Solving Problems in Convex Optimal Control by Progressive Decoupling in the Dynamics

R. Tyrrell Rockafellar[1]

Abstract

Primal and dual problems of optimal control in the format of convex problems of Bolza can simultaneously be solved by a specialized application of the progressive decoupling algorithm by translating them to a linkage problem format. Although explained and justified here only in pure form, the procedure has the unprecedented feature that, in each iteration, the minimization subproblem decomposes into a separate proximal minimization subproblem at each instant of time, along with another proximal minimization concerning a trajectory's endpoints.

Keywords: *progressive decoupling algorithm, dynamical decoupling, convex optimal control, dual problems of Bolza, generalized Hamiltonians, generalized Euler-Lagrange conditions, generalized transversality conditions.*

Version of 12 April 2024

Dedicated to Hélene Frankowska for her 70th birthday

MSC classifications: 49M27, 49M29, 49N05, 65K10

[1]University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195-4350;
E-mail: *rtr@uw.edu*, URL: sites.math.washington.edu/~rtr/mypage.html

1 Introduction

The idea of solving a large-scale problem of optimization by reducing it to solving much smaller problems in a sequence of iterations is age-old. In a common version, the given problem is comprised of separate "components" that might be solved independently of each other, but can't because of "linkages" between them which require attention. A scheme might succeed, however, if it could proceed by iterations in which the linkages are temporarily decoupled.

Such a scheme is the *progressive decoupling algorithm* (PDA) put forward in [15]. Here we explain how it can be applied to problems in a dynamical format in which the "decoupling" means that in each iteration a separate optimization problem is solved for each instance in time. The results from this *dynamical decoupling* are projected to produce a sequence of trajectories that converges eventually to an optimal trajectory.

This is a novel approach which could be articulated on many levels with different degrees of detail and numerical sophistication, but here, in a first presentation, we focus just on the basics. In particular, we emphasize convexity and the help from duality that comes with it. What comes out is an "ideal procedure" only, because important practical details for its execution, such as stopping criteria for inexactly computed iterations, aren't addressed. But that could be a beacon for further analysis and experimentation.

The general *linkage problem* behind this concerns a Hilbert space $\mathcal{H}$, a closed subspace $\mathcal{S} \subset \mathcal{H}$ and a closed (lower semicontinuous) convex functional $\Phi : \mathcal{H} \to \overline{I\!R} = [-\infty, \infty]$ that is proper (not $\equiv \infty$ and never taking on $-\infty$), with the set $\operatorname{dom} \Phi = \{ z \mid \Phi(z) < \infty \}$ then being convex and nonempty, although not necessarily closed. The task is to

$$(P_{\text{link}}) \qquad \text{minimize } \Phi(z) \text{ over } z \in \mathcal{S},$$

where the feasibility of z is tied not only to z being in $\mathcal{S}$, but also to its belonging to $\operatorname{dom} \Phi$.

The *progressive decoupling algorithm* for solving this problem emerges from the optimality condition associated with it in terms of the closed subspace $\mathcal{S}^{\perp}$ orthogonal to $\mathcal{S}$ in $\mathcal{H}$ and the subgradients of Φ in the sense of convex analysis, namely

$$\begin{aligned} &z \in \mathcal{S} \text{ and } \exists\, w \in \mathcal{S}^{\perp} \text{ with } w \in \partial \Phi(z), \text{ or} \\ &\text{equivalently, } z \in \operatorname{argmin}_z \{ \Phi(z) - \langle w, z \rangle \}. \end{aligned} \tag{1.1}$$

This is sufficient not only for z to solve (P) as the *primal* problem, but at the same time for w to solve the *dual* problem for the convex functional Φ^* conjugate to Φ,

$$(D_{\text{link}}) \qquad \text{maximize } -\Phi^*(w) \text{ over } w \in \mathcal{S}^{\perp},$$

and then

$$\text{minimum value in } (P_{\text{link}}) = \text{ maximum value in } (D_{\text{link}}) \tag{1.2}$$

The condition in (1.1) is furthermore necessary for the optimality of z under a constraint qualification,[2] but here we will always be working in a setting where we already know from the outset that some z and w satisfying (1.1) do exist.

