega), \quad u|_{\partial\Omega} = u_D. \end{cases}$$

Here, $W^{1,p}(\Omega)$ is the standard Sobolev space $\{u : \Omega \rightarrow \mathbb{R}; (u, \nabla u) \in L^p(\Omega; \mathbb{R}^{n+1})\}$ and $u_D \in W^{1-1/p,p}(\partial\Omega)$ determines the prescribed trace of u on the boundary $\partial\Omega$ of Ω . Through the paper, $\Omega \subset \mathbb{R}^n$ will be a bounded Lipschitz domain and $L^p(\Omega; \mathbb{R}^n)$ a Lebesgue space of measurable functions $y : \Omega \rightarrow \mathbb{R}^n$ such that $\|y\| := (\int_{\Omega} |y(x)|^p dx)^{1/p} < +\infty$; we assume $1 < p < +\infty$. Furthermore, we will assume that $F : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}$ and $G : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ are Carathéodory functions satisfying

$$(1.1) \quad c|s|^p - C \leq F(x, s) \leq C(1 + |s|^p), \quad |G(x, u)| \leq a(x) + C|u|^q$$

for some $a \in L^1(\Omega)$, $c > 0$ and $C \in \mathbb{R}$, and $1 < q < pn/(n-p)$ (if $p \geq n$, then simply $q < +\infty$). These assumptions guarantee that $\Phi : W^{1,p}(\Omega) \rightarrow \mathbb{R}$ is well defined, continuous, and coercive. However, Φ is not weakly lower semi-continuous in case $F(x, \cdot)$ is not convex. This causes faster and faster oscillations of the gradient of minimising sequences of (P) and a failure of existence of a solution to (P); for numerical analysis of direct minimisation of discretized problems we refer to [C, CCK, CKL, CL, L]. To relax this problem we can either consider a weak lower semi-continuous envelope of Φ or extend Φ continuously on a suitable hull of the original space $W^{1,p}(\Omega)$. The former option requires to make a convexification of $F(x, \cdot)$ which may be practically difficult to obtain even in very special cases; cf. [CP] or [BC] for a numerical approach. Here we will follow the latter option using the continuous extension which is easy to calculate and yields also more information about oscillations of the gradient of minimising sequences of (P). For other works in this spirit see e.g. [NW1, NW2, P1, P2, R1, R2, R3].

2. Young measures and relaxed variational problems

To construct a suitable extension of (P), we will first introduce a certain convex σ -compact envelope of the Lebesgue space $L^p(\Omega; \mathbb{R}^n)$.

Let $\text{rca}(\mathbb{R}^n) \cong C_0(\mathbb{R}^n)^*$ denote the space of Radon measures on \mathbb{R}^n , $C_0(\mathbb{R}^n)$ being the space of continuous functions on \mathbb{R}^n vanishing at infinity, cf. [DS]. Then we consider $L_w^\infty(\Omega; \text{rca}(\mathbb{R}^n)) \cong L^1(\Omega; C_0(\mathbb{R}^n))^*$ (the

subscript “w” stands for the adjective “weakly measurable”) and define the set of L^p -Young measures

$$(2.1) \quad \mathcal{Y}^p(\Omega; \mathbb{R}^n) := \left\{ \nu \equiv \{\nu_x\}_{x \in \Omega} \in L_w^\infty(\Omega; \text{rca}(\mathbb{R}^n)); \right. \\ \left. \nu_x \in \text{rca}_1^+(\mathbb{R}^n) \text{ for a.a. } x \in \Omega, \right. \\ \left. \int_\Omega \int_{\mathbb{R}^n} |s|^p \nu_x(ds) dx < +\infty \right\}.$$

Here, $\text{rca}_1^+(\mathbb{R}^n) := \{\mu \in \text{rca}(\mathbb{R}^n); \mu \geq 0, \int_{\mathbb{R}^n} \mu(ds) = 1\}$ is the set of probability Radon measures. Also, we have used the usual convention writing ν_x instead of $\nu(x)$. A natural (norm, weak*)-continuous and dense embedding of $L^p(\Omega; \mathbb{R}^n)$ into $\mathcal{Y}^p(\Omega; \mathbb{R}^n)$ is defined by $y \mapsto \nu := \{\delta_{y(x)}\}_{x \in \Omega}$ with $\delta_s \in \text{rca}_1^+(\mathbb{R}^n)$ denoting the Dirac measure supported at $s \in \mathbb{R}^n$. The set $\mathcal{Y}^p(\Omega; \mathbb{R}^n)$ thus forms a convex σ -compact hull of $L^p(\Omega; \mathbb{R}^n)$; see e.g. [R2, Chap. 3] for more details.

A relaxation of (P) can now be done simply by a continuous extension of all data involved in (P) from $L^p(\Omega; \mathbb{R}^n)$ to its convex hull, namely $\mathcal{Y}^p(\Omega; \mathbb{R}^n)$. This creates the problem

$$(RP) \quad \begin{cases} \text{minimise } \bar{\Phi}(u, \nu) := \int_\Omega \left[\int_{\mathbb{R}^n} F(x, s) \nu_x(ds) + G(x, u(x)) \right] dx, \\ \text{subject to } \int_{\mathbb{R}^n} s \nu_x(ds) = \nabla u(x) \text{ for a.a. } x \in \Omega, \\ u \in W^{1,p}(\Omega), \nu \in \mathcal{Y}^p(\Omega; \mathbb{R}^n), u|_{\partial\Omega} = u_D. \end{cases}$$

Let us note that the problem (RP) has a convex structure (i.e. minimises a convex functional $\bar{\Phi}$ on a convex set of feasible pairs (u, ν) provided $G(x, \cdot)$ is convex; the convexity of $F(x, \cdot)$ is not needed for the convexity of $\bar{\Phi}$). The important fact is that (RP) is a correct relaxation of (P).

Proposition 1. *If (1.1) holds, then*

- (i) (RP) always possesses a solution,
- (ii) $\inf(P) = \min(RP)$,
- (iii) every minimising sequence of (P) contains a subsequence converging (when embedded by $u \mapsto (u, \{\delta_{\nabla u(x)}\}_{x \in \Omega})$) to a solution to (RP), and
- (iv) conversely, any solution to (RP) can be attained by a minimising sequence of (P).

Sketch of the proof. Following [R2], we take a sufficiently large (but separable) linear space of test integrands $H := C(\bar{\Omega}) \cdot \{F\} + C(\bar{\Omega}) \otimes C_p(\mathbb{R}^m)$ with $C_p(\mathbb{R}^n) := \{v \in C(\mathbb{R}^n); \lim_{|s| \rightarrow \infty} v(s)|s|^{-p} = 0\}$ and “ \otimes ” the tensorial

product, i.e. $H = \{g_0(x)F(x, s) + \sum_{\text{finite}} g_i(x)v_i(s); g_0, g_i \in C(\bar{\Omega}), v_i \in C_p(\mathbb{R}^n)\}$. It is natural to endow H by the norm

$$(2.2) \quad \|h\|_H := \sup_{(x,s) \in \Omega \times \mathbb{R}^n} \frac{|h(x, s)|}{1 + |s|^p}.$$

We put $Y_H^p(\Omega; \mathbb{R}^n) = \{\eta \in H^*; \exists \{y_i\}_{i \in \mathbb{N}} \subset L^p(\Omega; \mathbb{R}^n) : \eta = \text{w}^*\text{-}\lim_{i \rightarrow \infty} i_H(u_k)\}$ where the embedding $i_H : L^p(\Omega; \mathbb{R}^n) \rightarrow H^*$ is defined by $\langle i_H(y), h \rangle = \int_{\Omega} h(x, y(x))dx$. Then we introduce an auxiliary relaxed problem

$$(RP_H) \quad \begin{cases} \text{minimise } \bar{\Phi}_0(u, \eta) := \langle \eta, F \rangle + \int_{\Omega} G(x, u(x))dx, \\ \text{subject to } \langle \eta, g \otimes \text{id} \rangle = \nabla y \cdot g \quad \forall g \in C(\bar{\Omega}; \mathbb{R}^n), \\ u \in W^{1,p}(\Omega), \eta \in Y_H^p(\Omega; \mathbb{R}^n), u|_{\partial\Omega} = u_D, \end{cases}$$

with $\text{id} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ denoting the identity. It is known [R2, Sect. 5.2] that (RP_H) is a correct relaxation of (P) in the sense that (RP_H) always possesses a solution, $\inf(P) = \min(RP_H)$, every minimising sequence of (P) contains a subsequence converging, when embedded by $u \mapsto (u, i_H(\nabla u))$, to a solution to (RP_H) , and conversely any solution to (RP_H) can be attained by a minimising sequence of (P) .

