Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Optimization problems with concentration and oscillation effects: relaxation theory and numerical approximation

by

Martin Kružík and Tomáš Roubíček

Preprint-Nr.: 57 1998

OPTIMIZATION PROBLEMS WITH CONCENTRATION AND OSCILLATION EFFECTS: RELAXATION THEORY AND NUMERICAL APPROXIMATION

Martin Kružík[1] & **Tomáš Roubíček**[2,1]

[1] *Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, CZ-182 08 Praha 8, Czech Republic.*

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

Abstract. The paper deals with optimal control problems for which minimizing (sub)sequences of controls do not converge weakly in L^1. For such problems, here governed by ordinary differential equations, the relaxed (generalized) solutions in terms of DiPerna-Majda measures are defined, correctness of the relaxation is shown and a numerical approximation is developed and tested on model examples.

Key words. Optimal control, impulse control, oscillations, concentrations, Young measures, DiPerna-Majda measures, weak L^1-compactness, numerical approximation.

AMS (MOS) subject classification. 49J15, 49N25, 65K10.

1 Motivation – examples

The paper treats optimization problems which are coercive only in a non-reflexive space L^1. In other words, the minimized functional has only the linear growth. More generally, the concentration effects can occur if the growth of the minimized functional is not strictly greater than the growth of the constraints. Such problems naturally appear in many applications, for instance in impulsive control theory (see e.g. Blaquire [4], Bryson, Ho [5, Sect. 3.7], Getz, Martin [9], Liu [16], Rempala, Zabczyk [20], Seierstad [25], Warga [29], etc.) or in variational problems of the type of Plateau's nonparametric minimal (hyper)surface (see e.g. Giusti [10] or Kačur and Souček [12] and references therein).

For simplicity, we will confine ourselves to problems of optimal control. Our relaxation method presented in Section 3 is systematically based on functional-analytical approach, using DiPerna and Majda's [6] generalization (cf. also [14, 21]) of classical Young measures [31] summarized in Section 2. Developing a theory of approximation of such measures in Section 4, it leads in Section 5 naturally to an efficient numerical method for the relaxed problem which keeps convex structure (if the relaxed problem has it) in the approximate relaxed problems, so that sometimes even finite linear-quadratic solvers can be used, see Section 6 below.

Let us however begin with a few illustrative examples. We remark that, contrary to usual impluse control [5, 9, 20] we admit unlimitted number of impulses and concentration effects combined with oscillation ones and also, in contrary to [25], the general theory will admitt vector-valued control u.

1.1 Example. Let us first consider the following Bolza-type optimal control problem for a system governed by an initial-value problem for a linear ordinary differential equation:

$$\left.\begin{array}{ll}\text{Minimize} & J(y,u) := \displaystyle\int_0^T (2-2t+t^2)|u(t)|\,\mathrm{d}t + (y(T)-1)^2 \\ \text{subject to} & \dfrac{\mathrm{d}y}{\mathrm{d}t} = u, \quad y(0)=0\ , \\ & y \in W^{1,1}(0,T)\,, \quad u \in L^1(0,T)\,, \end{array}\right\} \tag{1}$$

where u is a control ranging the Lebesgue space of integrable functions $L^1(0,T)$ and y is the corresponding state living in the Sobolev space $W^{1,1}(0,T)$ of functions whose (distributional) derivative belongs to $L^1(0,T)$, and the time horizon $T > 1$ is fixed. Roughly speaking, a minimizing control u must drive the state y sufficiently close to 1 at the terminal time $t = T$. Note also that the coefficient $a(t) := t^2 - 2t + 2$ attains its minimum value at the point $t = 1$ so that the optimal control is forced to concentrate around $t = 1$ provided $T > 1$. Considering, for $\varepsilon \in (0, T-1)$ and $\ell \in \mathbb{R}$, the control

$$u^\varepsilon(t) = \begin{cases} \ell/\varepsilon & \text{if } t \in (1, 1+\varepsilon)\ , \\ 0 & \text{otherwise}\ , \end{cases} \tag{2}$$

the corresponding state y^ε is obviously given by

$$y^\varepsilon(t) = \begin{cases} 0 & \text{if } t \in (0,1\,]\ , \\ \ell(t-1)/\varepsilon & \text{if } t \in (1, 1+\varepsilon\,]\ , \\ \ell & \text{if } t \in (1+\varepsilon, T\,]\ , \end{cases} \tag{3}$$

and $J(y^\varepsilon, u^\varepsilon) = \ell \min_{t\in[0,T]} a(t) + (\ell-1)^2 + \mathcal{O}(\varepsilon^2)$. Since $\min_{t\in[0,T]} a(t) = 1$, the sequence $\{(y^\varepsilon, u^\varepsilon)\}_{\varepsilon>0}$ will minimize J provided $\ell = 1/2$; then obviously $\lim_{\varepsilon\to0} J(y^\varepsilon, u^\varepsilon) = 3/4 = \inf J$. On the other hand, this value $\inf J$ cannot be achieved, i.e. the optimal control does not exist. Indeed, supposing $u \in L^1(0,T)$ optimal, then obviously $u \neq 0$, and we can always take some nonvanishing "part" of this control and add the corresponding area in a closer neighborhood of 1. This does not affect $y(T)$ but make $\int_0^T (2-2t+t^2)|u(t)|\,\mathrm{d}t$ lower, contradicting the optimality of u. This nonattainment is here because of the concentration effect. More precisely, the sequence $\{u^\varepsilon\}_{\varepsilon>0} \subset L^1(0,T)$ is not uniformly integrable (see e.g. [7] for a definition). As we mentioned above the reason for concentration effects is the coercivity of the optimization problem only in L^1-space, i.e. the linear growth in terms of u of the cost functional. Concentration effects are excluded if the growth of the cost functional is strictly greater than the growth of the controlled system, as proved in [21, Corollary 4.3.5].

1.2 Example. Moreover, the concentration effects can be combined with oscillation ones. This can be seen from the following optimal control problem:

$$\left.\begin{array}{lll}\text{Minimize} & J(y,u) := \int_0^T (2-2t+t^2)|u(t)| + y_2^2(t)\,\mathrm{d}t + (y_1(T)-1)^2 & \\ \text{subject to} & \dfrac{\mathrm{d}y_1}{\mathrm{d}t} = |u|, & y_1(0)=0\ , \\ & \dfrac{\mathrm{d}y_2}{\mathrm{d}t} = \dfrac{\max(0,u)}{\alpha} + \dfrac{\min(0,u)}{1-\alpha}, & y_2(0)=0\ , \\ & y \equiv (y_1,y_2) \in W^{1,1}(0,T;\mathbb{R}^2), & u \in L^1(0,T), \end{array}\right\} \quad (4)$$

with some $\alpha \in (0,1)$. Obviously, $\|y_2\|_{L^2(0,T)}$ tends to be as small as possible. Taking the control

$$u^\varepsilon(t) = \begin{cases} \ell/\varepsilon & \text{if } t \in (1, 1+\alpha\varepsilon), \\ -\ell/\varepsilon & \text{if } t \in (1+\alpha\varepsilon, 1+\varepsilon), \\ 0 & \text{otherwise}\ , \end{cases} \quad (5)$$

the corresponding state $y^\varepsilon = (y_1^\varepsilon, y_2^\varepsilon)$ has the component y_1^ε as in (3) while y_2^ε is small, namely $\|y_2\|_{L^2(0,T)} = \mathcal{O}(\varepsilon)$. Argumenting as in the previous case, for $\ell = 1/2$ we get a minimizing sequence and the infimum of the problem (4) is again $3/4$.

1.3 Example. Concentrations do not have to occur only at isolated points but they can be smeared out along the whole interval. This can be demonstrated on the following problem:

$$\left.\begin{array}{ll}\text{Minimize} & J(y,u) := \int_0^1 e_\theta(u(t)) + (y(t)-t)^2\,\mathrm{d}t \\ \text{subject to} & \dfrac{\mathrm{d}y}{\mathrm{d}t} = u, \quad y(0)=0\ , \\ & y \in W^{1,1}(0,1)\ , \quad u \in L^1(0,1)\ , \end{array}\right\} \quad (6)$$

where $e_\theta : \mathbb{R} \to \mathbb{R}$ is defined by

$$e_\theta(s) = \begin{cases} (1+\theta)|s| & \text{if } s \in [-1,1]\ , \\ \mathrm{e}^{(1-|s|)} + \theta|s| & \text{otherwise}\ . \end{cases} \quad (7)$$

with a parameter $\theta \geq 0$. For $\theta > 0$, the problem (6) is coercive in $L^1(0,T)$ but we know the exact behavior of minimizing sequences only for $\theta = 0$; then the infimum of (6) is 0. Indeed, consider the control

$$u^\varepsilon(t) = \begin{cases} \varepsilon^{-1} & \text{if } t \in [l\varepsilon - \frac{\varepsilon^2}{2}, l\varepsilon + \frac{\varepsilon^2}{2}], \\ 0 & \text{otherwise}\ , \end{cases}$$

where $\varepsilon > 0$ is small and $1 \leq l \leq 1/\varepsilon - 1$. The corresponding state is

$$y^\varepsilon(t) = \begin{cases} (l-1)\varepsilon^{-1} + \varepsilon^{-1}(t - l\varepsilon + \frac{\varepsilon^2}{2}) & \text{if } t \in [l\varepsilon - \frac{\varepsilon^2}{2}, l\varepsilon + \frac{\varepsilon^2}{2}] \\ (l-1)\varepsilon^{-1} & \text{otherwise}\ . \end{cases}$$

An easy calculation shows that $\lim_{\varepsilon\to 0} J(y^\varepsilon, u^\varepsilon) = 0$. On the other hand, there is no $y \in W^{1,1}(0,1),\ \ u \in L^1(0,1)$ satisfying the state equation for which $J(y,u) = 0$.

1.4 Example. By a combination of (4) and (6), one gets the following example:

$$\left.\begin{array}{lll}\text{Minimize} & J(y,u) := \displaystyle\int_0^1 e_\theta(u(t)) + (y_1(t)-t)^2 + y_2(t)^2\,\mathrm{d}t & \\ \text{subject to} & \dfrac{\mathrm{d}y_1}{\mathrm{d}t} = |u|\ , & y_1(0)=0\ , \\ & \dfrac{\mathrm{d}y_2}{\mathrm{d}t} = \dfrac{\max(0,u)}{\alpha} + \dfrac{\min(0,u)}{1-\alpha}\ , & y_2(0)=0\ , \\ & y \equiv (y_1,y_2) \in W^{1,1}(0,T;\mathbb{R}^2)\ , & u\in L^1(0,T)\ , \end{array}\right\} \tag{8}$$

Here again the minimum does not exist and utilizing a similar construction as in (4) and (6) we can see that the infimum is zero if $\theta = 0$. Figure 1 sketches one possible minimizing sequence for $\theta = 0$ and $\alpha = 1/3$.