The progressive decoupling algorithm exploits the minimization form of the subgradient relation in (1.1) while stabilizing it with a proximal term. The kth iteration begins with some $z^k \in \mathcal{S}$ and $w^k \in \mathcal{S}^{\perp}$ (the initial z^0 and w^0 in these subspaces can be chosen arbitrarily) and determines

$$\hat{z}^k = \operatorname{argmin}_z \Big\{ \Phi(z) - \langle w^k, z \rangle + \frac{r}{2} ||z - z^k||^2 \Big\}, \text{ where } r > 0. \tag{1.3}$$

[2]For example, if the projection of $\operatorname{dom} \Phi$ on $\mathcal{S}^{\perp}$ has the property that no ray emanating from the origin of $\mathcal{S}^{\perp}$ immediately exits from it; see [8, Theorem 18(c)].

It then updates by

$$z^{k+1} = \mathrm{proj}_{\mathcal{S}}\, \hat{z}^k, \qquad w^{k+1} = w^k - r[\hat{z}^k - z^{k+1}] = w^k - r\, \mathrm{proj}_{\mathcal{S}^\perp}\, \hat{z}^k. \tag{1.4}$$

where $\mathrm{proj}_{\mathcal{S}}$ is the mapping that projects elements of $\mathcal{H}$ onto the subspace $\mathcal{S}$. The generated sequence $\{(z^k, w^k)\}$ is guaranteed to converge in the weak topology of $\mathcal{H}$ to some particular $(\bar{z}, \bar{w})$ satisfying (1.1) as long as at least one such pair exists [15, Theorem 1]. The "decoupling" is brought about by the absence of the linkage constraint $z \in \mathcal{S}$ in (1.3).

The problems of optimal control that we propose to solve by this procedure are the generalized *problems of Bolza* introduced in [5] and underpinned with existence theorems and duality theorems in [7]. They concern trajectories of states over time represented by *arcs* as functions $x(\cdot) : [t_0, t_1] \to I\!R^n$ such that the derivative $\dot{x}(t)$ exists almost everywhere and $x(t) = x(t_0) + \int_{t_0}^{t} \dot{x}(s)ds$. The fundamental problem is to

$$(P_{\rm arcs}) \qquad \text{minimize } \mathcal{J}_{L,l}(x(\cdot)) = \int_{t_0}^{t_1} L(x(t), \dot{x}(t))dt + l(x(t_0), x(t_1)),$$

where L and l are closed proper convex functions on $I\!R^n \times I\!R^n$. Constraints are captured implicitly in this neoclassical format by ∞ values of L and l. Again there is a dual problem, which is to

$$(D_{\rm arcs}) \qquad \text{maximize } -\mathcal{J}_{\widetilde{L},\widetilde{l}}(p(\cdot)) = -\int_{t_0}^{t_1} \widetilde{L}(p(t), \dot{p}(t))dt - \widetilde{l}(p(t_0), p(t_1))$$

in which $\widetilde{L}$ and $\widetilde{l}$ are obtained from the convex functions L^* and l^* conjugate to L and l by

$$\widetilde{L}(p, q) = L^*(q, p), \qquad \widetilde{l}(p_0, p_1) = l^*(p_0, -p_1). \tag{1.5}$$

Controls, just like constraints, are implicit in these statements through the ways that L might be specified. For instance, in the case where

$$L(x, y) = \min_u \{\, K(Cx, u) \,|\, Ax + Bu = y \} \tag{1.6}$$

for a closed proper convex function K, problem $(P_{\rm arcs})$ would in effect seek a control function $u(\cdot)$ to

$$\text{minimize } \int_{t_0}^{t_1} K(Cx(t), u(t))dt + l(x(t_0), x(t_1)) \text{ with } \dot{x}(t) = Ax(t) + Bu(t). \tag{1.7}$$

The choice of K could enforce by ∞ values a constraint like $u(t) \in U$ or $u(t) \in U(x(t))$, say. The associated dual problem, determined from (1.5) and (1.6), seeks instead to

$$\text{maximize } -\int_{t_0}^{t_1} \widetilde{K}(B^*p(t), v(t))dt - \widetilde{l}(p(t_0), p(t_1)) \text{ with } \dot{p}(t) = -A^*p(t) + C^*v(t), \tag{1.8}$$

where

$$\widetilde{K}(B^*p, v) = K^*(v, B^*p), \tag{1.9}$$

and thus brings in dual control variables along with the primal control variables. All the technical details behind this, like measurable selections, are worked out in [5], where control models are explored directly, as they are also [11] and [13].