Moreover, it is known [R2, Sect. 5] that, since $p > 1$, there cannot appear concentration of energy in the sense that the set $\{x \mapsto F(x, \nabla u_i(x)); i \in \mathbb{N}\}$ is not only bounded in $L^1(\Omega)$ but even relatively weakly compact in $L^1(\Omega)$ provided $\{u_i\}_{i \in \mathbb{N}}$ is a minimising sequence for (P) .

Therefore, we can equally modify (RP_H) by adding the restriction that η is attainable by sequences non-concentrating energy in the above sense (those η is called p -nonconcentrating). This p -nonconcentrating $\eta \in Y_H^p(\Omega; \mathbb{R}^n)$ has a Young measure representation ν in the sense

$$(2.3) \quad \langle \eta, h \rangle = \int_{\Omega} \int_{\mathbb{R}^n} h(x, s) \nu_x(ds) dx$$

for any $h \in H$. The Young measure ν is defined uniquely because H is rich enough, namely $H \supset C(\bar{\Omega}) \otimes C_p(\mathbb{R}^n)$, cf. [S] or also Ball [B] or [R2, Chap. 3]. This shows that (RP) and (RP_H) are equivalent.

As to (iii), having a minimizing sequence $\{u_i\}_{i \in N}$ for (P) , by local sequential weak* compactness of $Y_H^p(\Omega; \mathbb{R}^n)$ there is its subsequence (denoted again as $\{u_i\}$) such that $(u_i, i_H(\nabla u_i)) \rightarrow (u, \eta)$ weakly* and (u, η) solves (RP_H) . In particular, η must be p -nonconcentrating, and we can consider its Young-measure representation ν . Then (u, ν) solves (RP) and $(u_i, \delta(\nabla u_i)) \rightarrow (u, \nu)$ weakly* because $\mathcal{Y}^p(\Omega; \mathbb{R}^m)$ is a coarser σ -compactification of $L^p(\Omega; \mathbb{R}^n)$ than $Y_H^p(\Omega; \mathbb{R}^n)$, cf. [R2] for details.

Conversely, if (u, ν) solves (RP), then (u, η) with $\eta \in Y_H^p(\Omega; \mathbb{R}^n)$ given by (2.3) solves (RP_H) , and then (cf. [R2, Sect. 5.2]) there is a minimising sequence $\{u_i\}_{i \in \mathbb{N}}$ for (P) attaining (u, η) . This sequence attains (u, ν) , too, which proves (iv). \square

3. Convex approximation of Young measures

We will now construct a convex finite-dimensional approximation of the set $\mathcal{Y}^p(\Omega; \mathbb{R}^n)$. For this, we need to discretise both Ω and \mathbb{R}^n , i.e. the x - and s -variables.

As to the discretisation of Ω , we now suppose Ω polyhedral and consider, for every $l \in \mathbb{N}$, a triangulation \mathcal{T}_l of Ω such that $\lim_{l \rightarrow \infty} \max_{\Delta \in \mathcal{T}_l} \text{diam}(\Delta) = 0$ and $\mathcal{T}_l \subset \mathcal{T}_{l+1}$, i.e. the nested triangulations refine everywhere on Ω . Then we define P_l by

$$(3.1) \quad [P_l h](x, s) := \sum_{i=1}^{(k+1)^n} \frac{1}{\text{meas}_n(\Delta)} \int_{\Delta} h(\xi, s) d\xi \quad \text{if } x \in \Delta \in \mathcal{T}_l.$$

Requiring $P_l : H \rightarrow H$, we must slightly enlarge the linear space H defined in the proof of Proposition 1, namely we will consider

$$(3.2) \quad H := C(\bar{\Omega}) \cdot \{F\} + \bigcup_{l \in \mathbb{N}} C_{\mathcal{T}_l}(\Omega) \otimes C_p(\mathbb{R}^m),$$

where $C_{\mathcal{T}_l}(\Omega) := \{g \in L^\infty(\Omega); \forall \Delta \in \mathcal{T}_l : g|_{\Delta} \in C(\bar{\Delta})\}$ denotes the space of element-wise continuous functions. Endowing H again by the norm (2.2), one can see that $\|P_l h\|_H \leq \|h\|_H$ and $P_l \circ P_l = P_l$, so that $P_l : H \rightarrow H$ is a continuous projector. Moreover, since the collection of triangulations is countable, $\bigcup_{l \in \mathbb{N}} C_{\mathcal{T}_l}(\Omega)$ is separable if endowed by the L^∞ -norm and so is also H .

As to the discretisation of \mathbb{R}^n , we now choose $m \in \mathbb{R}^+$ and $k \in \mathbb{N}$ and make a uniform partition of the hypercube $B_m^n := [-m, m]^n \subset \mathbb{R}^n$ to k^n similar hypercubes of the side $2m/k$ and consider thus a discretisation of the whole range \mathbb{R}^n as outlined on Fig. 1 for $n = 1$ and $n = 2$; of course, some parts of this partition of \mathbb{R}^n are inevitably unbounded.

Then we use the Q1 finite elements inside the hypercube B_m^n ; let us remind that these elements are just the tensorial products of linear functions in each coordinate. In such a way, we get the collection $\{v_i; i = 1, \dots, (k + 1)^n\}$ of the basis functions $B_m^n \rightarrow \mathbb{R}$ defined by taking 1 at a selected grid point and 0 at all resting $(k + 1)^n - 1$ grid points. We can prolong the basis functions v_i outside the hypercube B_m^n just by putting $v_i(s) := v_i(\text{Pr}(s))$ where $\text{Pr}(s) \equiv \text{Pr}((s_j)_{j=1, \dots, m}) := (\min(m, \max(-m, s_j)))_{j=1, \dots, m}$ is the orthogonal projection of s onto the hypercube B_m^n . Thus we get $v_i \in$

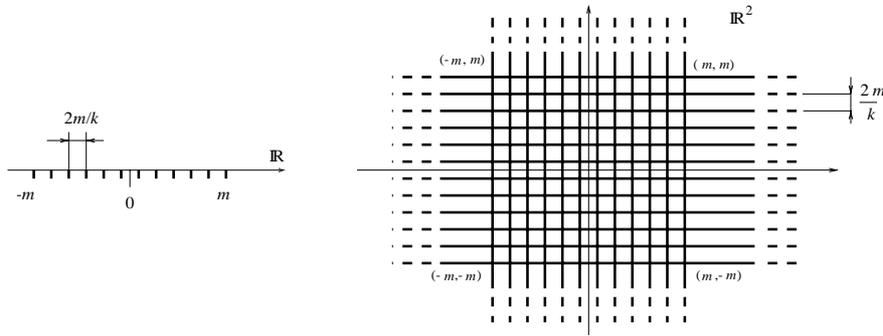


Fig. 1. A discretisation of \mathbb{R}^n for $n = 1$ and $n = 2$

$C_p(\mathbb{R}^n)$. Note that always $v_i(s) \geq 0$ and $\sum_{i=1}^{(k+1)^n} v_i(s) = 1$ for all $s \in \mathbb{R}^n$, which are actually the only essential properties of the chosen finite-element discretisation of \mathbb{R}^n we will need. Then, supposing $F(\cdot, s) \in C_{\mathcal{T}_l}(\Omega)$, we can construct a projector $P_{k,m} : H \rightarrow H$ defined by

$$(3.3) \quad [P_{k,m}h](x, s) := \sum_{i=1}^{(k+1)^n} h(x, s_i)v_i(s).$$

The projector $P_{k,m}$ is bounded if H is endowed with the norm (2.2) (here $\|P_{k,m}\|$ can be greater than 1) and commutes with P_l .

We now consider a discretisation parameter $d = (l, k, m)$ and create the discretisation of $\Omega \times \mathbb{R}^n$ as described above. The set of all d 's is naturally ordered by $d_1 \preceq d_2$ meaning that the discretisation created by d_2 is a refinement of the one created by d_1 . In particular, $d_1 = (l_1, k_1, m_1) \preceq d_2 = (l_2, k_2, m_2)$ just says that $l_1 \leq l_2$, $m_1 \leq m_2$, and $m_1k_2/(m_2k_1)$ is an integer. The set of all d 's is obviously directed by this ordering and thus we can use it to index nets; in particular, we will often work with nondecreasing sequences of d 's and then write $d \rightarrow \infty$ having equivalently just the meaning that $\min(l, k, m) \rightarrow \infty$. For given $d = (l, k, m)$ we will now consider a projector $P_d := P_l \circ P_{k,m} = P_{k,m} \circ P_l$, i.e.

$$(3.4) \quad [P_dh](x, s) = \sum_{i=1}^{(k+1)^n} \frac{1}{\text{meas}_n(\Delta)} \int_{\Delta} h(\xi, s_i)v_i(s)d\xi \text{ if } x \in \Delta \in \mathcal{T}_l.$$

By the formula (2.3), we can identify the set $\mathcal{Y}^p(\Omega; \mathbb{R}^n)$ with a subset of $Y_H^p(\Omega; \mathbb{R}^n)$ with H from (3.2). For brevity, we will not distinguish between $\mathcal{Y}^p(\Omega; \mathbb{R}^n)$ and its image in $Y_H^p(\Omega; \mathbb{R}^n)$.