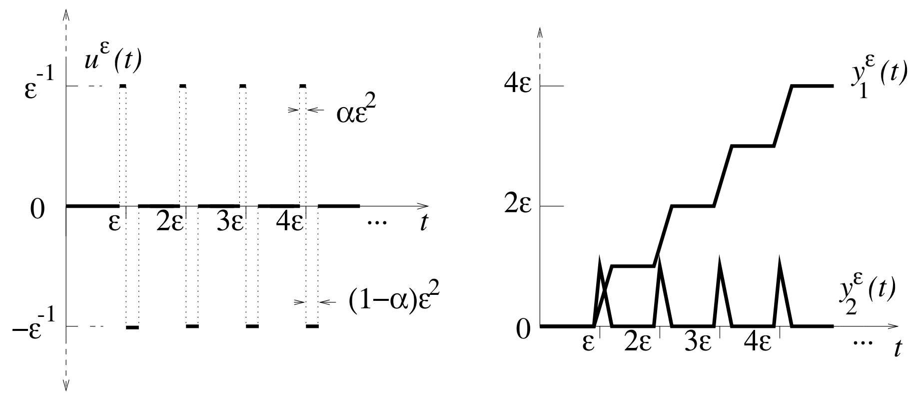

Figure 1: *A "nearly" optimal control u^ε and the response y^ε.*

Therefore, Examples 1.1-1.4 show that it is desirable to look for some generalized solutions to (1), (4), (6) and (8). It is clear that standard relaxed controls (cf. e.g. [30]) in terms of Young measures [31] cannot record properly concentration effects and thus we have to make the relaxation by a suitable generalization – here we use the DiPerna and Majda [6] measures, briefly summarized in Section 2. Then, in Section 3, we will use these measures for a relaxation (by a continuous extension) of optimal control problems from the class containing also Examples 1.1–4. To handle the relaxed problems numerically, in Section 4 we construct a convex finite-dimensional subset of DiPerna-Majda measures. For this, we will need an auxiliary envelope finer than DiPerna-Majda measures so that piecewise constant test functions can be admitted, constructed by using the so-called generalized Young functionals; cf. [15, 21, 22]. Then, in Section 5, we construct approximate relaxed problems and prove their convergence. Finally, numerical results for Examples 1.1-1.4 are presented in Section 6.

2 DiPerna-Majda measures in brief

First, we have to construct a suitable locally compact convex hull of the involved Lebesgue spaces $L^p(0,T;\mathbb{R}^m)$ used for $p = 1$ in Examples 1.1-1.4. As already mentioned, we will use a special extension proposed by DiPerna and Majda [6]. In what follows $C^0(\mathbb{R}^m)$ stands for the space of bounded continuous functions while $C(\mathbb{R}^m)$ for the space of continuous functions on $\mathbb{R}^m$, etc.

Let S^{m-1} denote the unit sphere in $\mathbb{R}^m$. Let further

$$\mathcal{R} \;=\; \Big\{ w\in C^0(\mathbb{R}^m);\ \exists w_0\in C^0(\mathbb{R}^m),\ w_1\in C(S^{m-1}) : \lim_{|s|\to\infty} w_0(s)=0\ ,\quad w(s)=w_0(s)+w_1\left(\frac{s}{|s|}\right)\frac{|s|^p}{1+|s|^p} \Big\}\ ; \tag{9}$$

the set $\mathcal{R}$ is a complete separable subring of the ring $C^0(\mathbb{R}^m)$ of all bounded continuous functions on $\mathbb{R}^m$. The corresponding compactification of $\mathbb{R}^m$, denoted $\gamma\mathbb{R}^m$, is then homeomorphic with a unit ball in $\mathbb{R}^m$, or equivalently with a simplex Δ in $\mathbb{R}^m$. This means that every $w\in\mathcal{R}$ admits a uniquely defined continuous extension on $\gamma\mathbb{R}^m$ (denoted then again by w without any misunderstanding) and conversely for every $w\in C(\gamma\mathbb{R}^m)$ the restriction on $\mathbb{R}^m$ lives in $\mathcal{R}$, cf. e.g. [8].

Further, we will denote by $\mathcal{Y}([0,T],\sigma;\gamma\mathbb{R}^m)$ the subset of $L^\infty_{\rm w}([0,T],\sigma;{\rm rca}(\gamma\mathbb{R}^m))$ consisting from the mappings $\hat\nu: t\mapsto\hat\nu_t$ such that $\hat\nu_t$ is a probability measure on $\gamma\mathbb{R}^m$ for σ-a.a. $t\in[0,T]$; here σ is a positive Radon measure on $[0,T]$ and $L^\infty_{\rm w}([0,T],\sigma;{\rm rca}(\gamma\mathbb{R}^m))$ denotes the Banach space of all weakly σ-measurable (i.e., for any $w\in\mathcal{R}$, the mapping $[0,T]\to\mathbb{R}: t\mapsto\int_{\gamma\mathbb{R}^m} w(s)\hat\nu_t({\rm d}s)$ is σ-measurable in the usual sense) σ-essentially bounded mappings from $(0,T)$ to the set of Radon measures ${\rm rca}(\gamma\mathbb{R}^m)$ on $\gamma\mathbb{R}^m$.

DiPerna and Majda [6] showed that, having a bounded sequence $\{u_k\}_{k\in\mathbb{N}}$ in $L^p(0,T;\mathbb{R}^m)$, there exists its subsequence (denoted by the same indices), $\sigma\in{\rm rca}([0,T])$ positive and $\hat\nu\in\mathcal{Y}([0,T],\sigma;\gamma\mathbb{R}^m)$ such that, for any $g\in C([0,T])$ and any $w\in\mathcal{R}$,

$$\lim_{k\to\infty}\int_0^T g(t)v(u_k(t)){\rm d}t \;=\; \int_0^T\int_{\gamma\mathbb{R}^m} g(t)w(s)\hat\nu_t({\rm d}s)\sigma({\rm d}t)\ , \tag{10}$$

where $v(s)=w(s)(1+|s|^p)$; note that w on the right-hand side of (10) denotes in fact a continuous extension on $\gamma\mathbb{R}^m$. We say that such a pair $(\sigma,\hat\nu)\in{\rm rca}([0,T])\times\mathcal{Y}([0,T],\sigma;\gamma\mathbb{R}^m)$ is attainable by a sequence $\{u_k\}_{k\in\mathbb{N}}\subset L^p(0,T;\mathbb{R}^m)$. The set of all attainable pairs $(\sigma,\hat\nu)$ is denoted by $\mathcal{DM}^p(0,T;\mathbb{R}^m)$, and its subset containing measures attainable by sequences contained in the ball of the radius $\rho>0$ in $L^p(0,T;\mathbb{R}^m)$ is denoted by $\mathcal{DM}^p_\rho(0,T;\mathbb{R}^m)$.

It is well-known (see [18, 19, 21, 24]) that, for any $v\in C(\mathbb{R}^m)$ such that $\lim_{|s|\to\infty} v(s)/|s|^p=0$ and any $g\in L^\infty(0,T)$, the limit in (10) can also be described as

$$\lim_{k\to\infty}\int_0^T g(t)v(u_k(t)){\rm d}t \;=\; \int_0^T\int_{\mathbb{R}^m} g(t)v(s)\nu_t({\rm d}s)\,{\rm d}t\ , \tag{11}$$

where $\nu=\{\nu_t\}_{t\in[0,T]}$ is a so-called L^p-Young measure, i.e. $\nu\in\mathcal{Y}([0,T];\mathbb{R}^m)$ and $\int_0^T\int_{\mathbb{R}^m}|s|^p\nu_t({\rm d}s)\,{\rm d}x<+\infty$; cf. [13]. We then say that ν is generated by $\{u_k\}_{k\in\mathbb{N}}$.

We will call a DiPerna-Majda measure $(\sigma,\hat\nu)\in\mathcal{DM}^p(0,T;\mathbb{R}^m)$ p-nonconcentrating if $\int_0^T\int_{\gamma\mathbb{R}^m\setminus\mathbb{R}^m}\hat\nu_t({\rm d}s)\sigma({\rm d}t)=0$. We should also mention that to any $(\sigma,\hat\nu)\in\mathcal{DM}^p(0,T;\mathbb{R}^m)$

we can assign the so-called p-nonconcentrating modification, i.e., $(\sigma^\circ,\hat\nu^\circ)\in\mathcal{DM}^p(0,T;\mathbb{R}^m)$ such that

$$\int_0^T\int_{\gamma\mathbb{R}^m} g(t)w(s)\hat\nu_t(\mathrm{d}t)\sigma(\mathrm{d}t)=\int_0^T\int_{\gamma\mathbb{R}^m} g(t)w(s)\hat\nu_t^\circ(\mathrm{d}t)\sigma^\circ(\mathrm{d}t)$$

for any $g\in C([0,T])$ and any $w\in\mathcal{R}$ vanishing at infinity. It was proved in ([13, 14]) that σ° has a density with respect to the Lebesgue measure and that $(\sigma^\circ,\hat\nu^\circ)$ can be expressed through an L^p-Young measure representation; cf. [21, Prop. 3.4.15]. Therefore, (10) can be now written as

$$\lim_{k\to\infty}\int_0^T g(t)v(u_k(t))\mathrm{d}t=\int_0^T\int_{\mathbb{R}^m} g(t)v(s)\nu_t(\mathrm{d}s)\mathrm{d}t+\int_0^T\int_{\gamma\mathbb{R}^m\setminus\mathbb{R}^m} g(t)w(s)\hat\nu(\mathrm{d}s)\sigma(\mathrm{d}t). \quad (12)$$

For a numerical implementation, it will be important that we have the following complete characterization of $\mathcal{DM}^p(0,T;\mathbb{R}^m)$ at our disposal; see [14, 21] for the proof.

2.1 Theorem. *Let $\mathcal{R}$ be as in (9) and $(\sigma,\hat\nu)\in\mathrm{rca}([0,T])\times\mathcal{Y}([0,T],\sigma;\gamma\mathbb{R}^m)$. Then the pair $(\sigma,\hat\nu)$ belongs to $\mathcal{DM}^p(0,T;\mathbb{R}^m)$ if and only if the following three properties are satisfied simultaneously:*

1. *σ is positive,*
2. *$\sigma_{\hat\nu}\in\mathrm{rca}([0,T])$ defined by $\sigma_{\hat\nu}(\mathrm{d}t)=(\int_{\mathbb{R}^m}\hat\nu_t(\mathrm{d}s))\sigma(\mathrm{d}t)$ is absolutely continuous with respect to the Lebesgue measure ($d_{\sigma_{\hat\nu}}$ will denote its density),*
3. *for a.a. $t\in(0,T)$ it holds*

$$\int_{\mathbb{R}^m}\hat\nu_t(\mathrm{d}s)>0,\qquad d_{\sigma_{\hat\nu}}(t)=\left(\int_{\mathbb{R}^m}\frac{\hat\nu_t(\mathrm{d}s)}{1+|s|^p}\right)^{-1}\int_{\mathbb{R}^m}\hat\nu_t(\mathrm{d}s)\ .$$

Let us define the natural imbedding $i:L^p(0,T;\mathbb{R}^m)\to\mathcal{DM}^p(0,T;\mathbb{R}^m)$ as

$$i(u)=(\sigma,\hat\nu)\,,\qquad\text{where } \hat\nu_t=\delta_{u(t)} \text{ and } d_\sigma(t)=1+|u(t)|^p \text{ for a.a. } t\in[0,T]. \quad (13)$$

Alternatively, DiPerna and Majda [6] however worked with measures from $\mathrm{rca}([0,T]\times\gamma\mathbb{R}^m)$; let us put here

$$\mathrm{DM}^p(0,T;\mathbb{R}^m)=\Big\{\eta\in\mathrm{rca}([0,T]\times\gamma\mathbb{R}^m);\ \exists\{u_k\}_{k\in\mathbb{N}}\subset L^p(0,T;\mathbb{R}^m) \quad (14)$$
$$\forall z\in C([0,T]\times\gamma\mathbb{R}^m):\ \langle\eta,z\rangle=\lim_{k\to\infty}\int_0^T z(t,u_k(t))(1+|u_k(t)|^p)\mathrm{d}t\Big\}$$

where, of course, $\langle\eta,z\rangle=\int_{[0,T]\times\gamma\mathbb{R}^m}z(t,s)\eta(\mathrm{d}t\mathrm{d}s)$. Further, there is a one-to-one correspondence between $\mathcal{DM}^p(0,T;\mathbb{R}^m)$ and $\mathrm{DM}^p(0,T;\mathbb{R}^m)$ given by

$$\int_{[0,T]\times\gamma\mathbb{R}^m}g(t)w(s)\eta(\mathrm{d}t\mathrm{d}s)=\int_0^T g(t)\int_{\gamma\mathbb{R}^m}w(s)\hat\nu_t(\mathrm{d}s)\sigma(\mathrm{d}t)\ ,\qquad g\in C([0,T]),\ w\in\mathcal{R}. \quad (15)$$

Thus we can also denote by $\mathrm{DM}^p_\rho(0,T;\mathbb{R}^m)$ the image of $\mathcal{DM}^p_\rho(0,T;\mathbb{R}^m)$ via this correspondence.