Theorems in [7] provide constraint qualifications under which both $(P_{\rm arcs})$ and $(D_{\rm arcs})$ have solutions, and the minimum value in $(P_{\rm arcs})$ equals the maximum value in $(D_{\rm arcs})$. But this is in the case where the arcs $x(\cdot)$ and $p(\cdot)$ are only asked to be absolutely continuous, that is, with their derivatives $\dot{x}(\cdot)$ and $\dot{p}(\cdot)$ being $\mathcal{L}^1$ functions. Our aim of applying the progressive decoupling algorithm to solve them requires, because of the Hilbert space context of the method, that $\dot{x}(\cdot)$ and $\dot{p}(\cdot)$ be $\mathcal{L}^2$ functions. We must therefore adapt the existence and duality theory in [7] to operate for $\mathcal{L}^2$ derivatives instead of $\mathcal{L}^1$ derivatives, in addition to managing to reformulate the primal and dual problems of Bolza as primal and dual problems in the linkage scheme.

That's the agenda for Section 2. In the $\mathcal{L}^2$ adaptation there we build on growth conditions on the Hamiltonian function

$$H(x,p) = \sup_y \{\, p{\cdot}y - L(x,y)\}, \ \text{ with } \ L(x,y) = \sup_p \{\, p{\cdot}y - H(x,p)\}, \tag{1.10}$$

that were introduced in the development of a Hamilton-Jacobi theory in [20], [21], to be able to cover convex value functions like

$$V(\tau,\xi) = \min_{x(\cdot)} \Big\{ \int_\tau^T L(x(t),\dot{x}(t))dt + g(x(T)) \,\Big|\, x(\tau) = \xi \Big\}. \tag{1.11}$$

The standard conditions for getting "viscosity" Hamilton-Jacobi solutions turned out to be incompatible with the convex duality, so growth had to be somehow calibrated differently. Note that the defining minimization in (1.11) is an example of $(P_{\rm arcs})$ with $[t_0,t_1]$ replaced by $[\tau,T]$ and with (τ,ξ) treated as a parameter element. A method for solving $(P_{\rm arcs})$, as proposed here, might therefore have implications also for solving Hamilton-Jacobi equations.

In Section 3 the progressive decoupling algorithm returns to center stage for application to the control problems in their linkage re-expressions. A key issue is developing a formula for the projection mapping in the update (1.4) in that particular context.

Finally, in Section 4, the potential for extensions to broader classes of control problems is surveyed along with extra features that might be added to the dynamical decoupling procedure and its ways of implementation.

2 Translating the control problems to the linkage format

The interval $[t_0,t_1]$ is fixed in what follows, so we let $\mathcal{L}^p_n$ stand for $\mathcal{L}^p([t_0,t_1],\mathbb{R}^n)$. Likewise, we let $\mathcal{A}^p_n$ stand then for the space of arcs $x(\cdot) : [t_0,t_1] \to \mathbb{R}^n$ such that the derivative function $\dot{x}(\cdot)$ belongs to $\mathcal{L}^p_n$. In the existence theory in [7] for solutions to problems $(P_{\rm arcs})$ and $(D_{\rm arcs})$, the minimization was over all $x(\cdot)$ and $p(\cdot)$ in $\mathcal{A}^1_n$. It's important for our purposes, however, to be able to restrict the arcs to lie in $\mathcal{A}^2_n$ without jeaprodizing the results that were obtained without that restriction.

Quadratic aspects of growth conditions can be brought in to achieve that. It will be convenient to express them in terms of the function

$$j(z) = \tfrac{1}{2}|z|^2, \qquad \text{where } |\cdot| \text{ is the canonical norm when } z \in \mathbb{R}^n, \tag{2.1}$$

this being the only convex function that is self-conjugate, $j^* = j$. The growth conditions will first be imposed on the Hamiltonian function H in (1.10), where a primal-dual symmetry is evident, but their effect on L and $\widetilde{L}$ will afterward be clarified.