The adjoint projector P_d^* obviously maps H^* into itself.

Lemma 1. We have $P_d^*Y_H^p(\Omega; \mathbb{R}^n) \subset \mathcal{Y}^p(\Omega; \mathbb{R}^n)$.

Proof. First, for $\nu \in \mathcal{Y}^p(\Omega; \mathbb{R}^n)$, one can see as in [R2, Examples 3.5.4–5] that

$$(3.5) \quad [P_d^* \nu]_x = \sum_{i=1}^{(k+1)^n} a_i(x) \delta_{s_i} \quad \text{with} \\ a_i(x) = \frac{1}{\text{meas}_n(\Delta)} \int_{\Delta} \int_{\mathbb{R}^n} v_i(s) \nu_{\xi}(ds) d\xi \quad \text{if } x \in \Delta \in \mathcal{T}_l,$$

where $\delta_s \in \text{rca}(\mathbb{R}^n)$ denotes the Dirac measure supported at $s \in \mathbb{R}^n$. Note that always $a_i(x) \geq 0$ and $\sum_{i=1}^{(k+1)^n} a_i(x) = 1$ because ν_x is a probability measure and because the basis functions satisfy $v_i(s) \geq 0$ and $\sum_{i=1}^{(k+1)^n} v_i(s) = 1$. Hence, $P_d^* \nu \in \mathcal{Y}^\infty(\Omega; \mathbb{R}^n)$ so that P_d^* maps $\mathcal{Y}^p(\Omega; \mathbb{R}^n)$ into itself.

Second, for a general $\eta \in Y_H^p(\Omega; \mathbb{R}^n)$, there is its so-called p -nonconcentrating modification $\nu \in \mathcal{Y}^p(\Omega; \mathbb{R}^n)$ defined (here even uniquely) by (2.3) for any $h \in C(\bar{\Omega}) \otimes C_p(\mathbb{R}^n)$; cf. [R2, Sect. 3.4]. In fact, (2.3) holds for any $h \in H$ having growth lesser than p , in particular for $P_d h$ which is even bounded. Then, for any $h \in H$,

$$\langle P_d^* \eta, h \rangle = \langle \eta, P_d h \rangle = \int_{\Omega} \int_{\mathbb{R}^n} [P_d h](x, s) \nu_x(ds) dx \\ = \int_{\Omega} \int_{\mathbb{R}^n} h(x, s) [P_d^* \nu]_x(ds) dx,$$

i.e. $P_d^* \eta = P_d^* \nu \in \mathcal{Y}^p(\Omega; \mathbb{R}^n)$ by (3.5). \square

Remark 1. In view of (3.5), the “discrete” Young measures from $P_d^* \mathcal{Y}^p(\Omega; \mathbb{R}^n)$ are element-wise constant (also called homogeneous) and can be supported only at a finite number, namely $(k+1)^n$, grid points of \mathbb{R}^n . The Dirac measures in (3.5) are called atoms.

Remark 2. The set of “approximate Young measures” $P_d^* \mathcal{Y}^p(\Omega; \mathbb{R}^n) = P_d^* Y_H^p(\Omega; \mathbb{R}^n)$ is convex; in fact, its convexity follows from the convexity of $\mathcal{Y}^p(\Omega; \mathbb{R}^n)$ and linearity of P_d^* even without knowing the explicit formula (3.5).

4. Convex approximation of relaxed problems

The above set of “discrete” Young measures $P_d^* \mathcal{Y}^p(\Omega; \mathbb{R}^n)$ can be used to construct a finite-dimensional approximation of various relaxed optimisation problems, in particular of (RP). Often, such relaxed problems have a convex structure, contrary to the “original” problems. The convexity of

$P_d^* \mathcal{Y}^p(\Omega; \mathbb{R}^n)$ then allows us to keep “as much” convex structure in the approximate problems as possible, which is very advantageous for both efficiency and reliability. For instance, in zero-sum games, a saddle-point search is much more effective if a convex/concave structure is preserved [R2, Sect. 7.3].

For any discretisation parameter $d = (l, k, m)$ we define an approximation of the relaxed problem (RP) by

$$(RP_d) \begin{cases} \text{minimise } \bar{\Phi}(u, \nu) := \int_{\Omega} \left[\int_{\mathbb{R}^n} F(x, s) \nu_x(ds) + G(x, u(x)) \right] dx, \\ \text{subject to } \int_{\mathbb{R}^n} s \nu_x(ds) = \nabla u(x) \text{ for a.a. } x \in \Omega, \\ u \in U_l, \nu \in P_d^* \mathcal{Y}^p(\Omega; \mathbb{R}^n), u|_{\partial\Omega} = u_D, \end{cases}$$

where $U_l := \{u \in W^{1,p}(\Omega); \forall \Delta \in \mathcal{T}_l : u|_{\Delta} \text{ is affine}\}$.

Proposition 2. *Suppose (I.1) and let u_D be compatible in the sense that $u_D = u_0|_{\partial\Omega}$ for some $u_0 \in U_{l_0}$, $l_0 \in N$. Then,*

- (i) *for d sufficiently large, (RP_d) always possesses a solution.*
- (ii) *For $d = (l, k, m) \rightarrow \infty$ with $l \rightarrow \infty$ sufficiently slowly in comparison with $k, m \rightarrow \infty$, one has $\min(RP_d) \rightarrow \min(RP)$ and every cluster point of sequence of solutions to (RP_d) solves (RP).*

Proof. For any $d = (l, k, m) \succeq d_0 = (l_0, k_0, m_0)$ with $m_0 \geq \max_{i=1, \dots, n} \|\partial u_0 / \partial x_i\|_{L^\infty(\Omega)}$, one can see that the set of feasible pairs (u, ν) for (RP_d) is nonempty, contains a pair (u_0, ν^0) with a suitable ν^0 supported at vertices of the hypercube B_m^n and satisfies $\int_{\mathbb{R}^n} s \nu_x^0(ds) = \nabla u_0(x)$. Then, the existence of solutions to (RP_d) claimed in (i) follows by standard coercivity, continuity, and compactness arguments; note that the feasible set of (RP_d) is finite-dimensional.

Let us now consider a sequence $u_l \in U_l$ with $u_l|_{\partial\Omega} = u_D$ such that $(u_l, \delta(\nabla u_l))$ converges to a solution of (RP); cf. [R2, Proposition 5.5.1] with a modification for the Dirichlet boundary conditions. This shows that $\min(RP_{(l, \infty, \infty)}) \rightarrow \min(RP)$; for the definition of $(RP_{(l, \infty, \infty)})$ see Sect. 6 below. Unfortunately, the pair $(u_l, \delta(\nabla u_l))$ need not be feasible for (RP_d) , yet we can modify slightly $\delta(\nabla u_l)$ by replacing it with a suitable ν^l so that (u_l, ν^l) is admissible for (RP_d) and $|\bar{\Phi}(u_l, \nu^l) - \Phi(u_l)| \leq \varepsilon(l, k, m)$ is arbitrarily small provided k and m are sufficiently large (depending possibly on l); cf. also (6.3). Thus we obtain convergence of minima claimed at the point (ii).

The convergence of solutions in terms of cluster points then follows by standard compactness, continuity, and coercivity arguments; we omit the details. \square

To guarantee the smoothness of $\bar{\Phi}$ with respect to the geometry of $W^{1,p}(\Omega) \times L^\infty(\Omega; \text{rca}(\mathbb{R}^n))$, we must suppose the differentiability of $G(x, \cdot)$ only; let $G'(x, \cdot)$ denote its derivative. Namely, we want to assure $u \mapsto \int_\Omega G(x, u(x))dx : W^{1,p}(\Omega) \rightarrow \mathbb{R}$ to be a C^1 -function. For this, it suffices to assume (cf. [R4] or [R2, Sect. 5.3]) that

$$(4.1) \quad \exists a_1 \in L^{q/(q-1)}(\Omega), \exists b \in \mathbb{R} : |G'(x, u)| \leq a_1(x) + b|u|^{q-1},$$

$$(4.2) \quad \exists a_2 \in L^{q/(q-2)}(\Omega), \exists b \in \mathbb{R} : |G'(x, u_1) - G'(x, u_2)| \leq (a_2(x) + b|u_1|^{q-2} + b|u_2|^{q-2})|u_1 - u_2|;$$

note that (4.2) requires $q \geq 2$ which is always possible provided $p > 2n/(n + 2)$.