Let us remark that the above results as well as the approximation theory introduced below hold for a multidimensional domain with a finite Lebesgue measure in place of the interval $(0,T)$, too.

3 Optimal-control problems and their relaxation

Let us now consider the following Bolza problem covering the special cases (1), (4), (6) and (8):

$$
\text{(P)} \qquad \begin{cases} \text{Minimize} & J(y,u) := \int_0^T a(t,y,u) + b(t,u)\,\mathrm{d}t + f(y(T)) \\ \text{subject to} & \frac{\mathrm{d}y}{\mathrm{d}t} = c(t,y,u) + d(t,u), \quad y(0) = y_0 \ , \\ & y \in W^{1,1}(0,T;\mathbb{R}^n) \ , \qquad u \in L^p(0,T;\mathbb{R}^m) \ , \end{cases}
$$

where $a : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is a Carathéodory function (i.e. $a(t,\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuous for a.a. $t \in (0,T)$ and $a(\cdot,r,s) : (0,T) \to \mathbb{R}$ is measurable for all r and s), also $c : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a Carathéodory function, $b \in C([0,T] \times \mathbb{R}^m)$ and $f \in C(\mathbb{R}^n)$ satisfying

$$\max(|a(t,r,s)|, |c(t,r,s)|) \le \alpha_{1+\varepsilon}(t) + \beta(|r|^{1/\varepsilon} + |s|^{p/(1+\varepsilon)}), \tag{16}$$

$$b_0 \in C([0,T];\mathcal{R}), \qquad \text{where} \ \ b_0(t,s) := b(t,s)/(1+|s|^p) \ , \tag{17}$$

$$|c(t,r,s)| \le (\alpha_1(t) + \beta|s|^p)(1+|r|), \tag{18}$$

$$|a(t,r_1,s) - a(t,r_2,s)| \le (\alpha_1(t) + \beta|r_1|^{1/\varepsilon} + \beta|r_2|^{1/\varepsilon} + \beta|s|^{p/(1+\varepsilon)})|r_1 - r_2|, \tag{19}$$

$$|c(t,r_1,s) - c(t,r_2,s)| \le (\alpha_1(t) + \beta|r_1|^{1/\varepsilon} + \beta|r_2|^{1/\varepsilon} + \beta|s|^{p/(1+\varepsilon)})|r_1 - r_2|, \tag{20}$$

$$d_0 \in C([0,T];\mathcal{R}^n), \qquad \text{where} \ \ d_0(t,s) := d(t,s)/(1+|s|^p) \ , \tag{21}$$

$$f \ge 0 \ , \quad a(t,r,s) + b(t,s) \ge \varepsilon|s|^p \ , \tag{22}$$

with some $\varepsilon > 0$, $\beta \in \mathbb{R}$, $\alpha_q \in L^q(0,T)$.

Due to concentrations in the control u, the response y may fall out $W^{1,1}(0,T;\mathbb{R}^n)$, keeping its variation bounded. Instead of a bounded-variation space $\mathrm{BV}(0,T;\mathbb{R}^n)$, we will rather work with a bit finer extension $W^1_\mu(0,T;\mathbb{R}^n)$ of $W^{1,1}(0,T;\mathbb{R}^n)$ introduced by Souček [26], namely

$$W^1_\mu(0,T;\mathbb{R}^n) \ := \ \Big\{ (y,\dot{y}) \in L^1(0,T;\mathbb{R}^n) \times \mathrm{rca}([0,T];\mathbb{R}^n); \tag{23}$$

$$\exists \{y_k\}_{k\in\mathbb{N}} \subset W^{1,1}(0,T;\mathbb{R}^n) : \ \ y_k \to y, \ \frac{\mathrm{d}y_k}{\mathrm{d}t} \rightharpoonup \dot{y} \text{ weakly*} \Big\} ;$$

more precisely, [26] used $n = 1$ but a multidimensional domain instead of $[0,T]$. See also [21, Example 5.1.8] for a comparison with $\mathrm{BV}(0,T;\mathbb{R}^n)$. The canonical embedding $j : W^{1,1}(0,T;\mathbb{R}^n) \to W^1_\mu(0,T;\mathbb{R}^n)$ is defined by $j(y) = (y, \mathrm{d}y/\mathrm{d}t)$. It is shown in [26] that, if normed by $\|(y,\dot{y})\| := \|y\|_{L^1(0,T;\mathbb{R}^n)} + \|\dot{y}\|_{\mathrm{rca}([0,T];\mathbb{R}^n)}$, $W^1_\mu(0,T;\mathbb{R}^n)$ is a Banach space containing, just by definition (23), $j(W^{1,1}(0,T;\mathbb{R}^n))$ densely. Moreover, there exist unique $y_T, y_0 \in \mathbb{R}^n$ such that the per-partes formula

$$\int_0^T \left(y\frac{\mathrm{d}v}{\mathrm{d}t} + \dot{y}v \right) \mathrm{d}t = y_T v(T) - y_0 v(0)$$

holds for any $v \in C^1([0,T])$; it is then natural to call y_T and y_0 the trace of $(y,\dot y)$ and write $y_T = (y,\dot y)|_{t=T}$ and $y_0 = (y,\dot y)|_{t=0}$. The mapping $(y,\dot y) \mapsto (y_T, y_0) : W^1_\mu(0,T;\mathbb{R}^n) \to \mathbb{R}^n \times \mathbb{R}^n$ is weakly* continuous and the balls in $W^1_\mu(0,T;\mathbb{R}^n)$ are weakly* compact; cf. [26, Theorems 2(ii) and 6].

Using DiPerna-Majda's extension of $L^p(0,T;\mathbb{R}^m)$ and Souček's extension of $W^{1,1}(0,T;\mathbb{R}^n)$, the relaxed problem will look as:

$$
\text{(RP)} \left\{ \begin{array}{ll} \text{Minimize} & \hat J(y,\dot y,\sigma,\hat\nu) := \displaystyle\int_0^T\!\!\int_{\gamma\mathbb{R}^m} \frac{a(t,y(t),s)+b(t,s)}{1+|s|^p}\hat\nu_t(\mathrm{d}s)\sigma(\mathrm{d}t) + f\left((y,\dot y)|_{t=T}\right) \\ \text{subject to} & \dot y = \displaystyle\int_{\gamma\mathbb{R}^m} \frac{c(t,y(t),s)+d(t,s)}{1+|s|^p}\hat\nu_t(\mathrm{d}s)\sigma, \qquad (y,\dot y)|_{t=0} = y_0\,, \\ & (y,\dot y) \in W^1_\mu(0,T;\mathbb{R}^n), \qquad (\sigma,\hat\nu) \in \mathcal{DM}^p(0,T;\mathbb{R}^m)\ . \end{array} \right.
$$

Of course, the state equation is now understood in the sense of measures on $[0,T]$. The following assertion and Proposition 3.2 below justify that (RP) is actually a correct relaxation of (P).

3.1 Proposition. *Let (16)–(18), (20), and (21) be valid, let y_k solves*

$$
\frac{\mathrm{d}y_k}{\mathrm{d}t} = c(t,y_k,u_k) + d(t,u_k)\ , \qquad y_k(0) = y_0\ , \tag{24}
$$

and let the sequence $\{u_k\}_{k\in\mathbb{N}}$ attains $(\sigma,\hat\nu)$ in the sense (10). Then $(y_k, \frac{\mathrm{d}}{\mathrm{d}t}y_k) \rightharpoonup (y,\dot y)$ weakly in $W^1_\mu(0,T;\mathbb{R}^n)$ and $y_k(T) \to (y,\dot y)_{t=T}$, and $(y,\dot y)$ is the unique solution to*

$$
\dot y = \int_{\gamma\mathbb{R}^m} \frac{c(t,y(t),s)+d(t,s)}{1+|s|^p}\hat\nu_t(\mathrm{d}s)\sigma\ , \qquad (y,\dot y)|_{t=0} = y_0. \tag{25}
$$

Proof. First, we show the apriori estimate $\|y_k\|_{W^{1,1}(0,T;\mathbb{R}^n)} \le C$. Using (18) and (20), from (24) one gets

$$
\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}|y_k| &\le \left|\frac{\mathrm{d}}{\mathrm{d}t}y_k\right| = |c(y_k,u_k) + d(u_k)| \\ &\le (\alpha_1 + \beta|u_k|^p)(1+|y_k|) + \|d_0\|_{C(0,T;C(\mathbb{R}^m))}(1+|u_k|^p) \end{aligned} \tag{26}
$$

from which we get by Gronwall's inequality that $\|y_k\|_{L^\infty(0,T;\mathbb{R}^n)}$ is bounded independently of k because $|u_k|^p$ is bounded in $L^1(0,T)$. By (26) one then gets also boundedness of $\|\frac{\mathrm{d}}{\mathrm{d}t}y_k\|_{L^1(0,T;\mathbb{R}^n)}$. Altogether, we proved boundedness of $\|y_k\|_{W^{1,1}(0,T;\mathbb{R}^n)}$.

Then, we can select a subsequence (denoted by the same indices) such that $\{j(y_k)\}$ converges weakly* in $W^1_\mu(0,T;\mathbb{R}^n)$. Let us denote by $(y,\dot y)$ the limit of this subsequence. Then also $y_k(T) \to (y,\dot y)_{t=T}$ and $y_k(0) \to (y,\dot y)_{t=0}$. Since $y_k(0) = y_0$, we thus obtain $(y,\dot y)_{t=0} = y_0$.

Also, we know that

$$
\frac{\mathrm{d}y_k}{\mathrm{d}t} \to \dot y \qquad \text{weakly* in } \mathrm{rca}([0,T];\mathbb{R}^n) \tag{27}
$$

and

$$d(\cdot, u_k) \to \int_{\gamma\mathbb{R}^m} \frac{d(t,s)}{1+|s|^p} \hat{\nu}_t(\mathrm{d}s)\sigma \qquad \text{weakly* in } \mathrm{rca}([0,T];\mathbb{R}^n)\ , \tag{28}$$

which follows directly from (10) provided $d = (d_i)_{i=1}^n$, $d_i = g_i \otimes v_i$ for some $g_i \in C([0,T])$ and $v_i \in \mathcal{R}$, while for general d satisfying (21) we must still use the fact that $d(t,s)/(1+|s|^p)$ can be approximated uniformly on $[0,T] \times \mathbb{R}^n$ by functions from $C([0,T]) \otimes \mathcal{R}^n$, as implicitly used already in the definition (14).