Assumption 2.1 (primal-dual growth conditions on the Hamilitonian).
(a) *There are constants $\sigma > 0$, $\alpha > 0$, $\beta > 0$, γ, such that*

$$H(x,p) \le \sigma j(p) + (\alpha|p| + \beta)|x| + \gamma. \tag{2.2}$$

(b) *There are constants $\tilde{\sigma} > 0$, $\tilde{\alpha} > 0$, $\tilde{\beta} > 0$, $\tilde{\gamma}$, such that*

$$H(x,p) \geq -\tilde{\sigma} j(x) - (\tilde{\alpha}|x| + \tilde{\beta})|p| - \tilde{\gamma}. \tag{2.3}$$

In general, the Hamiltonian expression $H(x,p)$, as obtained in (1.1) by taking the partial conjugate of the convex function L, is concave in x and convex in p, and could take on both ∞ and $-\infty$. These conditions require, however, that $-\infty < H(x,p) < \infty$ for all x and p. The dual Hamiltonian associated instead with $\widetilde{L}$, namely

$$\widetilde{H}(p,x) = \sup_q \{\, q{\cdot}x - \widetilde{L}(p,q)\}, \;\; \text{with} \;\; \widetilde{L}(p,q) = \sup_x \{\, q{\cdot}x - \widetilde{H}(p,x)\}, \tag{2.4}$$

comes out then as

$$\widetilde{H}(p,x) = -H(x,p), \tag{2.5}$$

whereas otherwise closure operations might have to be brought in. The finite concave-convex function H is not just continuous but locally Lipschitz continuous everywhere on $\mathbb{R}^n \times \mathbb{R}^n$. The relationship in (2.5) makes clear that the primal effect on L of the assumptions on H is completely reflected in their dual effect on $\widetilde{L}$.

Proposition 2.2 (growth assumption equivalences)**.** *Condition (a) in Assumption 2.1 is equivalent to having*

$$L(x,y) \geq \sigma^{-1} j\Big(\max\{0, |y| - \alpha|x|\}\Big) - \beta|x| - \gamma, \tag{2.6}$$

while condition (b) is equivalent to having

$$\widetilde{L}(p,q) \geq \tilde{\sigma}^{-1} j\Big(\max\{0, |q| - \tilde{\alpha}|p|\}\Big) - \tilde{\beta}|p| - \tilde{\gamma}. \tag{2.7}$$

Proof. To establish the first claim, we take conjugates on both sides of (2.2). We see that

$$\begin{aligned} L(x,y) + \beta|x| + \gamma &\geq \sup_p \{\, p{\cdot}y - \sigma j(p) - \alpha|x||p|\} = \sup_{\pi\geq 0}\Big\{ \sup_{|p|=\pi} \{\, p{\cdot}y - \sigma j(p) - \alpha|x||p|\}\Big\} \\ &= \sup_{\pi\geq 0}\Big\{\, \pi{\cdot}|y| - \sigma j(\pi) - \alpha|x|\pi\Big\} = \sigma^{-1} j\Big(\max\{0, |y| - \alpha|x|\}\Big), \end{aligned}$$

because the conjugate of the convex function on $\mathbb{R}$ that's the sum of j and the indicator of $[0,\infty)$ is the function that equals j on $[0,\infty)$ and 0 on $(-\infty, 0]$. The truth of the second claim follows then by the symmetry in (2.4) and (2.5). □

Corollary 2.3 (feasibility implication)**.** *Under Assumption 2.1, an arc $x(\cdot)$ for which the integral component of the functional $\mathcal{J}_{L,l}$ in problem (P_{arcs}) is not ∞ must be in $\mathcal{A}^2_n$. Likewise, an arc $p(\cdot)$ for which the integral component of the functional $\mathcal{J}_{\widetilde{L},\widetilde{l}}$ in problem (D_{arcs}) is not ∞ must be in $\mathcal{A}^2_n$.*

Proof. The finiteness of $\int_{t_0}^{t_1} L(x(t),\dot{x}(t))dt$ necessitates that of $\int_{t_0}^{t_1} j(\max\{0, |\dot{x}(t)| - \alpha|x(t)|\})dt$ and implies the existence of a nonnegative function $\mu(\cdot) \in \mathcal{L}^2_1$ such that $|\dot{x}(t)| \leq \mu(t) + \alpha|x(t)|$. Since $x(\cdot)$ is bounded over $[t_0, t_1]$, the function $\mu(\cdot) + \alpha x(\cdot)$ likewise belongs to $\mathcal{L}^2_1$, and the same then holds for $|\dot{x}(\cdot)|$, so that $\dot{x}(\cdot) \in \mathcal{L}^2_n$. Similarly for $\widetilde{L}$. □