Using convexity of the set of feasible elements of (RP_d) and the supposed smoothness of $\bar{\Phi}$, one can derive the necessary and sufficient optimality conditions for a solution to (RP_d) in the form of a Weierstrass-type maximum principle.

Proposition 3. *Suppose (1.1), (4.1), and (4.2), and let (u, ν) solve (RP_d) . Then there is some $\lambda \in L^\infty(\Omega; \mathbb{R}^n)$ element-wise constant such that the maximum principle*

$$(4.3) \quad \max_{s \in \mathbb{R}^n} \mathcal{H}_\lambda^d(x, s) = \int_{\mathbb{R}^n} \mathcal{H}_\lambda^d(x, s) \nu_x(ds) \quad \text{for a.a. } x \in \Omega$$

and the identity

$$(4.4) \quad \int_\Omega (\lambda \cdot \nabla z - G'(u) \cdot z) dx = 0 \quad \forall z \in U_l, z|_{\partial\Omega} = 0,$$

are satisfied, where the “discrete Hamiltonian” $\mathcal{H}_\lambda^d \in H$ is given by

$$(4.5) \quad \mathcal{H}_\lambda^d := P_d(\lambda \otimes \text{id} - F).$$

Moreover, if $G(x, \cdot)$ is convex for a.a. $x \in \Omega$, then conversely every (u, ν) solves (RP_d) provided it is feasible for (RP_d) and satisfies (4.3)–(4.5) for some $\lambda \in L^\infty(\Omega; \mathbb{R}^n)$.

Proof. See [R2, Proposition 5.5.3] with the modification of $P_d : H \rightarrow H$ given now by (3.4) for corresponding optimality conditions for $(RP_{H,d})$ defined as (RP_H) but with $P_d^* Y_H^p(\Omega; \mathbb{R}^n)$ instead of $Y_H^p(\Omega; \mathbb{R}^n)$. This gives then (4.3)–(4.4) for $(RP_{H,d})$. However, the problems (RP_d) and $(RP_{H,d})$ are in fact identical because $P_d^* Y_H^p(\Omega; \mathbb{R}^n) = P_d^* \mathcal{Y}^p(\Omega; \mathbb{R}^n)$, see Lemma 1.

Supposing $G(x, \cdot)$ convex, $\bar{\Phi}$ is convex with respect to the geometry of $W^{1,p}(\Omega) \times L^\infty(\Omega; \text{rca}(\mathbb{R}^n))$ and optimality conditions (4.3)–(4.5) are also sufficient. Note that λ (possibly non-constant along elements $\Delta \in \mathcal{T}_l$) can be replaced by its element-wise averages without affecting (4.3)–(4.5). \square

Let us remark that λ is the Lagrange multiplier with respect to the constraint $\int_{\mathbb{R}^n} s\nu(ds) = \nabla u$. These sufficient optimality conditions gives the theoretical basis for our adaptive-refinement scheme. Another important fact is that, due to a special construction of the projector P_d , the maximum principle is satisfied if (and only if), on each element $\Delta \in \mathcal{T}_l$, it holds $\mathcal{H}_\lambda^d(x, s_i) \leq \int_{\mathbb{R}^n} \mathcal{H}_\lambda^d(x, s)\nu_x(ds)$ for every grid point s_i with $i = 1, \dots, (k + 1)^n$. Thus (4.3) can be checked in a finite number of points.

5. An active-set strategy in s -variable

The above mentioned substantial advantages of convex approximation (RP_d) are deteriorated by a need of a very large dimensionality of the approximate relaxed problems. Realize that the dimension of $P_d^*\mathcal{Y}^p(\Omega; \mathbb{R}^n)$ with $d = (l, k, m)$ is $\text{card}(\mathcal{T}_l)(k + 1)^n$.

The maximum principle (4.3) predicts generically that the solutions to (RP_d) will be typically supported not at all $(k + 1)^n$ grid points but only at a very few ones on each element. This is a basis for our adaptive technique where we approximate (RP_d) by $(\text{RP}_{d,A})$, requiring the involved discrete Young measures to be supported only on a (suitably) prescribed set $A \subset \Omega \times \mathbb{R}^n$. The support of a Young measure ν is defined as

$$\text{Supp}(\nu) := \{(x, s) \in \Omega \times \mathbb{R}^n; s \in \text{supp}(\nu_x)\}$$

where $\text{supp}(\nu_x) \subset \mathbb{R}^n$ is the support of the measure $\nu_x \in \text{rca}(\mathbb{R}^n)$ in the usual sense; this defines $\text{Supp}(\nu)$ up to zero-measure sets on Ω . Then the problem $(\text{RP}_{d,A})$ reads

$$(\text{RP}_{d,A}) \quad \begin{cases} \text{minimise } \bar{\Phi}(u, \nu) := \int_{\Omega} \left[\int_{\mathbb{R}^n} F(x, s)\nu_x(ds) + G(x, u(x)) \right] dx, \\ \text{subject to } \int_{\mathbb{R}^n} s\nu_x(ds) = \nabla u(x) \text{ for a.a. } x \in \Omega, \\ u \in U_l, \nu \in P_d^*\mathcal{Y}^p(\Omega; \mathbb{R}^n), \text{Supp}(\nu) \subset A, u|_{\partial\Omega} = u_D. \end{cases}$$

Proposition 4. *Suppose (1.1), (4.1), and (4.2) and let (u, ν) solve (RP_d) with λ being the corresponding multiplier. Let $A \subset \Omega \times \mathbb{R}^n$ satisfy*

$$(5.1) \quad \left\{ (x, s) \in \Omega \times \mathbb{R}^n; s \text{ is a grid point, } \mathcal{H}_\lambda^d(x, s) = \max_{\tilde{s} \in \mathbb{R}^n} \mathcal{H}_\lambda^d(x, \tilde{s}) \right\} \subset A.$$

Then, every solution to $(\text{RP}_{d,A})$ solves also (RP_d) .

Proof. Obviously, every feasible point $(\tilde{u}, \tilde{\nu})$ of $(\text{RP}_{d,A})$ solves also (RP_d) provided $\bar{\Phi}(\tilde{u}, \tilde{\nu}) \leq \min(\text{RP}_d)$.

Each solution (u, ν) to (RP_d) is simultaneously a feasible point for $(\text{RP}_{d,A})$ because it satisfies the maximum principle (4.3) and so $\text{Supp}(\nu) \subset A$ for our special choice of A . Therefore $\min(\text{RP}_{d,A}) \leq \min(\text{RP}_d)$.

Thus, having a solution $(\tilde{u}, \tilde{\nu})$ to $(\text{RP}_{d,A})$, we have proved $\bar{\Phi}(\tilde{u}, \tilde{\nu}) = \min(\text{RP}_{d,A}) \leq \min(\text{RP}_d)$ and hence $(\tilde{u}, \tilde{\nu})$ solves also (RP_d) . \square

If we knew a priori the correct “discrete Hamiltonian” \mathcal{H}_λ^d , we could immediately use Proposition 4 to solve a presumably much “smaller” problem $(\text{RP}_{d,A})$ instead of (RP_d) . As we can at most guess \mathcal{H}_λ^d , we need a tool how to verify a posteriori whether a solution to $(\text{RP}_{d,A})$ solves also (RP_d) and need a certain effective method to get an enough good guess of the set A . Let us first deal with the first task, the second will be addressed in Sect. 6.

Proposition 5. *Suppose (1.1), (4.1) and (4.2). Let $G(x, \cdot)$ be convex for a.a. $x \in \Omega$, let (u, ν) solve $(\text{RP}_{d,A})$ with a corresponding Lagrange multiplier λ . Suppose that the discrete maximum principle (4.3)–(4.5) is satisfied. Then (u, ν) solves (RP_d) .*

Proof. Any (u, ν) feasible for $(\text{RP}_{d,A})$ is also feasible for (RP_d) . Moreover, (u, ν) satisfies all the optimality conditions (4.3)–(4.5) because the adjoint equation (4.4), which is independent of A if formulated for $(\text{RP}_{d,A})$, is satisfied for the obtained multiplier λ and because the satisfaction of the maximum principle (4.3) is assumed. Since $G(x, \cdot)$ is convex, the necessary optimality conditions (4.3)–(4.5) are also sufficient for (RP_d) , see Proposition 3. Therefore (u, ν) solves (RP_d) . \square

Remark 3. For n small the approximate relaxed problem (RP_d) can be solved directly; see [R1] for variational problems or [R2, Sects. 4.3 and 7.3] for an optimal control and a zero-sum game problems with $n = 1$. However, for n substantially larger than 1 such methods cannot be effectively used and ultimately a certain adaptivity idea is needed.