By the compact embedding $W^{1,1}(0,T;\mathbb{R}^n) \subset L^q(0,T;\mathbb{R}^n)$, we have $y_k \to y$ in $L^q(0,T;\mathbb{R}^n)$ for any $q < +\infty$. Using $q = 1/\varepsilon$, from (16) one gets

$$c(y_k, u_k) \to \int_{\gamma\mathbb{R}^m} \frac{c(t,y(t),s)}{1+|s|^p} \hat{\nu}_t(\mathrm{d}s)\sigma \qquad \text{weakly in } L^{1+\varepsilon}(0,T;\mathbb{R}^n)\ ; \tag{29}$$

note that $t \mapsto c(t,y(t),s)$ need not be continuous but, since $c(t,r,\cdot)$ has a lesser growth than the p-th power due to (16), we can work with the L^p-Young measure representation of $(\sigma, \hat{\nu})$ from (11) and use [21, Lemma 3.6.7].

Altogether, by (27)-(29) we can pass to the limit in (24), which gives just (25).

The solution to (25) is unique. Indeed, taking two solutions $(y_1, \dot{y}_1)$ and $(y_2, \dot{y}_2)$ and subtracting (25) for $(y_1, \dot{y}_1)$ and $(y_2, \dot{y}_2)$, we get by (19) that

$$\begin{aligned} \dot{y}_{12} &= \int_{\gamma\mathbb{R}^m} \frac{c(t,y_1(t),s) - c(t,y_2(t),s)}{1+|s|^p} \hat{\nu}_t(\mathrm{d}s)\sigma \\ &\leq \left(\alpha_1 + \beta|y_1|^{1/\varepsilon} + \beta|y_2|^{1/\varepsilon} + \beta \int_{\gamma\mathbb{R}^m} \frac{|s|^{p/(1+\varepsilon)}}{1+|s|^p} \hat{\nu}_t(\mathrm{d}s)\sigma \right) |y_{12}|, \end{aligned}$$

where $y_{12} := y_1 - y_2$ and $\dot{y}_{12} := \dot{y}_1 - \dot{y}_2$. As $\int_{\gamma\mathbb{R}^m} \frac{|s|^{p/(1+\varepsilon)}}{1+|s|^p} \hat{\nu}_t(\mathrm{d}s)\sigma \in L^{1+\varepsilon}(0,T)$, one can see that $\dot{y}_{12}$ has a density, and thus $\dot{y}_{12} = \mathrm{d}y_{12}/\mathrm{d}t$ a.e. on $[0,T]$. As $y_{12}(0) = 0$, by Gronwall's inequality one gets $y_{12} = 0$ on $[0,T]$. □

Let us define $\pi : L^p(0,T;\mathbb{R}^m) \to W^{1,1}(0,T;\mathbb{R}^n)$ by $y = \pi(u)$ where y is the solution to the state equation $\mathrm{d}y/\mathrm{d}t = c(t,y,u) + d(t,u)$, $y(0) = y_0$. Then, if we define $\hat{\pi} : \mathcal{DM}^p(0,T;\mathbb{R}^m) \to W^1_\mu(0,T;\mathbb{R}^n)$ by $(y, \dot{y}) = \hat{\pi}(\sigma, \hat{\nu})$ where $(y, \dot{y})$ is the solution to the relaxed state equation (25), we have $j(\pi(u)) = \hat{\pi}(i(u))$ or, saying otherwise, $j \circ \pi = \hat{\pi} \circ i$, where the imbeddings i and j were defined respectively by (13) and by $j(y) = (y, \mathrm{d}y/\mathrm{d}t)$.

Let us note that in (16) we assumed a sub-critical growth of a and c because a correct extension of problems with terms of critical growth p in u interacting nonadditively with y would bring delicate problems, see [17]. Here, as y need not be continuous, discontinuous test integrands of the form $c(t,y(t),s)$ would have to be admitted, which would require to work with a locally convex envelope of $L^p(0,T;\mathbb{R}^m)$ strictly finer than $\mathrm{DM}^p(0,T;\mathbb{R}^m)$, cf. [21, Example 3.3.11].

The set of $(\sigma, \hat{\nu}) \in \mathcal{DM}^p(0,T;\mathbb{R}^m)$ involved in (RP) is not convex, which is an essential drawback especially for numerical treatment in Sections 4–6. Moreover, we did not defined any (locally) compact topology on $\mathcal{DM}^p(0,T;\mathbb{R}^m)$. It lead us to a modification of (RP) by exploiting $\mathrm{DM}^p(0,T;\mathbb{R}^m)$ defined in (14) and endowed by the weak* topology of $\mathrm{rca}([0,T]\times\gamma\mathbb{R}^m)$, so that we get:

$$
\text{(RP')}\quad\left\{\begin{array}{ll}
\text{Minimize} & \bar{J}(y,\dot{y},\eta):=\displaystyle\int_{[0,T]\times\gamma\mathbb{R}^m}\frac{a(t,y,s)+b(t,s)}{1+|s|^p}\eta(\mathrm{d}t\mathrm{d}s)+f\left((y,\dot{y})|_{t=T}\right)\\
\text{subject to} & \dot{y}=\displaystyle\int_{\gamma\mathbb{R}^m}\frac{c(t,y(t),s)+d(t,s)}{1+|s|^p}\eta(\cdot\mathrm{d}s),\quad (y,\dot{y})|_{t=0}=y_0\ ,\\
& (y,\dot{y})\in W^1_\mu(0,T;\mathbb{R}^n)\ ,\qquad \eta\in\mathrm{DM}^p(0,T;\mathbb{R}^m),
\end{array}\right.
$$

where $\int_{\gamma\mathbb{R}^m}\frac{z(t,s)}{1+|s|^p}\eta(\cdot\mathrm{d}s)\in\mathrm{rca}([0,T];\mathbb{R}^n)$ is defined by the identity

$$
\forall g\in C([0,T];\mathbb{R}^n):\qquad \left\langle\int_{\gamma\mathbb{R}^m}\frac{z(t,s)}{1+|s|^p}\eta(\cdot\mathrm{d}s),g\right\rangle=\int_{[0,T]\times\gamma\mathbb{R}^m}\frac{g(t)\cdot z(t,s)}{1+|s|^p}\eta(\mathrm{d}t\mathrm{d}s).
$$

Note that the problem (RP') may have a convex structure; it occurs, e.g., if $a(t,\cdot)$ is convex and $c(t,\cdot)$ is linear. We will denote $\bar{\pi}(\eta):=(y,\dot{y})$ the solution to the state problem in (RP').

3.2 Proposition. *Let (16)–(22) be valid. Then the relaxed problem* (RP') *has a solution,* inf(P)= min(RP'). *Moreover, every solution* $\eta\in\mathrm{DM}^p(0,T;\mathbb{R}^m)$ *to* (RP') *is attainable (in the sense used in (14)) by a minimizing sequence for* (P) *and, conversely, every minimizing sequence for* (P) *contains a converging (in the sense (14)) subsequence to a solution of* (RP').

Proof. Denote $\Phi(u)=J(\pi(u),u)$ and $\bar{\Phi}(\eta)=\bar{J}(\bar{\pi}(\eta),\eta)$. The existence of a solution to (RP') follows from the fact that, due to coercivity of Φ, the level sets of $\bar{\Phi}$ (i.e. sets $\{\eta\in\mathrm{DM}^p(0,T;\mathbb{R}^m);\ \bar{\Phi}(\eta)\le c\}$) are contained in a weakly* compact set $\mathrm{DM}^p_\rho(0,T;\mathbb{R}^m)$ and that $\bar{J}$ and $\bar{\pi}$ (and thus also $\bar{\Phi}$) are weakly* continuous; this follows from (16) and (19) and the (weak*,norm)-continuity of the mapping $\eta\mapsto y:\mathrm{DM}^p(0,T;\mathbb{R}^m)\to L^{1+\varepsilon}(0,T;\mathbb{R}^n)$ with $(y,\dot{y})=\bar{\pi}(\eta)$ proved essentially in Proposition 3.1.

Suppose that inf(P)> min(RP') and that $\eta\in\mathrm{DM}^p(0,T;\mathbb{R}^m)$ is a solution to (RP'). Then there is $\{u_k\}_{k\in\mathbb{N}}\subset L^p(0,T)$ that generates $\eta\in\mathrm{DM}^p(0,T;\mathbb{R}^m)$ in the sense (14). Further, we put $y_k=\pi(u_k)$. We have due to the mentioned continuity of $\bar{J}$ and $\bar{\pi}$ that $\lim_{k\to\infty}\Phi(u_k)=\lim_{k\to\infty}J(y_k,u_k)=\lim_{k\to\infty}J(\pi(u_k),u_k)=\bar{J}(\bar{\pi}(\eta),\eta)=\min(\mathrm{RP'})$. Thus, for k large enough we would get $\Phi(u_k)<\inf(\mathrm{P})$, which cannot be true. Thus we proved inf(P)≤min(RP'). If there is another sequence $\{\tilde{u}_k\}$ such that $\liminf_{k\to\infty}J(\pi(\tilde{u}_k),\tilde{u}_k)\le\lim_{k\to\infty}J(\pi(u_k),u_k)=\min(\mathrm{RP'})$ then there would be its subsequence generating some $\eta_0\in\mathrm{DM}^p(0,T;\mathbb{R}^m)$ that $\bar{J}(\bar{\pi}(\eta_0),\eta_0)<\bar{J}(\bar{\pi}(\eta),\eta)$, which gives a contradiction. This shows that $\{u_k\}$ is also a minimizing sequence for (P).

Similarly, if we suppose that inf(P)< min(RP'), then there would exist a minimizing sequence $\{u'_k\}_{k\in\mathbb{N}}$ of (P) necessarily bounded due to the coercivity of Φ implied by (22) generating some $\eta'\in\mathrm{DM}^p(0,T;\mathbb{R}^m)$ and, for $y'_k=\pi(u'_k)$, also $(y'_k,\frac{\mathrm{d}}{\mathrm{d}t}y'_k)\rightharpoonup(y',\dot{y}')=\bar{\pi}(\eta')$ weakly* in $W^1_\mu(0,T;\mathbb{R}^n)$ and $y'_k(T)\to y_T=(y',\dot{y}')_{t=T}$. Finally, we would have that $\bar{J}(\bar{\pi}(\eta'),\eta')<\bar{J}(\bar{\pi}(\eta),\eta)$, contrary to the assumption that $\bar{J}(\bar{\pi}(\eta),\eta)=\min(\mathrm{RP'})$.

Altogether, we showed that inf (P)= min (RP'). This also shows that every minimizing sequence of (P) contains a subsequence converging to a solution to (RP'). □

As there is a one-to-one mapping between $\mathcal{DM}^p(0,T;\mathrm{I\!R}^m)$ and $\mathrm{DM}^p(0,T;\mathrm{I\!R}^m)$ given by (15), we can also modify Proposition 3.2 for (RP). It gives that (RP) has a solution, inf(P) = min(RP), and there is the above specified relation between minimizing sequences for (P) and solution to (RP), the convergence being now understood in the sense of (10) instead of (14).