This feasibility result allows us to apply the main result of [7] to the primal and dual problems of Bolza even when the arcs in them are restricted from the beginning to lie in $\mathcal{A}^2_n$. The constraint qualifications utilized in this result are conditions on *attainability* in being concerned with the relationship between the convex sets

$$\begin{aligned} C_l &= \operatorname{dom} l = \{\, (x_0,x_1) \,|\, l(x_0,x_1) < \infty\}, \\ C_L &= \Big\{\, (x_0,x_1) \,\Big|\, \exists x(\cdot) \text{ with } \textstyle\int_{t_0}^{t_1} L(x(t),\dot{x}(t))dt < \infty, \; (x(t_0),x(t_1)) = (x_0,x_1)\Big\}, \end{aligned} \tag{2.8}$$

in the primal problem and, in parallel, the relationship between the convex sets

$$\begin{array}{c} C_{\widetilde{l}} = \operatorname{dom} \widetilde{l} = \{\, (p_0, p_1) \,|\, \widetilde{l}(p_0, p_1) < \infty \}, \\ C_{\widetilde{L}} = \left\{\, (p_0, p_1) \,\middle|\, \exists p(\cdot) \text{ with } \int_{t_0}^{t_1} \widetilde{L}(p(t), \dot{p}(t))dt < \infty,\ (p(t_0), p(t_1)) = (p_0, p_1) \right\}, \end{array} \tag{2.9}$$

in the dual problem. We denote the relative interior of a convex set C by $\operatorname{ri} C$, as usual.

Theorem 2.4 (solution existence and duality [7, Theorem 1])**.** *Suppose under Assumption 2.1 that*

$$\operatorname{ri} C_L \cap \operatorname{ri} C_l \neq \emptyset \;\; \textit{and} \;\; \operatorname{ri} C_{\widetilde{L}} \cap \operatorname{ri} C_{\widetilde{l}} \neq \emptyset. \tag{2.10}$$

Then solutions exist to both $(P_{\rm arcs})$ *and* $(D_{\rm arcs})$ *as problems articulated in* $\mathcal{A}_n^2$*, and*

$$\textit{minimum value in } (P_{\rm arcs}) \;=\; \textit{maximum value in } (D_{\rm arcs}). \tag{2.11}$$

This is supplemented by results from the Hamilton-Jacobi theory in [20] that exploit the role of the "max" terms in the growth conditions in Assumption 2.1, in understanding the possibilities for the sets C_L and $C_{\widetilde{L}}$.

Theorem 2.5 (special existence criterion [20, Corollary 4.4, Theorem 4.5])**.** *Under Assumption 2.1, the projection of the set* C_L *in either the* x_0 *component or the* x_1 *component is all of* $I\!R^n$*; likewise, the projection of the set* $C_{\widetilde{L}}$ *in either the* p_0 *component or the* p_1 *component is all of* $I\!R^n$*.*

Therefore, the constraint qualification conditions in (2.10) are sure to be satisfied if one condition out of each of the following is fulfilled:

(a) *Either there exists* x_0 *such that the function* $l(x_0, \cdot)$ *is finite, or there exists* x_1 *such that the function* $l(\cdot, x_1)$ *is finite,*

(b) *Either there exists* p_0 *such that the function* $\widetilde{l}(p_0, \cdot)$ *is finite, or there exists* p_1 *such that the function* $\widetilde{l}(\cdot, p_1)$ *is finite.*

The conditions in (2.10) of Theorem 2.4 and in Theorem 2.5(b) in terms of $\widetilde{L}$ and $\widetilde{l}$ can instead be expressed as conditions of another sort on L and l. We refer back to the cited papers behind these theorems for the details.

We are also able in our translation project to draw on the optimality conditions for problems $(P_{\rm arcs})$ and $(D_{\rm arcs})$ that were developed in [5] in terms of the subgradients of the convex functions that are involved. The *transversality* condition

$$(p(t_0), -p(t_1)) \in \partial l(x(t_0), x(t_1)), \tag{2.12}$$

which can be written dually as

$$(x(t_0), -x(t_1)) \in \partial \widetilde{l}(p(t_0), p(t_1)), \tag{2.13}$$

is important along with the *Euler-Lagrange* condition

$$(\dot{p}(t), p(t)) \in \partial L(x(t), \dot{x}(t)) \;\text{ a.e.}, \tag{2.14}$$

which can be written dually as

$$(\dot{x}(t), x(t)) \in \partial \widetilde{L}(p(t), \dot{p}(t)) \;\text{ a.e..} \tag{2.15}$$

And as an alternative to the Euler-Lagrange conditions there is the *Hamiltonian* condition

$$\dot{x}(t) \in \partial_p H(x(t), p(t)), \quad \dot{p}(t) \in \partial_x[-H](x(t), p(t)), \;\text{ a.e.} \tag{2.16}$$

where it should be recalled that the function $-H(\cdot, p)$ is convex, as is the function $H(x, \cdot)$.