6. Estimation of the Hamiltonian in a multilevel technique

Let us now focus to a good guess of the active set A . Inspiring by the condition (5.1), we put

$$(6.1) \quad A = \left\{ (x, s) \in \Omega \times \mathbb{R}^n : h(x, s) \geq \max_{\tilde{s} \in \mathbb{R}^n} h(x, \tilde{s}) - \varepsilon \right\},$$

where $\varepsilon > 0$ is some tolerance and h will play a role of some guess of the “correct” Hamiltonian \mathcal{H}_λ^d .

Proposition 6. *Suppose $\|h - \mathcal{H}_\lambda^d\|_{L^\infty(\Omega \times S)} \leq \varepsilon/2$ for some solution (u, ν) of (RP_d) with a corresponding Lagrange multiplier λ and for some $S \subset \mathbb{R}^n$ so large that, for a.a. $x \in \Omega$,*

$$(6.2) \quad \text{Argmax } \mathcal{H}_\lambda^d(x, \cdot) \cup \text{Argmax } h(x, \cdot) \subset S.$$

Let A be chosen according (6.1). Then every solution to $(\text{RP}_{d,A})$ solves (RP_d) .

Proof. In view of Proposition 4, it suffices to show that A from (6.1) satisfies (5.1). Let us take ν such that $\mathcal{H}_\lambda^d(x, s) = \max_{\tilde{s} \in \mathbb{R}^n} \mathcal{H}_\lambda^d(x, \tilde{s})$ for any $(x, s) \in \text{Supp}(\nu)$. Then, by (6.2), $\text{Supp}(\nu) \subset \Omega \times S$ up to a zero measure set. Taking $(x, s) \in \text{Supp}(\nu)$ and using again (6.2), we thus have $s \in S$ and therefore

$$h(x, s) \geq \mathcal{H}_\lambda^d(x, s) - \frac{\varepsilon}{2} = \max_{\tilde{s} \in \mathbb{R}^n} \mathcal{H}_\lambda^d(x, \tilde{s}) - \frac{\varepsilon}{2} \geq \max_{\tilde{s} \in \mathbb{R}^n} h(x, \tilde{s}) - \varepsilon.$$

Thus we showed that (5.1) is satisfied for A from (6.1). \square

The above assertion suggests to look for a good guess of (u, ν) and of λ to get a good guess of the Hamiltonian \mathcal{H}_λ^d . In a two-level approach, we consider two discretisations $d_0 := (l, k_0, m) \preceq (l, k_1, m) =: d_1$; note that it uses the same spatial triangulation (the x -discretisation is unchanged). Our argument will be supported by the problem created as a limit for $k \rightarrow \infty$, which for $m \in \mathbb{R}^+ \cup \{+\infty\}$ motivates the auxiliary problem

$$(\text{RP}_{(l, \infty, m)}) \quad \begin{cases} \text{minimise } \bar{\Phi}(u, \nu) := \int_{\Omega} \left[\int_{\mathbb{R}^n} F(x, s) \nu_x(ds) + G(x, u(x)) \right] dx, \\ \text{subject to } \int_{\mathbb{R}^n} s \nu_x(ds) = \nabla u(x) \text{ for a.a. } x \in \Omega, \\ u \in U_l, \nu \in P_l^* \mathcal{Y}^p(\Omega; \mathbb{R}^n), \\ \text{Supp}(\nu) \subset \Omega \times B_m^n, u|_{\partial\Omega} = u_D. \end{cases}$$

Proposition 7. *Suppose (1.1), (4.1) and (4.2) and fix l .*

- (i) *For m sufficiently large, we have $\min(\text{RP}_{(l, \infty, m)}) = \min(\text{RP}_{(l, \infty, \infty)})$ and, moreover, $\min(\text{RP}_d) \equiv \min(\text{RP}_{(l, k, m)}) \rightarrow \min(\text{RP}_{(l, \infty, m)})$ for $k \rightarrow \infty$.*
- (ii) *If, in addition, $F(x, \cdot)$ is smooth and $G(x, \cdot)$ is strictly convex, (u_k, ν_k) solves $(\text{RP}_{(l, k, m)})$ with λ_k being the corresponding multiplier, and if (u, ν) solves $(\text{RP}_{(l, \infty, m)})$ with λ being the corresponding multiplier, then $u_k \rightarrow u$ and $\lambda_k \rightarrow \lambda$.*

Proof. Take a solution (u, ν) to $\min(\text{RP}_{(l, \infty, \infty)})$ and the corresponding multiplier $\lambda \in L^\infty(\Omega; \mathbb{R}^n)$ element-wise constant. Likewise in Proposition 3, one

can see that ν_x is supported on the set of maximisers of $\mathcal{H}_\lambda^d := P_l(\lambda \otimes \text{id} - F)$ with P_l defined by (3.1); cf. also [R2, Sect. 5.5]. As F has a super-linear growth by (1.1) and λ is essentially bounded over Ω , this support is a priori bounded and so contained in B_m^n for m large enough. Then, $\min(\text{RP}_{(l,\infty,m)})$ is equivalent to $\min(\text{RP}_{(l,\infty,\infty)})$.

To show $\min(\text{RP}_{(l,k,m)}) \rightarrow \min(\text{RP}_{(l,\infty,m)})$, we will seek (u^k, ν^k) feasible for $(\text{RP}_{(l,k,m)})$ and converging to some (u, ν) solving $(\text{RP}_{(l,\infty,m)})$. We will construct (u^k, ν^k) by taking $u^k = u$ and by a suitable modification of $P_{k,m}^* \nu$. Note that unfortunately $(u, P_{k,m}^* \nu)$ need not be feasible for $(\text{RP}_{(l,k,m)})$. We choose ν^k in the form

$$(6.3) \quad \nu^k = (1 - \varepsilon_k) P_{k,m}^* \nu + \varepsilon_k \nu^{k,m}$$

with some $\varepsilon_k > 0$ small and with $\nu^{k,m} \in P_d^* \mathcal{Y}(\Omega; \mathbb{R}^n)$. Besides, we need that $\int_{\mathbb{R}^n} s \nu_x^k(ds) = \nabla u(x)$, so that $y_{k,m} := \int_{\mathbb{R}^n} s \nu_x^{k,m}(ds) = \int_{\mathbb{R}^n} s \nu_x^k(ds) + (1 - 1/\varepsilon_k) \delta_{k,m}$ with $\delta_{k,m} := \int_{\mathbb{R}^n} s [(P_{k,m}^* \nu)_x - \nu_x](ds)$. As $\|h - P_{k,m} h\|_{L^1(\Omega; C(B_m^n))} \rightarrow 0$ for $k \rightarrow \infty$, $P_{k,m}^* \nu \rightarrow \nu$ weakly* because $\langle \nu - P_{k,m}^* \nu, h \rangle = \langle \nu, h - P_{k,m} h \rangle \leq \|\nu\| \|h - P_{k,m} h\|_{L^1(\Omega; C(B_m^n))} \rightarrow 0$ for any $h \in H$ and ν with $\text{Supp}(\nu) \subset \Omega \times B_m^n$. Thus $\delta_{k,m} \rightarrow 0$. It allows us to choose ε_k converging to 0 sufficiently slowly so that also $(1 - 1/\varepsilon_k) \delta_{k,m} \rightarrow 0$. Supposing that m was taken even so large that $\text{supp}(\nu_x)$ is contained in the interior of B_m^n , we have $y_{k,m}(x) \in B_m^n$ provided k is large enough. Then there exists a Young measure $\nu^{k,m}$ supported on the grid points of B_m^n and satisfying $y_{k,m} := \int_{\mathbb{R}^n} s \nu_x^{k,m}(ds)$. From (6.3), one gets obviously $\nu^k \in P_d^* Y_H^p(\Omega; \mathbb{R}^n)$ and $\nu^k \rightarrow \nu$ weakly* because $P_{k,m}^* \nu \rightarrow \nu$ and because $\varepsilon_k \rightarrow 0$.