To illustrate the general considerations, we can return to Examples 1.1-1.4. For $m = 1$ the compactification $\gamma\mathrm{I\!R}$ is just the standard two-point compactification $\gamma\mathrm{I\!R} \cong \bar{\mathrm{I\!R}} := \mathrm{I\!R} \cup \{+\infty, -\infty\}$. In terms of the Diperna-Majda measure $(\sigma, \hat{\nu}) \in \mathcal{DM}^p(0,T;\mathrm{I\!R}^m)$, the (unique) optimal relaxed control for the special case of (1) has then the form

$$\sigma(\mathrm{d}t) \;=\; \mathrm{d}t + \ell\delta_1 \;, \qquad \hat{\nu}_t \;=\; \begin{cases} \delta_0 & \text{if } t \neq 1 \;, \\ \delta_{+\infty} & \text{if } t = 1 \;, \end{cases} \tag{30}$$

while for the case of (4) it has the form:

$$\sigma(\mathrm{d}t) \;=\; \mathrm{d}t + \ell\delta_1 \;, \qquad \hat{\nu}_t \;=\; \begin{cases} \delta_0 & \text{if } t \neq 1 \;, \\ \alpha\delta_{+\infty} + (1-\alpha)\delta_{-\infty} & \text{if } t = 1 \;, \end{cases} \tag{31}$$

with $\ell = 1/2$, where $\delta_t \in \mathrm{rca}([0,T])$ or $\delta_s \in \mathrm{rca}(\bar{\mathrm{I\!R}})$ denotes the Dirac measure supported at $t \in [0,T]$ or at $s \in \bar{\mathrm{I\!R}}$, respectively. Eventually, we know solutions to the relaxed problems of (6) and (8) explicitly only if $\theta = 0$: an optimal relaxed control for (6) with $\theta = 0$ has the form

$$\sigma(\mathrm{d}t) = 2\mathrm{d}t \;, \qquad \hat{\nu}_t = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_{+\infty} \;, \tag{32}$$

and an optimal relaxed control for (8) with $\alpha = 1/2$ and $\theta = 0$ is

$$\sigma(\mathrm{d}t) = 2\mathrm{d}t \;, \qquad \hat{\nu}_t = \frac{1}{2}\Big(\delta_0 + \alpha\delta_{+\infty} + (1-\alpha)\delta_{-\infty}\Big) \;. \tag{33}$$

See also Figures 2–5 below for (an approximation of) such optimal relaxed controls; we calculated a bit more illustrative case $\theta > 0$, however.

4 Approximation of DiPerna-Majda measures

We want to apply the general approximation theory introduced in [21] resulting here to some suitable finite-dimensional convex subsets of $\mathrm{DM}^p(0,T;\mathrm{I\!R}^m)$. In such a way, the possible convex structure of (RP) is preserved for discrete problems, which is advantageous for optimization routines. We want to use the partition of the interval $[0,T]$ and then piecewise constant DiPerna-Majda measures. The standard construction from [21] thus needs piecewise continuous test functions, but the DiPerna-Majda measures introduced in

Sect. 3 admit only continuous test functions. This will force us to introduce an auxiliary finer extension (i.e. a finer convex local compactification) of $L^p(0,T;\mathbb{R}^m)$.

Let $\{\mathcal{T}^1_{d_1}\}_{d_1>0}$ denote an equidistant partition of $[0,T]$ and $\{\mathcal{T}^2_{d_2}\}_{d_2>0}$ a regular triangulation of $\gamma\mathbb{R}^m \cong \mathbb{R}^m \cup S^{m-1}$ with $\{s^l_{d_2}\}_{l=1}^{L(d_2)}$ the set of mesh points; here we use the fact that $\mathbb{R}^m \cup S^{m-1}$ is homeomorphic with a compact polyhedral domain in $\mathbb{R}^m$ that is then triangulated by a mesh of $L(d_2) \in \mathbb{N}$ points. We suppose that $\mathcal{T}^1_{d'}$ is a refinement of $\mathcal{T}^1_d$ if $d' < d$ and similarly for $\mathcal{T}^2_d$. In what follows we will denote the number of subintervals in $[0,T]$ by $M(d_1) = T/d_1$. In the following two assertions, we now formulate the results of this sections:

4.1 Proposition. *Let $\eta_d \in \mathrm{rca}([0,T] \times \gamma\mathbb{R}^m)$ be given by the formula*

$$\langle \eta_d, h\rangle = \sum_{j=1}^{M(d_1)} \sum_{l=1}^{L(d_2)} \int_{E^1_j} h_0(t, s^l_{d_2}) q_{jl} \ dt \ , \tag{34}$$

with $h_0 \in C([0,T] \times \gamma\mathbb{R}^m)$ being the continuous extension of $h(t,s)/(1+|s|^p)$, and with

$$\sum_{l=1}^{L(d_2)} \frac{q_{jl}}{1+|s^l_{d_2}|^p} = 1 \ , \qquad 1 \le j \le M(d_1), \quad \text{and} \tag{35}$$

$$q_{jl} \ge 0 \ , \qquad 1 \le j \le M(d_1), \ 1 \le l \le L(d_2). \tag{36}$$

Then $\eta_d \in \mathrm{DM}^p(0,T;\mathbb{R}^m)$.

Proof. Put $\lambda_{jk} = q_{jk}/\sum_{l=1}^{L(d_2)} q_{jl}$ for $j = 1, \ldots, M(d_1)$. Then $\lambda_{jl} \ge 0$ and $\sum_{l=1}^{L(d_2)} \lambda_{jl} = 1$. Then η_d given by (34) defines through the correspondence (15) the pair $(\sigma, \hat\nu)$ given by

$$\hat\nu_t = \sum_{l=1}^{L(d_2)} \lambda_{jl} \delta_{s^l_{d_2}} \ , \quad d_\sigma(t) = \sum_{l=1}^{L(d_2)} q_{jl}, \quad t \in E^1_j \ , \tag{37}$$

where d_σ denotes the density of σ which is thus absolutely continuous with respect to the Lebesgue measure. It remains to show that $(\sigma, \hat\nu)$ defined by (37) belongs to $\mathcal{DM}^p(0,T;\mathbb{R}^m)$.

By (35), we have

$$\sum_{l=1}^{L(d_2)} \frac{\lambda_{jl}}{1+|s^l_{d_2}|^p} = \frac{1}{\sum_{k=1}^{L(d_2)} q_{jk}} \sum_{l=1}^{L(d_2)} \frac{q_{jl}}{1+|s^l_{d_2}|^p} = \frac{1}{\sum_{l=1}^{L(d_2)} q_{jl}} \ ,$$

so that

$$q_{jl} = \lambda_{jl} \left(\sum_{k=1}^{L(d_2)} \frac{\lambda_{jk}}{1+|s^k_{d_2}|^p} \right)^{-1}$$

and, for $t \in E^1_j$, one gets

$$d_\sigma(t) = \sum_{l=1}^{L(d_2)} q_{jl} = \left(\sum_{l=1}^{L(d_2)} \frac{\lambda_{jl}}{1+|s^l_{d_2}|^p} \right)^{-1} = \left(\int_{\mathbb{R}^m} \frac{\hat{\nu}_t(\mathrm{d}s)}{1+|s|^p} \right)^{-1} .$$

Now one can easily use Theorem 2.1. First, σ is positive. The condition (35) ensures that $\int_{\mathbb{R}^m} \hat{\nu}_t(\mathrm{d}s) > 0$ and, finally, the measure $\sigma_{\hat{\nu}}$ from Theorem 2.1 is absolutely continuous with the density $\left(\int_{\mathbb{R}^m} \frac{\hat{\nu}_t(\mathrm{d}s)}{1+|s|^p}\right)^{-1} \int_{\mathbb{R}^m} \hat{\nu}_t(\mathrm{d}s)$. $\square$

Note that the subset of $\mathrm{DM}^p(0,T;\mathbb{R}^m)$ containing measures satisfying (34)-(36) is convex. Indeed, for arbitrary $\lambda \in [0,1]$ and η^1_d, η^2_d fulfilling (34)-(36), the measure $\lambda\eta^1_d + (1-\lambda)\eta^2_d$ obeys again (34)-(36) with $q_{jl} := \lambda q^1_{jl} + (1-\lambda) q^2_{jl}$, where q^1_{jl} and q^2_{jl} correspond to η^1_d and η^2_d, respectively.

4.2 Proposition. *For any $\eta \in \mathrm{DM}^p(0,T;\mathbb{R}^m)$ there is $\{\eta_d\}_{d>0} \subset \mathrm{DM}^p(0,T;\mathbb{R}^m)$, η_d in the form (34)-(36), such that w^*-$\lim_{d\to 0} \eta_d = \eta$.*

As already mentioned, the proof of this assertion requires a construction of a finer local compactification of $L^p(0,T;\mathbb{R}^m)$. We define $\mathrm{Car}^p(0,T;\mathbb{R}^m) = \{h : (0,T) \times \mathbb{R}^m \to \mathbb{R}$ Carathéodory; $\exists \alpha \in L^1(0,T),\ \beta \in \mathbb{R}^+;\ |h(t,s)| \le \alpha(t) + \beta |s|^p\}$. Following [21], we will endow $\mathrm{Car}^p(0,T;\mathbb{R}^m)$ by the norm

$$\|h\|_{\mathrm{Car}^p(0,T;\mathbb{R}^m)} := \inf_{\substack{\forall (t,s)\in(0,T)\times\mathbb{R}^m: \\ |h(t,s)|\le\alpha(t)+\beta|s|^p}} \|\alpha\|_{L^1(0,T)} + \beta \ . \tag{38}$$

We will consider a separable linear subspace H of $\mathrm{Car}^p(0,T;\mathbb{R}^m)$ containing a coercive function, e.g., $h(t,s) = |s|^p$. Following [21] we define the embedding $i_H : L^p(0,T;\mathbb{R}^m) \to H^*$ by

$$\langle i_H(u), h \rangle := \int_0^T h(t,u(t))\,\mathrm{d}t \tag{39}$$

for any $u \in L^p(0,T;\mathbb{R}^m)$ and $h \in H$. Moreover, we put

$$Y^p_H(0,T;\mathbb{R}^m) := \mathrm{w}^*\text{-cl}\, i_H(L^p(0,T;\mathbb{R}^m)). \tag{40}$$

The set $Y^p_H(0,T;\mathbb{R}^m)$ will be addressed as the set of generalized Young functionals. We say that $\eta \in Y^p_H(0,T;\mathbb{R}^m)$ is generated by $\{u_k\}_{k\in\mathbb{N}}$ if $\eta = \mathrm{w}^*\text{-}\lim_{k\to\infty} i_H(u_k)$. We recall that $Y^p_H(0,T;\mathbb{R}^m)$ makes a convex locally compact envelope of $L^p(0,T;\mathbb{R}^m)$; for a detailed investigation of it we refer to [21].