Theorem 2.6 (characterizations of optimality [5, Theorems 5 and 6])**.** *The following are equivalent for a pair of arcs $x(\cdot)$ and $p(\cdot)$:*

(a) *$x(\cdot)$ solves ($P_{\rm arcs}$), $p(\cdot)$ solves ($D_{\rm arcs}$), and the optimal values are equal as in (2.11).*

(b) *The transversality condition (2.12) holds along with the Euler-Lagrange condition (2.14).*

(c) *The transversality condition (2.13) holds along with the Euler-Lagrange condition (2.15).*

(d) *Either transversality condition holds along with the Hamiltonian condition (2.16).*

With the assurance that the arcs in the primal and dual problems ($P_{\rm arcs}$) and ($D_{\rm arcs}$) can be restricted harmlessly to $\mathcal{A}^2_n$, and that these facts about existence and optimality will nonetheless persist, we can proceed next to pose the problems equivalently in the mold of ($P_{\rm link}$) and ($D_{\rm link}$). For this we take the Hilbert space to be

$$\mathcal{H} = \mathcal{L}^2_n \times \mathcal{L}^2_n \times I\!R^n \times I\!R^n \tag{2.17}$$

with the inner product and norm being given by

$$\begin{gathered}\Big\langle (z(\cdot), w(\cdot), c_0, c_1), (z'(\cdot), w'(\cdot), c'_0, c'_1)\Big\rangle = \textstyle\int_{t_0}^{t_1} z(t)\cdot z'(t)dt + \int_{t_0}^{t_1} w(t)\cdot w'(t)dt + c_0\cdot c'_0 + c_1\cdot c'_1,\\ ||(z(\cdot), w(\cdot), c_0, c_1)||^2 = \textstyle\int_{t_0}^{t_1} |z(t)|^2 dt + \int_{t_0}^{t_1} |w(t)|^2 dt + |c_0|^2 + |c_1|^2.\end{gathered} \tag{2.18}$$

We take the convex function Φ on $\mathcal{H}$ to be

$$\Phi(z(\cdot), w(\cdot), c_0, c_1) = \int_{t_0}^{t_1} L(z(t), w(t))dt + l(c_0, c_1) \tag{2.19}$$

and the linkage subspace $\mathcal{S}$ of $\mathcal{H}$ to be

$$\mathcal{S} = \{\, (x(\cdot), \dot{x}(\cdot), x(t_0), x(t_1)) \,|\, x(\cdot) \in \mathcal{A}^2_n\}. \tag{2.20}$$

Proposition 2.7 (linkage format verification)**.** *The convex function Φ is closed and proper, with its conjugate given by*

$$\Phi^*(z'(\cdot), w'(\cdot), c'_0, c'_1) = \int_{t_0}^{t_1} L^*(z'(t), w'(t))dt + l^*(c'_0, c'_1). \tag{2.21}$$

The subspace $\mathcal{S}$ is closed, with its complement given by

$$\mathcal{S}^\perp = \{\, (\dot{p}(\cdot), p(\cdot), p(t_0), -p(t_1)) \,|\, p(\cdot) \in \mathcal{A}^2_n\}. \tag{2.22}$$

Proof. The conjugate Φ^* is calculated by maximimizing $\langle (z(\cdot), w(\cdot), c_0, c_1), (z'(\cdot), w'(\cdot), c'_0, c'_1)\rangle - \Phi(z(\cdot), w(\cdot), c_0, c_1)$ over all $(z(\cdot), w(\cdot), c_0, c_1) \in \mathcal{H}$. That separates into a maximization over (c_0, c_1), yielding the conjugate expression $l^*(c'_0, c'_1)$, and the maximization producing the conjugate of the integral functional associated with the integrand L. That's known from [4] to be the integral functional functional associated with the conjugate integrand L^* under elementary conditions that are satisfied here, because L is autonomous (independent of time t), and the space $\mathcal{L}^2_n \times \mathcal{L}^2_n = \mathcal{L}^2_{2n}$ contains all the constant functions from $[t_0, t_1]$ to $I\!R^{2n}$. Specifically, from the properness of L and L^* there exist (ζ, ω) and (ζ', ω') in $I\!R^n \times I\!R^n$ such for $(z_0(\cdot), w_0(\cdot)) \equiv (\zeta, \omega)$ and $(z'_0(\cdot), w'_0(\cdot)) \equiv (\zeta', \omega')$ we have $\int_{t_0}^{t_1} L(z_0(t), w_0(t))dt < \infty$ and $\int_{t_0}^{t_1} L^*(z'_0(t), w'_0(t))dt < \infty$, and that fulfills the requirement.