For (u, ν) a solution to $(\text{RP}_{(l,\infty,\infty)})$, u solves the convexified problem, i.e. minimises $\text{co}_l \Phi(u) := \int_{\Omega} [P_l F]^{**}(x, \nabla u) + G(x, u) dx$ over U_l , where the double star denotes the convex hull. As $G(x, \cdot)$ is assumed strictly convex, so is also $\text{co}_l \Phi$ and thus u is determined uniquely.

Let us now consider a solution (u_k, ν_k) to $(\text{RP}_{(l,k,m)})$ with the corresponding multiplier λ_k . By Proposition 2(ii), every cluster point u of the sequence $\{u_k\}$ must minimise $\text{co}_l \Phi$. As u is determined uniquely, we get $u_k \rightarrow u$ strongly in $W^{1,p}(\Omega)$ because u and u_k live in a finite-dimensional space U_l of element-wise affine functions on the fixed triangulation \mathcal{T}_l .

Then, from (4.3) with $d = (l, \infty, \infty)$ derived in [R2, Proposition 5.3.2], we can see that ν_x minimises $s \mapsto [P_l F](x, s) - \lambda(x) \cdot s$ so that its first momentum $\int_{\mathbb{R}^n} s \nu_x(ds) = \nabla u(x)$ minimises the convexified functional, i.e. $s \mapsto [P_l F]^{**}(x, s) - \lambda(x) \cdot s$. As $F(x, \cdot)$ is smooth, also $[P_l F](x, \cdot)$ is smooth, and by [F, Corollary 3] so is $[P_l F]^{**}(x, \cdot)$. Here, we used also the

super-linear growth of $[P_l F](x, \cdot)$. Hence

$$(6.4) \quad \frac{\partial}{\partial s} [P_l F]**(x, \nabla u(x)) - \lambda(x) = 0.$$

As u is unique, so is λ .

Moreover, from (4.3) for $d = (l, k, m)$ one can derive similarly as (6.4) that

$$(6.5) \quad \lambda_k(x) \in \partial_s [P_d F]**(x, \nabla u_k(x)),$$

where ∂_s denotes the sub-differential with respect to the s -variable. As $P_d F = P_{k,m} P_l F \rightarrow P_l F$ pointwise in $\Omega \times B_m^n$ for $k \rightarrow \infty$, the same holds for the convex hulls. As $\nabla u_k(x) \rightarrow \nabla u(x)$ and $\nabla u(x)$ lives in the interior of B_m^n , from (6.4) and (6.5) we infer $\lambda_k(x) \rightarrow \lambda(x)$ for a.a. $x \in \Omega$. \square

By Proposition 7, we can presume that the problems (RP_{d_0}) and (RP_{d_1}) do not differ from each other too much if d_0 is already large enough. Then we can take $h = \mathcal{H}_{\lambda_0}^{d_0}$ with u_0 and λ_0 corresponding to a “smaller” problem (RP_{d_0}) as a good guess of the Hamiltonian $\mathcal{H}_{\lambda_1}^{d_1}$ for (RP_{d_1}) .

Proposition 8. *Suppose (1.1), (4.1), and (4.2). Let $F(x, \cdot)$ be smooth, $G(x, \cdot)$ be strictly convex, and let m be sufficiently large so that the equivalence in Proposition 7(i) holds.*

(i) *For any $\varepsilon > 0$, we have*

$$(6.6) \quad \begin{aligned} &\exists k_0 \forall d_1 = (l, k_1, m) \succeq d_0 \\ &= (l, k_0, m) : \|\mathcal{H}_{\lambda_0}^{d_0} - \mathcal{H}_{\lambda_1}^{d_1}\|_{L^\infty(\Omega \times B_m^n)} \leq \varepsilon/2, \end{aligned}$$

where λ_1 and λ_0 are multipliers corresponding to some solutions to (RP_{d_1}) and (RP_{d_0}) , respectively.

(ii) *For any k_0 , there is S bounded such that $S \supset \text{Argmax } \mathcal{H}_\lambda^d(x, \cdot)$ provided $d \succeq (l, k_0, m)$ and λ is the multiplier corresponding to a solution (u, ν) of (RP_d) .*

Proof. By Proposition 7(ii), $\{\lambda_k\}_{k \in \mathbb{N}}$ is a Cauchy sequence, so that also the sequence $\{\lambda_k \otimes \text{id} - F\}_{k \in \mathbb{N}}$ is Cauchy in $L^\infty(\Omega \times S)$ for any S bounded, in particular for $S = B_m^n$. Then also $\{P_{k,m} P_l(\lambda_k \otimes \text{id} - F)\}_{k \in \mathbb{N}} = \{\mathcal{H}_{\lambda_k}^{(l,k,m)}\}_{k \in \mathbb{N}}$ is Cauchy in $L^\infty(\Omega \times B_m^n)$, from which (6.6) follows.

As $F(x, \cdot)$ has a super-linear growth and $\{\lambda_k\}_{k \in \mathbb{N}}$ is Cauchy, $\text{Argmax } \mathcal{H}_{\lambda_k}^{(l,k,m)}$ is a priori bounded provided k is large enough. \square

Corollary 1. *For any $\varepsilon > 0$ there is $k_0 \in \mathbb{N}$ such that, for all $d_1 = (l, k_1, m) \succeq d_0 = (l, k_0, m)$, any solution of $(\text{RP}_{d_1, A})$ solve also (RP_{d_1}) provided*

$$(6.7) \quad A = \left\{ (x, s) \in \Omega \times \mathbb{R}^n; \mathcal{H}_{\lambda_0}^{d_0}(x, s) \geq \max_{\tilde{s} \in \mathbb{R}^n} \mathcal{H}_{\lambda_0}^{d_0}(x, \tilde{s}) - \varepsilon \right\}$$

and λ_0 is the multiplier corresponding to some solution of (RP_{d_0}) .

Proof. Take k_0 sufficiently large so that (6.6) holds, then take S so large that $S \supset \text{Argmax } \mathcal{H}_{\lambda}^d(x, \cdot)$ provided $d \succeq (l, k_0, m)$ (see Proposition 8(ii)) and eventually use Proposition 6 with $h = \mathcal{H}_{\lambda_0}^{d_0}$ and $\mathcal{H}_{\lambda}^d = \mathcal{H}_{\lambda_1}^{d_1}$. \square

The above idea yields a multi-level algorithm with a chain $d_0 \preceq d_1 \preceq d_2 \preceq \dots \preceq d_j \preceq \dots \preceq d_J$ with $d_j = (l, k_j, m)$. Thus we come to the following conceptual algorithm.

- (1) **Initialisation:** Choose $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{J-1} > 0$, put $\varepsilon := \varepsilon_0, j := 0, h := 0$.
- (2) **Activation:** compute A from (6.1).
- (3) **Optimisation routine:** solve $(\text{RP}_{d, A})$. It gives $(u, \nu) \in U_l \times P_d^* \mathcal{Y}^p(\Omega; \mathbb{R}^n)$ and the multiplier $\lambda \in L^\infty(\Omega; \mathbb{R}^n)$.
- (4) If (4.3)–(4.5) is satisfied on the level d_j , then go to (6), else continue.
- (5) **Correction:** increase the tolerance by putting $\varepsilon := 2\varepsilon$. Go to (2).
- (6) If $j < J$ then continue, else end.
- (7) **Refinement:** put $j := j + 1, h := \mathcal{H}_{\lambda}^{d_j-1}, \varepsilon := \varepsilon_j$. Go to (2).

The flow chart of this algorithm is schematically drawn on Fig. 2.

Remark 4. Since $h = 0$ is taken for initialisation, on the coarsest grid d_0 the algorithm puts $A = \Omega \times B_m^n$. As the maximum principle (4.3)–(4.5) forms necessary optimality conditions, the algorithm always performs the “refinement” after the very first run of the “optimisation routine”.

Remark 5. In the “correction” step, always the tolerance is made greater so that, after a finite number of steps, it achieves such a value that (6.1) activates all grid points, i.e. $A = \Omega \times B_m^n$ is accepted. Thus the algorithm eventually must come to a satisfaction of the discrete maximum principle. As this maximum principle is supposed also sufficient, the algorithm must after a finite number of steps come to a situation that all assumptions of Proposition 4 are satisfied and thus it found a solution on the discretisation level d_j .