Let us define $\Upsilon^p_{\mathcal{R}} := \{v \in C(\mathbb{R}^m);\ v(s) = v_0(s)(1+|s|^p)\ ,\ v_0 \in \mathcal{R}\}$ with $\mathcal{R}$ from (9) and then

$$H := G_0 \otimes \Upsilon^p_{\mathcal{R}} = \left\{ h;\ h(t,s) = \sum_{i=1}^k g(t)v(s)\ ,\quad g\in G_0\ ,\ v\in\Upsilon^p_{\mathcal{R}}\ ,\ k\in\mathbb{N} \right\}\ ,$$

where

$$G_0 := \bigcup_{d>0} G_d \quad \text{with} \quad G_d := \left\{ g \in L^\infty(0,T); \ \ \forall E \in \mathcal{T}^1_d : \ \ g|_E \in C(\bar{E}) \right\} \ .$$

This choice of G_0 was already used in [21, Sec. 3.5]; note that such H is separable in the norm (38) because both $\mathcal{R} \subset C(\mathrm{I\!R}^m)$ and $G_0 \subset L^\infty(0,T)$ are separable.

Following [21, Sec. 3.5] we define a projector $P^1_{d_1} : H \to H$ by

$$[P^1_{d_1} h](t,s) := \frac{1}{|E|} \int_E h(\tau, s)\, \mathrm{d}\tau \ , \qquad t \in E, \quad h \in H;$$

$E \subset [0,T]$ denotes the current subinterval of the equidistant partition $\mathcal{T}^1_{d_1}$. Obviously, $P^1_{d_1}$ makes an interval-wise constant approximation of $h(\cdot, s)$. Analogously, the projector $\hat{P}^1_{d_1} : G_0 \to G_0$ is defined by

$$[\hat{P}^1_{d_1} g](t) = \frac{1}{|E|} \int_E g(\tau)\, \mathrm{d}\tau \ , \qquad t \in E, \quad g \in G_0.$$

Furthermore, we define $\hat{P}^2_d : C(\gamma \mathrm{I\!R}^m) \to C(\gamma \mathrm{I\!R}^m)$ by assigning to each $v \in C(\gamma \mathrm{I\!R}^m)$ an element-wise affine interpolation coinciding with v at any mesh point of $\mathcal{T}^2_{d_2}$, i.e.

$$[\hat{P}^2_{d_2} h_0](t,s) := \sum_{l=1}^{L(d_2)} \frac{h(t, s^l_{d_2})}{1 + |s^l_{d_2}|^p} v^l_{d_2}(s) \ ,$$

with $\{v^l_d\}_{1 \le l \le L(d_2)}$ denoting a basis of the finite element subspace $\hat{P}^2_{d_2} C(\gamma \mathrm{I\!R}^m) \subset C(\gamma \mathrm{I\!R}^m)$ such that, for all $s \in \gamma \mathrm{I\!R}^m$, $\sum_{l=1}^{L(d_2)} v^l_{d_2}(s) = 1$ and $v^l_{d_2}(s) \ge 0$, and that $v^l_{d_2}(s^l_{d_2}) = 1$ for any mesh point $s^l_{d_2}$ of $\mathcal{T}^2_{d_2}$; recall that $L(d_2)$ denoted the number of the mesh points of $\mathcal{T}^2_{d_2}$. Then we define the projector $P^2_{d_2} : H \to H$ by

$$\left[P^2_{d_2} h \right](t,s) := \left[\hat{P}^2_{d_2} (h_0(t, \cdot)) \right](s)(1 + |s|^p) \ , \tag{41}$$

see also [21, Example 3.5.5]. Eventually, for $d = (d_1, d_2)$ we define the projector $P_d : H \to H$ by

$$P_d := P_{(d_1, d_2)} := P^1_{d_1} P^2_{d_2} = P^2_{d_2} P^1_{d_1} \ .$$

4.3 Lemma. *For any $h(t,s) = g(t) w(s)(1 + |s|^p)$ with $g \in G_0$ and $w \in \mathcal{R}$, the following holds:*

$$\lim_{d_1 \to 0} \|h - P^1_{d_1} h\|_{\mathrm{Car}^p(0,T;\mathrm{I\!R}^m)} = 0 \ , \tag{42}$$

$$\lim_{d_2 \to 0} \|h - P^2_{d_2} h\|_{\mathrm{Car}^p(0,T;\mathrm{I\!R}^m)} = 0 \ , \tag{43}$$

$$\lim_{d \to 0} \|h - P_d h\|_{\mathrm{Car}^p(0,T;\mathrm{I\!R}^m)} = 0 \ . \tag{44}$$

Proof. Take some $\tilde{d}_1 > 0$ fixed. Then g is continuous on the closure of $\tilde{E} \in \mathcal{T}^1_{\tilde{d}_1}$ and thus, for any $t_1, t_2 \in \tilde{E}$, $|g(t_1) - g(t_2)| \leq \mu_{\tilde{E}}(|t_1 - t_2|)$ with $\mu_{\tilde{E}}$ being the continuity modulus of g restricted on the closure of $\tilde{E}$. We recall that $\lim_{r\to 0} \mu_{\tilde{E}}(r) = 0$ for any $\tilde{E} \in \mathcal{T}^1_{\tilde{d}_1}$. One can estimate for any $d_1 < \tilde{d}_1$

$$\begin{aligned}
\|h - P^1_{d_1} h\|_{\mathrm{Car}^p(0,T;\mathbb{R}^m)} &\leq \sup_{s\in\gamma\mathbb{R}^m} |w(s)| \|g - \hat{P}^1_{d_1} g\|_{L^1(0,T)} \\
&= \|w\|_{C(\gamma\mathbb{R}^m)} \sum_{E\in\mathcal{T}^1_{d_1}} \int_E \left| g(t) - \frac{1}{|E|}\int_E g(\tau)\,\mathrm{d}\tau \right| \mathrm{d}t \\
&\leq \|w\|_{C(\gamma\mathbb{R}^m)} \sum_{E\in\mathcal{T}^1_{d_1}} \frac{1}{|E|} \int_E \int_E |g(t) - g(\tau)|\,\mathrm{d}\tau\,\mathrm{d}t \\
&\leq \|w\|_{C(\gamma\mathbb{R}^m)} \sum_{\tilde{E}\in\mathcal{T}^1_{\tilde{d}_1}} \sum_{E\subset\tilde{E}} \frac{1}{|E|} \int_E \int_E \mu_{\tilde{E}}(d_1)\,\mathrm{d}\tau\,\mathrm{d}t \\
&\leq T\|w\|_{C(\gamma\mathbb{R}^m)} \sum_{\tilde{E}\in\mathcal{T}^1_{\tilde{d}_1}} \mu_{\tilde{E}}(d_1) \to 0 \text{ as } d_1 \to 0 \,.
\end{aligned}$$

The proof of (43) is similar; we exploit the uniform continuity of w (precisely of its extension) on the compact set $\gamma\mathbb{R}^m$. Eventually, (44) follows from (42) and (43) by [21, Prop. 3.5.3]. □

We can define the adjoint operator P^*_d to P_d obviously by $\langle \xi, P_d h\rangle = \langle P^*_d \xi, h\rangle$ for any $\xi \in Y^p_H(0,T;\mathbb{R}^m)$ and $h \in H$. Moreover, as $C([0,T]) \subset G_0$, $Y^p_H(0,T;\mathbb{R}^m)$ makes a finer convex local compactification of $L^p(0,T;\mathbb{R}^m)$ than $\mathrm{DM}^p(0,T;\mathbb{R}^m)$ which is equivalent with $Y^p_{H_0}(0,T;\mathbb{R}^m)$ for $H_0 := C([0,T]) \otimes \Upsilon^p_{\mathcal{R}}$; cf. [21]. This means that there exists an affine weakly* continuous surjection

$$\psi : Y^p_H(0,T;\mathbb{R}^m) \to \mathrm{DM}^p(0,T;\mathbb{R}^m) \,.$$

fixing $L^p(0,T;\mathbb{R}^m)$ in the sense $\psi \circ i_H = i_{H_0}$. This mapping is not injective, however.

4.4 Lemma. *It holds $P^*_d Y^p_H(0,T;\mathbb{R}^m) \subset Y^p_H(0,T;\mathbb{R}^m)$. Moreover, for any $\xi \in Y^p_H(0,T;\mathbb{R}^m)$, it holds $\mathrm{w}^*\text{-}\lim_{d\to 0} \xi_d = \xi$ with $\xi_d = P^*_d \xi \in Y^p_H(0,T;\mathbb{R}^m)$. Then also $\mathrm{w}^*\text{-}\lim_{d\to 0} \psi(\xi_d) = \psi(\xi)$.*

Proof. Utilizing [21, Prop. 3.5.9] we get that $P^*_d Y^p_H(0,T;\mathbb{R}^m) \subset Y^p_H(0,T;\mathbb{R}^m)$. By Lemma 4.3 one gets

$$\begin{aligned}
\langle \xi - \xi_d, h\rangle &= \langle \xi, h - P_d h\rangle \\
&\leq \|\xi\|_{\mathrm{Car}^p(0,T;\mathbb{R}^m)^*} \|h - P_d h\|_{\mathrm{Car}^p(0,T;\mathbb{R}^m)} \to 0. \qquad (45)
\end{aligned}$$

This shows $\xi_d \to \xi$ weakly* in H^*. Further, as ψ is a weakly* continuous surjection, $\psi(\xi_d) \to \psi(\xi)$ weakly* in H^*_0. □

Proof of Proposition 4.2. For the orientation, let us display our situation by the following diagram:

$$\begin{array}{ccccc} Y_H(0,T;\mathbb{R}^m) & \xrightarrow{P_d^*} & P_d^*Y_H(0,T;\mathbb{R}^m) & \subset & Y_H(0,T;\mathbb{R}^m)\\ \psi\big\downarrow & & \big\downarrow\psi & & \\ \mathrm{DM}^p(0,T;\mathbb{R}^m) & \supset & \{\eta\in\mathrm{DM}^p(0,T;\mathbb{R}^m);\ (34)\text{ holds}\} & & \end{array}$$

Let $\eta \in \mathrm{DM}^p(0,T;\mathbb{R}^m)$. Then, by the above diagram, there is $\xi \in Y_H^p(0,T;\mathbb{R}^m)$ (generally not unique) such that $\eta = \psi(\xi)$. We put

$$\eta_d = \psi(\xi_d)\ , \qquad \xi_d = P_d^*\xi\ .$$

By Lemma 4.4, we know that w*-$\lim_{d\to 0}\xi_d = \xi$, $\eta_d \in \mathrm{DM}^p(0,T;\mathbb{R}^m)$ and w*-$\lim_{d\to 0}\eta_d = \eta \in \mathrm{DM}^p(0,T;\mathbb{R}^m)$. Now it remains to show that η_d can be expressed in the form (34)-(36).

As H is separable, for any $\xi \in Y_H^p(0,T;\mathbb{R}^m)$ there is uniquely defined $\xi^\circ \in Y_H^p(0,T;\mathbb{R}^m)$, the so-called p-nonconcentrating modification of ξ, such that $\langle \xi, h\rangle = \langle \xi^\circ, h\rangle$ for any $h \in H$, $\lim_{|s|\to\infty} h(t,s)/|s|^p = 0$ for a.a. $t \in [0,T]$. Therefore, defining $\tilde{\xi} = \xi - \xi^\circ$ we have

$$\langle \xi, P_d h\rangle = \langle \xi^\circ, P_d h\rangle + \left\langle \tilde{\xi}, P_d h\right\rangle = \langle P_d^*\xi^\circ, h\rangle + \left\langle P_d^*\tilde{\xi}, h\right\rangle\ .$$

Due to [21, Prop. 3.4.15.] and [21, Sec. 3.5], $P_d^*\xi^\circ$ has the unique representation by an interval-wise homogeneous aggregated L^p-Young measure, say ν_d,

$$\nu_{d,t} = \sum_{l=1}^{L'(d_2)} \tilde{\lambda}_{jl}\delta_{s^l_{d_2}}\ , \text{ for any } t \in E^1_j \in \mathcal{T}_{d_1}\ ,$$

where we suppose that $s^l_{d_2} \in \mathbb{R}^m$ for $l = 1,\dots,L'(d_2)$ (while $s^l_{d_2} \in \gamma\mathbb{R}^m \setminus \mathbb{R}^m$ for $l = L'(d_2)+1,\dots,L(d_2)$), $\sum_{l=1}^{L'(d_2)}\tilde{\lambda}_{jl} = 1$ and $\tilde{\lambda}_{jl} \geq 0$. As L^p-Young measures can be embedded into p-nonconcentrating DiPerna-Majda measures (see [21, Remark 3.2.16.]), we obtain an interval-wise homogeneous aggregated p-nonconcentrating DiPerna-Majda measure $\eta^\circ_d \cong (\sigma^\circ_d, \hat{\nu}^\circ_d)$ with

$$d_{\sigma^\circ_d}(t) = \left(\sum_{l=1}^{L'(d_2)} \frac{\lambda_{jl}}{1+|s^l_{d_2}|^p}\right)^{-1}\ , \qquad \hat{\nu}^\circ_{d,t} = \sum_{l=1}^{L'(d_2)} \lambda_{jl}\delta_{s^l_{d_2}}\ , \text{ for any } t\in E^1_j\in\mathcal{T}_{d_1}\ ,$$

where $d_{\sigma^\circ_d}$ stands for the density of σ°_d and $\sum_{l=1}^{L'(d_2)}\lambda_{jl} = 1$ and $\lambda_{jl} \geq 0$. Finally, we define for any $j = 1,\dots,M(d_1)$ and any $l = 1,\dots,L'(d_2)$

$$q_{jl} = \lambda_{jl}\left(\sum_{l=1}^{L'(d_2)} \frac{\lambda_{jl}}{1+|s^l_{d_2}|^p}\right)^{-1} \tag{46}$$

and the measure η°_d having the representation

$$\langle \eta^\circ_d, h\rangle = \sum_{j=1}^{M(d_1)}\sum_{l=1}^{L'(d_2)} \int_{E^1_j} h_0(t, s^l_{d_2}) q_{jl}\,\mathrm{d}t \tag{47}$$

for any $h(t,s)=h_0(t,s)(1+|s|^p)$ with $h_0\in C([0,T])\otimes\mathcal{R}$. Note that, for any $1\le j\le M(d_1)$,

$$\sum_{l=1}^{L'(d_2)}\frac{q_{jl}}{1+|s^l_{d_2}|^p}=1\ . \tag{48}$$

It remains to approximate $\tilde\eta_d=\psi(P^*_d\tilde\xi)$, For $h(t,s)=g(t)w(s)(1+|s|^p)$ with $g\in C([0,T])$ and $w\in\mathcal{R}$, we can write

$$\langle\tilde\eta_d,h\rangle=\sum_{j=1}^{M(d_1)}\sum_{l=L'(d_2)+1}^{L(d_2)}\int_{E^1_j}g_j(t)w(s^l_{d_2})q_{jl}\mathrm{d}t\ , \tag{49}$$

where $g_j=P^1_{d_1}g(t)$ for $t\in E^1_j$, $L'(d_2)+1\le l\le L(d_2)$, $E^1_j\in\mathcal{T}^1_{d_1}$, $E^2_l\in\mathcal{T}^2_{d_2}$ and

$$q_{jl}=\int_{E^1_j}\int_{E^2_l}\hat\nu_t(\mathrm{d}s)\sigma(\mathrm{d}t)$$

with $\eta\cong(\sigma,\hat\nu)\in\mathcal{DM}^p(0,T;\mathrm{I\!R}^m)$. Combining now (47) and (49) and taking into account (46) and (48), we get that $\psi(\xi_d)$ is described in the form (34)-(36), which gives the desired result. □

5 Numerical approximation of the relaxed problem

As we cannot implement an arbitrary DiPerna-Majda measure $\eta\in\mathrm{DM}^p(0,T;\mathrm{I\!R}^m)$ on computers, we restrict ourselves to piecewise constant DiPerna-Majda measures given by (34) for given discretization parameter $d=(d_1,d_2)$, i.e. for given discretizations $\mathcal{T}^1_{d_1}$ and $\mathcal{T}^2_{d_2}$. For these measures the approximation of problem (RP') looks as follows:

$$(\mathrm{RP}_d)\quad\left\{\begin{array}{ll}\text{Minimize} & \tilde J(y,\dot y,q)=\displaystyle\sum_{l=1}^{L(d_2)}\int_0^T\frac{a(t,y,s^l_{d_2})+b(t,s^l_{d_2})}{1+|s^l_{d_2}|^p}q_l(t)\,\mathrm{d}t+f\left((y,\dot y)|_{t=T}\right)\\ \text{subject to} & \dot y=\displaystyle\sum_{l=1}^{L(d_2)}\frac{c(t,y,s^l_{d_2})+d(t,s^l_{d_2})}{1+|s^l_{d_2}|^p}q_l(t)\ ,\qquad(y,\dot y)|_{t=0}=y_0\ ,\\ & q_l(t)\ge0\quad\text{and}\quad\displaystyle\sum_{l=1}^{L(d_2)}\frac{q_l(t)}{1+|s^l_{d_2}|^p}=1,\quad t\in[0,T],\\ & q\ \ \text{piecewise constant on}\ \mathcal{T}^1_{d_1},\\ & (y,\dot y)\in W^1_\mu(0,T;\mathrm{I\!R}^n),\qquad q\in L^\infty(0,T;\mathrm{I\!R}^{L(d_2)}).\end{array}\right.$$

In accord with the formula (34), we will identify $q\in L^\infty(0,T;\mathrm{I\!R}^{L(d_2)})$ admissible for (RP_d) with $\eta\in\mathrm{DM}^p(0,T;\mathrm{I\!R}^m)$ given by the formula

$$\langle\eta,h\rangle=\sum_{l=1}^{L(d_2)}\int_0^T\frac{h(t,s^l_{d_2})}{1+|s^l_{d_2}|^p}q_l(t)\,\mathrm{d}t\ . \tag{50}$$

Then, in particular, $\bar{J}(y,\dot{y},\eta)=\tilde{J}(y,\dot{y},q)$. For simplicity, we suppose that all the integrals as well as the initial-value problem in (RP$_d$) can be evaluated exactly so that (RP$_d$) can already be implemented, as it is indeed the case in Section 6 below. The following two assertion establish the convergence of (RP$_d$) to (RP').

5.1 Proposition. *The discrete relaxed problems* (RP$_d$) *possess solutions. Moreover,* $\lim_{d\to 0}\min$(RP$_d$)= min(RP').

Proof. The existence of a solution to the discrete problem follows from the same arguments as in the proof of Proposition 3.2. Due to the coercivity of the problem its solutions identified via (50) are contained in a set $\psi(P_d^*(Y^p_{H,\rho}(0,T;\mathrm{I\!R}^m)))$ for some $\rho>0$ sufficiently large, where $Y^p_{H,\rho}(0,T;\mathrm{I\!R}^m):=\{\xi\in H^*;\xi=\text{w*-}\lim i_H(u_k),\ \|u_k\|_{L^p(0,T;\mathrm{I\!R}^m)}\le\rho\}$ is weakly* compact. As both P_d^* and ψ are weakly* continuous, $\psi(P_d^*(Y^p_{H,\rho}(0,T;\mathrm{I\!R}^m)))$ is weakly* compact, as well.

Let $\eta\in\mathrm{DM}^p(0,T;\mathrm{I\!R}^m)$ be a solution to (RP'). Then we know from Proposition 4.2 that w*-$\lim_{d\to0}\eta_d=\eta$ for some η_d admissible for (RP$_d$). Therefore, from the weak* continuity of $\eta\mapsto\bar{\Phi}(\eta)=\bar{J}(\bar{\pi}(\eta),\eta)$, we get $\lim_{d\to0}\bar{\Phi}(\eta_d)=\bar{\Phi}(\eta)$. Clearly, (as we use inner approximations) $\bar{\Phi}(\eta_d)\ge\bar{\Phi}(\eta)$. Now, if $\bar{\eta}_d\in\mathrm{DM}^p(0,T;\mathrm{I\!R}^m)$ is a solution to (RP$_d$), we have $\bar{\Phi}(\eta_d)\ge\bar{\Phi}(\bar{\eta}_d)\ge\bar{\Phi}(\eta)$. Finally, we have $\lim_{d\to0}\bar{\Phi}(\bar{\eta}_d)=\bar{\Phi}(\eta)$, or equivalently $\lim_{d\to0}\min$ (RP$_d$) = min (RP'). □

5.2 Proposition. *Let* η_d *be a solution to* (RP$_d$). *Then* $\{\eta_d\}_{d>0}$ *contains a weakly** *converging subsequence and the limit of each such a subsequence solves* (RP').

Proof. The problems (RP$_d$), $d>0$, are uniformly coercive. Indeed, by (22), we have $\bar{\Phi}(\eta)\ge\varepsilon\langle\eta,|s|^p/(1+|s|^p)\rangle$ for any $\eta\in\mathrm{DM}^p(0,T;\mathrm{I\!R}^m)$. Then, for some $u_0\in L^p(0,T;\mathrm{I\!R}^m)$ such that $i(u_0)$ is admissible for every (RP$_d$) we have $\Phi(u_0)\ge\bar{\Phi}(\eta)\ge\varepsilon\langle\eta,|s|^p/(1+|s|^p)\rangle$ for any solution η to (RP$_d$). Thus $\mathrm{DM}^p_\rho(0,T;\mathrm{I\!R}^m)$ with $\rho=\varepsilon^{-1/p}\Phi(u_0)^{1/p}$.

As $\mathrm{DM}^p_\rho(0,T;\mathrm{I\!R}^m)$ is sequentially compact $\{\eta_d\}_{d>0}$ contains a subsequence $\{\eta_{d'}\}_{d'>0}$ that converges for $d'\to0$ to some $\eta'\in\mathrm{DM}_\rho(0,T;\mathrm{I\!R}^m)$. We know due to the previous proposition that $\lim_{d'\to0}\min$ (RP$_{d'}$) $=\lim_{d'\to0}\bar{\Phi}(\eta_{d'})=$ min (RP'). As $\bar{\Phi}$ is weakly* continuous, we have $\lim_{d'\to0}\bar{\Phi}(\eta_{d'})=\bar{\Phi}(\eta')$ Therefore, η' minimizes $\bar{\Phi}$. □

6 Illustrative examples

The approximate relaxed problem (RP$_d$) is a minimization problem over the polyhedral set of the parameters $\{q_{jl}\}$ with $1\le j\le M(d_1)$ and $1\le l\le L(d_2)$. We calculated the Examples 1.1–4 whose objective functions are quadratic because they have the cost functional J additively splitted and quadratic in terms of the state y. Therefore, the resulting optimization problems can be solved even by a finite algorithm. Here we used Schittkowski's

QLD routine which is a part of the sequential quadratic programming package NLPQL [23]. All the computations were performed on SGI workstations. On each figure the left-hand side picture denotes the computed relaxed control and the right-hand side picture the corresponding state.