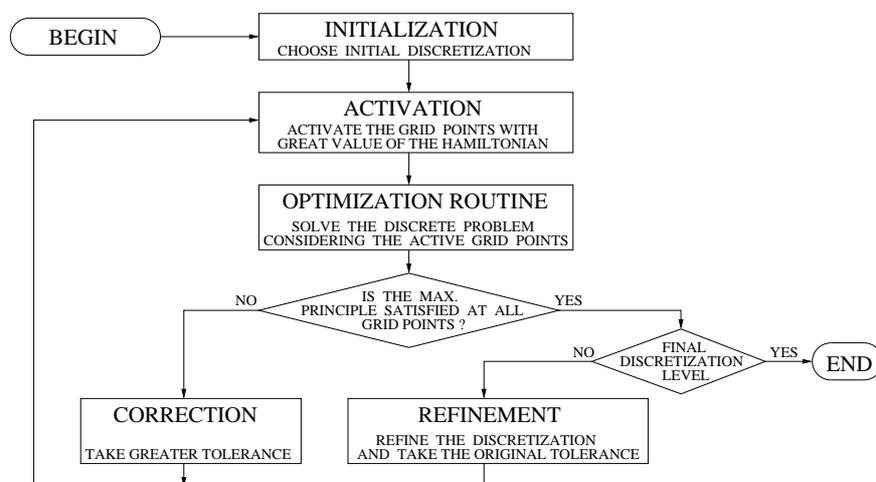


Fig. 2. A conceptual algorithm for adaptive refinement in s -variable

Remark 6. For d_j sufficiently large, we explained in Corollary 1 that $\mathcal{H}_{\lambda_{j-1}}^{d_j-1}$ represents a good estimate of the Hamiltonian $\mathcal{H}_{\lambda_j}^{d_j}$, and thus the correction step is supposed to be effective only few times (if at all) for the discretisation d_j . The fact that $\varepsilon > 0$ can be chosen small for large k_0 claimed in Corollary 1 is important, indicating high expected efficiency of $(\text{RP}_{d,A})$ in comparison with (RP_d) because, for $\varepsilon = 0$, only those atoms are activated (i.e. contained in A) where $\mathcal{H}_{\lambda_1}^{d_1}$ attains its maximum.

Remark 7. A-posteriori verification of the maximum principle (4.3)–(4.5) and activation of the set A , which are the only operations to be made on the whole grid of the discretisation d_j , are supposed “cheap” in comparison with the execution of the optimisation routine so that they will not delay substantially the calculations; this naturally need not be true for $(k+1)^n$ extremely large, i.e. for a very fine discretisations in the s -variable.

7. Numerical example

In this section we report on numerical experience and illustrate our algorithm in the shifted Tartar’s broken extremal example [NW2], namely

$$(7.1) \quad \begin{cases} \text{minimise} & \int_0^1 [(u'(x)^2 - 1)^2 + (u(x) - g(x))^2] dx, \\ \text{subject to} & u \in W^{1,p}(0,1), \quad u(0) = z_0, \quad u(1) = z_1. \end{cases}$$

This corresponds to (P) with $n = 1$, $\Omega := (0,1)$, $p = 4$, $F(x,s) := (s^2 - 1)^2$ and $G(x,u) := (u - g(x))^2$. For $g(x) := -\frac{3}{128}(x - x_b)^5 - \frac{1}{3}(x - x_b)^3$,

$z_0 := g(0)$, and $z_1 := (\frac{25}{24} - \frac{1}{12}x_b + \frac{1}{24}x_b^2)(1 - x_b)$, the relaxed problem (RP) has the unique solution [NW1]

$$u(x) = \begin{cases} g(x), \\ \frac{1}{24}(x - x_b)^3 + (x - x_b), \end{cases}$$

$$\nu_x = \begin{cases} \frac{1-u'(x)}{2}\delta_{-1} + \frac{1+u'(x)}{2}\delta_1 & \text{for } x \in (0, x_b), \\ \delta_{u'(x)} & \text{for } x \in (x_b, 1). \end{cases}$$

Since $G(x, \cdot)$ is quadratic, the discrete relaxed problem (RP_d) has a linear/quadratic structure. Thus, finite direct solvers can in principle be applied to solve it. However, our numerical experiments proved that this direct approach appears limited to very small number of unknowns hence we will focus to solve $(RP_{d,A})$ as proposed in Sect. 6.

In the implementation of the proposed algorithm we employed a uniform grid in both Ω and $[a, b] \supset B_m^1$ (the space of atoms) for m sufficiently large, cf. Proposition 7(i). The spatial grid of mesh-size 2^{-l} is defined by the $N := 2^l$ elements, so that $\mathcal{T}_l = \{\Delta = 2^{-l}(i - 1, i); i = 1, \dots, 2^l\}$. For each $\Delta \in \mathcal{T}_l$, the discrete set of atoms (i.e. grid points on $[a, b]$) at level $j = 0, 1, \dots, J$ is defined as $a + 2^{-j}(b - a)i/K$ for $0 \leq i \leq k_j := 2^j K$. This yields a presumably very coarse uniform grid $a, a + (b - a)/K, \dots, b$ of $k_0 + 1$ grid points between a and b for $j = 0$ and then halves each interval successively as j increases to the finest grid for $j = J$ with $k_J + 1$ grid points.

To describe the set of active atoms in our MATLAB realization, we defined a sparse matrix which, for each element $\Delta = 2^{-l}(i - 1, i)$ characterised the set of active atoms as non-zero entries in `atoms(i, :)`. For instance, the initialisation is given by `atoms=sparse(ones(N, 1) * 1 : K + 1)`.

For any reasonable set of active grid points (we will comment on that below), the optimisation problem `qp` consists of minimising

$$x \cdot Hx + f \cdot x \quad \text{subject to} \quad Ax \leq b,$$

where $A \cdot x \leq b$ is understood component-wise; here we adopted notation from the MATLAB `qp` manual which differs from previous sections.

Let the first $N + 1$ component of the D -dimensional unknown vector x describe the discrete displacements at the grid points and thereafter the coefficients of the $D - N - 1 = \text{nnz}(\text{atoms})$ active atoms. The finite element basis consists of $N + 1$ hat functions which vanish outside of $h(i - 1, i + 1)$ and equal one at ih . Then, the leading $(N + 1) \times (N + 1)$ -block of the $D \times D$ -matrix H is the tridiagonal matrix of the exact $L^2(0, 1)$ -products of the $N + 1$ hat functions with a typical local mass matrix $[2, 1; 1, 2]h/6$. The corresponding first $N + 1$ entries of f are the $L^2(0, 1)$ -products of hat functions with g evaluated by an elementwise 3-node Gaussian quadrature

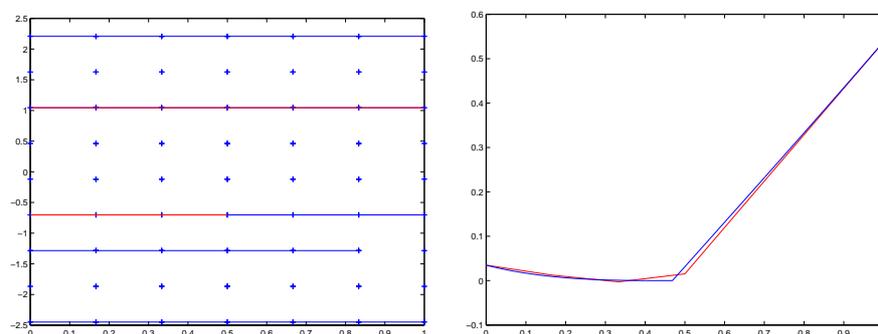


Fig. 3. Support of Young measure (left) and a displacement (right) on the level $j = 3$

rule. The remaining entries vanish in H . Thus, H is not positive definite. In order to regularise the problem, we added 10^{-5} times the $D \times D$ -unit matrix to the exact matrix H . The remaining contributions in f are given by the nonlinear function $(s^2 - 1)^2$ evaluated with s replaced by the atom linked to the current component.

The $2N \times D$ -matrix A as the $2N$ -dimensional vector b characterise the equality side restrictions: the element-wise constraints force the sum of the components of x to equal one and the measure to have the mean value equal the the slope on each element. Hence, the only non-zero entries of A and f are $f_{2i-1} = A_{2i-1,\ell} = 1$, $A_{2i,i} = 1/h = -A_{2i,i+1}$, and $A_{2i,\ell} = s$ if the i -th component of x corresponds to an atom s in the i -th element.

The remaining restrictions are directly incorporated by inequalities $0 \leq x_{N+2}, x_{N+3}, \dots, x_D \leq +\infty$ (the measure-coefficients are non-negative) and, finally, $u_D(0) \leq x_1 \leq u_D(0)$ and $u_D(1) \leq x_{N+1} \leq u_D(1)$ according to the geometric boundary condition.