The first case is Example 1.1. Here, in the notation of (P), we have $n = m = 1$, $a = c = 0$, $b(t,s) = (2 - 2t + t^2)|s|$, $d(t,s) = s$, $f(r) = (r-1)^2$, and $T = 1.5$.

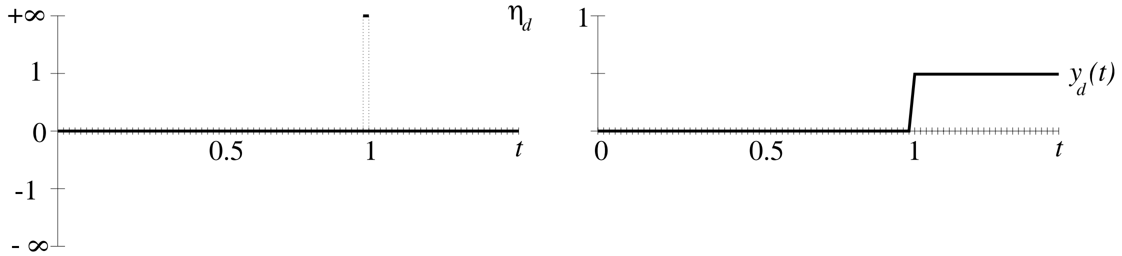

Figure 2: *Approximate solution to the relaxation of (1);* $M(d_1) = 80$*,* $L(d_2) = 20$*.*
Relaxed control (left) and corresponding state (right).

The second case is Example 1.2 with $\alpha = 1/3$. Here, $n = 2$, $m = 1$, $a(t,r,s) = r_2^2$, $b(t,s) = (2 - 2t + t^2)|s|$, $c = 0$, $d(t,s) = (|s|, \frac{1}{\alpha}\max(0,s) + \frac{1}{1-\alpha}\min(0,s))$, $f(r) = (r-1)^2$, and $T = 1.5$.

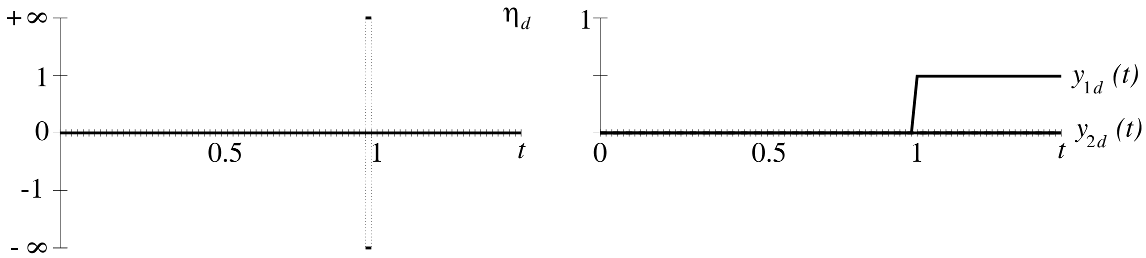

Figure 3: *Approximate solution to the relaxation of (4);* $M(d_1) = 80$*,* $L(d_2) = 20$*.*
Relaxed control (left) and corresponding state (right).

The third case is Example 1.3. Here $n = m = 1$, $a(t,r,s) = (r-t)^2$, $b(t,s) = e_\theta(s)$ with $\theta = 1/5$ where e_θ is given by (7), $c = 0$, $d(t,s) = s$, $f = 0$, and $T = 1.0$.

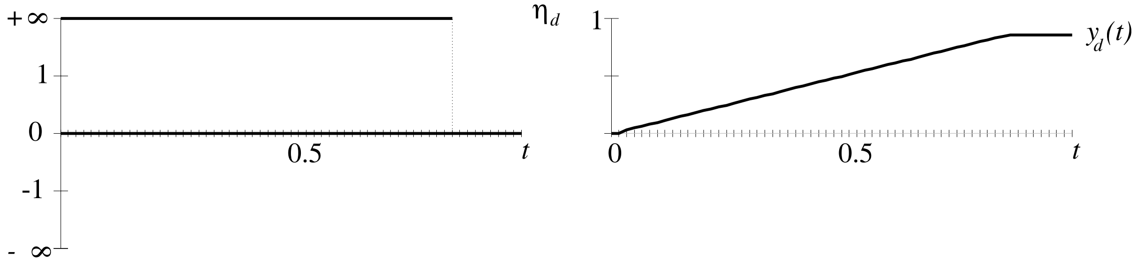

Figure 4: *Approximate solution to the relaxation of (6); $M(d_1) = 60$, $L(d_2) = 20$. Relaxed control (left) and corresponding state (right).*

Eventually, the last case is Example 1.4 with $\alpha = 1/3$. We have $n = 2$, $m = 1$, $a(t, r, s) = (r_1 - t)^2 + r_2^2$, $b(t, s) = e_\theta(s)$ where e_θ is given by (7) with $\theta = 1/5$, $c = 0$, $d(t, s) = (|s|, \frac{1}{\alpha}\max(0, s) + \frac{1}{1-\alpha}\min(0, s))$, $f = 0$, and $T = 1.0$

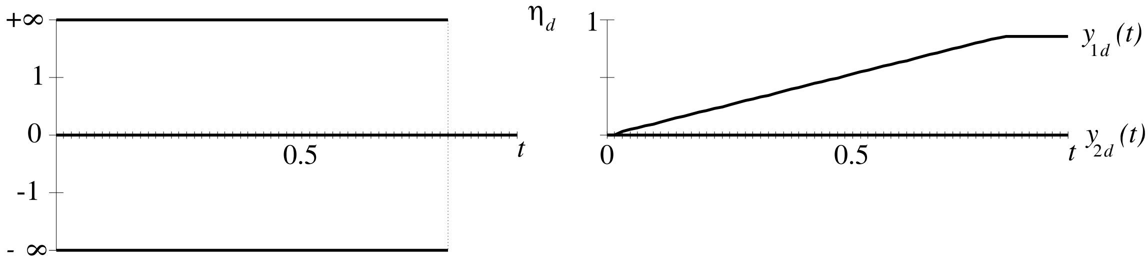

Figure 5: *Approximate solution to the relaxation of (8); $M(d_1) = 60$, $L(d_2) = 20$. Relaxed control (left) and corresponding state (right).*

Note that, in accord with (30)–(33), the calculated relaxed control is 2-atomic (see Figures 2 and 4) or 3-atomic (see Figures 3 and 5) for some t.

Acknowledgment: This work was supported by the grant No. A 107 5707 (Grant Agency of the Academy of Sciences) and by the grant No. 201/96/0228 (Grant Agency of the Czech Republic). The paper has been finished during a stay of M.K. in the Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany, whose support and hospitality is gratefully acknowledged.

References

[1] J.J. Alibert, G. Bouchitté: Non uniform integrability and generalized Young measures. *J. Convex Anal.* **4** (1997), 129–147.

[2] E.J. Balder: New existence results for optimal controls in the absence of convexity: the importance of extremality. *SIAM J. Control Optim.* **32** (1994), 890–916.

[3] J.M. Ball: A version of the fundamental theorem for Young measures. In: *PDEs and Continuum Models of Phase Transition.* (M.Rascle, D.Serre, M.Slemrod. eds.) Lecture Notes in Physics **344**, Springer, Berlin, 1989, pp.207–215.

[4] A. Blaquire: Impulsive optimal control with finite or infinite time horizon. *J. Optim. Th. Appl.* **46** (1985), 431–439.

[5] A.E. Bryson, Y.C. Ho: *Applied Optimal Control Theory.* Ginn & Co., Waltham, 1969.

[6] R.J. DiPerna, A.J. Majda: Oscillations and concentrations in weak solutions of the incompressible fluid equations. *Comm. Math. Physics* **108** (1987), 667–689.

[7] N. Dunford, J.T. Schwartz: *Linear Operators.* Part I, Interscience, New York, 1967.

[8] R. Engelking: *General topology.* 2^{nd} ed., PWN, Warszawa, 1985.

[9] W.M. Getz, D.H. Martin: Optimal control systems with state variable jump discontinuities. *J. Optim. Th. Appl.* **31** (1980), 195–205.

[10] E. Giusti: *Minimal Surfaces and Functions of Bounded Variations.* Birkhäuser, Boston, 1984.

[11] P.R. Halmos: *Measure Theory.* D. van Nostrand, Toronto, 1950.

[12] J. Kačur, J. Souček: Direct variational methods in nonreflexive spaces. *Mathematica Slovaca*, **29** (1979), 209–226.

[13] M. Kružík, T. Roubíček: Explicit characterization of L^p-Young measures. *J. Math. Anal. Appl.* **198** (1996), 830–843.

[14] M. Kružík, T. Roubíček: On the measures of DiPerna and Majda. *Math. Bohemica* **120** (1997), 411–430.

[15] M. Kružík, T. Roubíček: Some geometric properties of generalized Young functionals. *Proc. Royal Soc. Edinburgh, Sect.A* (accepted).

[16] X. Liu: Impulsive Control and Optimization. *Appl. Math. Comp.* **73** (1995), 77–98.

[17] B.M. Miller: The generalized solutions of nonlinear optimization problems with impulse control. *SIAM J. Control Optimization* **34** (1996), 1420–1440.

[18] P. Pedregal: Numerical approximation of parametrized measures. *Numer. Funct. Anal. Opt.* **16** (1995), 1049–1066.

[19] P. Pedregal: *Parametrized Measures and Variational Principles.* Birkhäuser, Basel, 1997.

[20] R. Rempala, J. Zabczyk: On the maximum principle for deterministic impulse control problems. *J. Optim. Th. Appl.* **59** (1988), 281–288.

[21] T. Roubíček: *Relaxation in Optimization Theory and Variational Calculus.* W. de Gruyter, Berlin, 1997.

[22] T. Roubíček: Nonconcentrating generalized Young functionals. *Commentationes Math. Univ. Carolinae* **38** (1997), 91–99.

[23] K. Schittkowski: NLPQL: A Fortran subroutine solving constrained nonlinear programming problems. *Annals of Operation Research* **5** (1985-6), pp. 485–500.

[24] M.E. Schonbek: Convergence of solutions to nonlinear dispersive equations. *Comm. in Partial Diff. Equations* **7** (1982), 959–1000.

[25] A. Seierstad: Existence of and optimal control with sparse jumps in the state variable. *J. Optim. Th. Appl.* **45** (1985), 265–293.

[26] J. Souček: Spaces of functions on domain Ω, whose k-th derivatives are measures defined on $\bar{\Omega}$. *Čas. pro Pěst. Mat.* **97** (1972), 10–46.

[27] L. Tartar: Compensated compactness and applications to partial differential equations, In: *Nonlinear Analysis and Mechanics* (R.J.Knops, ed.) Heriott-Watt Symposium IV, Pitman Res. Notes in Math. **39**, San Francisco, 1979.

[28] M. Valadier: Young measures, In: *Methods of Nonconvex Analysis* (A.Cellina, ed.) Lecture Notes in Math. **1446**, Springer, Berlin, 1990, 152–188.

[29] J. Warga: Variational problems with unbounded controls. *SIAM J. Control* **3** (1965), 424–438.

[30] J. Warga: *Optimal Control of Differential and Functional Equations.* Academic Press, New York, 1972.

[31] L.C. Young: Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Comptes Rendus de la Société des Sciences et des Lettres de Varsovie*, Classe III **30** (1937), 212–234.