The MATLAB `qp` routine provides us with a solution to the problem `qp` as well as with the Lagrange multiplier λ utilised in the definition of the Hamiltonian (4.5). Therefore, the proposed algorithm can easily be realized in MATLAB.

In the following, we comment on a numerical experiment obtained for $\epsilon_j = 2^{-j}/100$ and $a = -2 - \pi/7$, $b = 2 + \exp(1)/13$, $x_b = 1/2 - \pi/100$, $h = 1/3$, and $K = 1$. In Fig. 3 and 4 we show on the left side set of active atoms as thick lines and on the right the exact and approximate displacements obtained for our algorithm for $j = 3$ (in Fig. 3 with 6 elements and 9 atoms) and $j = 6$ (in Fig. 4 with 24 elements and 65 atoms).

The plus signs at the positions $(2^{-l}i, s)$ and $(2^{-l}(i-1), s)$ in the left plots illustrate the fine set of possible atoms s on the element i at the level j . There are two kinds of horizontal lines which link two of those plus signs on elements. The thinner lines trough midpoints such as $(2^{-l}(i-1/2), s)$ indicate that the atom s is active on the element $2^{-l}(i-1, i)$ while the

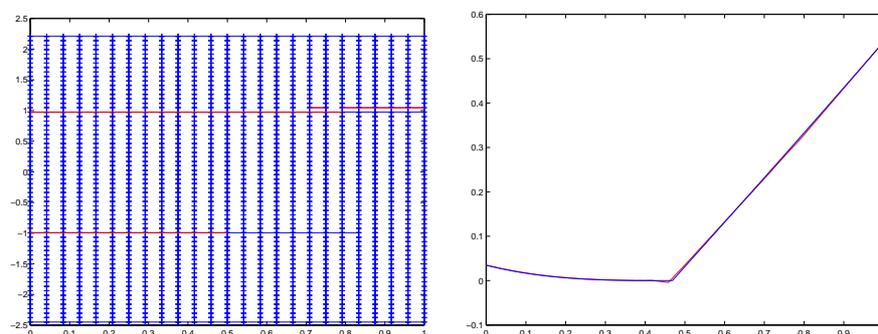


Fig. 4. Support of Young measure (left) and a displacement (right) on the level $j = 6$

thicker line implies, in addition, that the coefficient related to the atom s is significant (i.e. “actually” positive, namely larger than 10^{-3}).

The plus signs in the right displacement plots indicate the exact solution at the nodes of spacial grid.

The refinement in the s -variable (we halved the distance of atoms in each level) can naturally be combined with a refinement in space (we furthermore halved the elements at every second level). A statistical report on the algorithm’s practical performance is given in Table 1 where, for each level j , we displayed the total number of atoms and their average per element. In the second last column we show the sum $\text{cpu}(j)$ over all dimensions D which are considered in the termination of the qp sub-routine until the level j is reached. Assuming that the CPU-time for the algorithm is mainly determined by the qp applications, we count only the effort for those calls which is monotone in D . The exact CPU-time depends on the quality of the feasible solutions and is expected to be more expansive than linear growth. Then, we may compare the corresponding accumulated costs $\text{cpu}(j)$ with the idealised costs for one run of the qp sub-routine with the current dimension $D_j = N(K2^j + 1)$. The quotient $\text{cpu}(j)/D_j$ in the last column may be regarded as the guaranteed estimate of the reduction of computer costs by our strategy.

Remark 8. For the finest discretisation, there are $N + 1 + N(K2^J + 1)$ unknowns and the MATLAB optimisation toolbox solver qp (quadratic programming) performed reasonably only for less than hundred unknowns. In particular, we cannot offer a comparison of real CPU-time because we did not succeed with the qp sub-routine for higher dimensions. This underlines the reliability of our multilevel active set strategy.

Remark 9. Even for a small number of unknowns, the side restrictions appeared very stringently. It was not only advisable but necessary to provide the qp sub-routine with a feasible point. In our implementation, we took

Table 1. Performance of proposed algorithm with additional space-refinement

	j	D_j	$\text{nnz}(\text{atoms})/N$	$\text{cpu}(j)$	$\text{cpu}(j)/D_j$
	level	number of atoms	active atoms per element	estimated CPU-time	guaranteed cost reduction
	1	3	3	23	2.3
	2	4	4.8	59	1.9
Fig. 3 →	3	9	4.8	195	3.5
	4	17	4	256	1.2
	5	33	3.9	369	.93
Fig. 4 →	6	65	4.1	492	.31
	7	129	4.2	617	.20
	8	257	5.2	916	.07
	9	513	6.8	1292	.05

the current displacement approximation or initially an affine approximation (which, e.g., satisfies the Dirichlet conditions). Then, like in (6.3), we added further atoms (namely a or b) to guarantee that, on each element, the current slope belong to the convex set of active atoms. After that, a feasible set of coefficients existed and was constructed.

Remark 10. From various numerical tests, we concluded that it is faster and more reliable to choose the initial (i.e. coarsest) grid in the s -variable very coarse, say K below 10.

Remark 11. It appeared reasonable to decrease the tolerance, i.e., to lower ϵ_j from one to the next finer level.

Remark 12. The values for a , b and the brake point x_b are chosen irrational to avoid the perfect situation where ± 1 belongs to the possible set of atoms and the solution brakes just at the mesh point. In this case, the practical performance is even more reliable and efficient.

Acknowledgements. This research was partly covered by the grant No. A 107 5705 of the Academy of Sciences of Czech Republic and No. 201/96/0228 of the Grant Agency of the Czech Republic.

References

- [B] Ball, J.M. (1989): A version of the fundamental theorem for Young measures. In: M.Rasle, D.Serre, M.Slemrod (eds.), PDEs and Continuum Models of Phase Transition. Lecture Notes in Physics **344**, Springer, Berlin, pp.207–215
- [BC] Brighi, B., Chipot, M. (1994): Approximated convex envelope of a function. SIAM J. Numer. Anal. **31**, 128–148

- [CP] Carstensen, C., Plecháč, P. (1997): Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comp.* **66** 997–1026
- [C] Chipot, M. (1991): Numerical analysis of oscillations in conconvex problems. *Numer. Math.* **59**, 747–767
- [CCK] Chipot, M., Collins, C., Kinderlehrer, D. (1995): Numerical analysis of oscillations in multiple well problems. *Numer. Math.* **70**, 259–282
- [CKL] Collins, C., Kinderlehrer, D., Luskin, M. (1991): Numerical approximation of the solution of a variational problem with a double well potential. *SIAM J. Numer. Anal.* **28**, 321–332
- [CL] Collins, C., Luskin, M. (1991): Optimal order error estimates for finite element approximation of the solution of a nonconvex variational problem. *Math. Comp.* **57**, 621–637
- [DS] Dunford N., Schwartz, J.T. (1967): *Linear Operators, Part I*, Interscience, New York
- [F] Friesecke, G. (1994): A necessary and sufficient conditions for nonattainment and formation of microstructure almost everywhere in scalar variational problems. *Proc. Royal Soc. Edinburgh* **124 A**, 437–471
- [L] Luskin, M. (1996): On the computation of crystalline microstructure. *Acta Numerica* **5**, 191–257
- [NW1] Nicolaides, R.A., Walkington, N.J. (1992): Computation of microstructure utilizing Young measure representations. In: C.A.Rogers, R.A.Rogers (eds.), *Recent Advances in Adaptive and Sensory Materials and their Applications*. Technomic Publ., Lancaster, pp. 131–141
- [NW2] Nicolaides, R.A., Walkington, N.J. (1995): Strong convergence of numerical solutions to degenerate variational problems. *Math. Comp.* **64**, 117–127
- [P1] Pedregal, P. (1995): Numerical computation of parametrized measures. *Numer. Funct. Anal. Optim.* **16**, 1049–1066
- [P2] Pedregal, P. (1996): On the numerical analysis of non-convex variational problems. *Numer. Math.* **74**, 325–336
- [R1] Roubíček, T. (1994): Finite element approximation of a microstructure evolution. *Math. Methods in the Applied Sciences* **17**, 377–393
- [R2] Roubíček, T. (1997): *Relaxation in Optimization Theory and Variational Calculus*. W. de Gruyter, Berlin, New York
- [R3] Roubíček, T. (1996): Numerical approximation of relaxed variational problems. *J. Convex Anal.* **3**, 329–347
- [R4] Roubíček, T. (1998): Optimality conditions for nonconvex variational problems relaxed in terms of Young measures. *Kybernetika* **34**, 335–347
- [S] Schonbek, M.E. (1982): Convergence of solutions to nonlinear dispersive equations. *Comm. in Partial Diff. Equations* **7**, 959–1000