In the same way, then, the conjugate of Φ^* as calculated from (2.21) comes out as Φ, since $L^{**} = L$ and $l^{**} = l$. Conjugate functions are by nature always closed, so this confirms that Φ is closed, like Φ^*. In other words, the lower level sets of these convex functions are closed — in the norm topology and the weak topology of $\mathcal{H}$, since for convex sets closedness is the same in those two topologies.

As for the claim about $\mathcal{S}^{\perp}$, that subspace certainly does include the elements indicated on the right side of (2.22), inasmuch as

$$\begin{aligned}&\Big\langle (x(\cdot),\dot{x}(\cdot),x(t_0),x(t_1)),(\dot{p}(\cdot),p(\cdot),p(t_0),-p(t_1))\Big\rangle\\ &\qquad=\int_{t_0}^{t_1}[x(t)\cdot\dot{p}(t)+\dot{x}(t)\cdot p(t)]dt+x(t_0)\cdot p(t_0)-x(t_1)\cdot p(t_1)=0.\end{aligned}$$

Suppose now, on the other hand, that $(z(\cdot),w(\cdot),c_0,c_1)\in\mathcal{S}^{\perp}$, or in other words, that

$$0=\int_{t_0}^{t_1}[x(t)\cdot z(t)+\dot{x}(t)\cdot w(t)]dt+x(t_0)\cdot c_0+x(t_1)\cdot c_1\ \text{ for all }\ x(\cdot)\in\mathcal{A}_n^2. \tag{2.23}$$

Define $p(\cdot)$ by $p(t)=-c_1-\int_t^{t_1}z(s)ds$, so that $p(\cdot)\in\mathcal{A}_n^2$ with $c_1=-p(t_1)$ and $z(\cdot)=\dot{p}(\cdot)$. Then

$$\int_{t_0}^{t_1}x(t)\cdot z(t)dt+x(t_1)\cdot c_1=\int_{t_0}^{t_1}x(t)\cdot\dot{p}(t)dt-x(t_1)\cdot p(t_1)=-\int_{t_0}^{t_1}\dot{x}(t)\cdot p(t)dt-x(t_0)\cdot p(t_0),$$

and in consequence from (2.23)

$$0=\int_{t_0}^{t_1}\dot{x}(t)\cdot[w(t)-p(t)]dt+x(t_0)\cdot[c_0-p(t_0)]\ \text{ for all }\ x(\cdot)\in\mathcal{A}_n^2.$$

But for $x(\cdot)\in\mathcal{A}_n^2$, $\dot{x}(\cdot)$ can be any function in $\mathcal{L}_n^2$, and $x(t_0)$ can be any element of $I\!R^n$. Hence, both $w(\cdot)-p(\cdot)=0$ and $c_0-p(t_0)=0$. This shows that $(z(\cdot),w(\cdot),c_0,c_1)$ must be of the form $(\dot{p}(\cdot),p(\cdot),p(t_0),-p(t_1))$ indicated in (2.22). The argument also establishes that the formula in (2.22) defines a *closed* subspace of $\mathcal{H}$, and that confirms that the formula for $\mathcal{S}$ itself, being of the same type, defines a closed subspace. □

Theorem 2.8 (control problems as linkage problems)**.** *With $\mathcal{H}$, Φ and $\mathcal{S}$ chosen as in (2.17), (2.19), (2.20), the control problems (P_{arcs}) and (D_{arcs}) are equivalently posed as linkage problems (P_{link}) and (D_{link}). Criteria for the existence of solutions to these specialized linkage problems are provided accordingly by Theorems 2.4 and 2.5. Moreover the joint optimality condition for them in (1.1) can be identified with the optimality conditions for the control problems in Theorem 2.6.*

Proof. It's clear from (2.19) and (2.20) that Φ reduces to the Bolza functional $\mathcal{J}_{L,l}$ on $\mathcal{S}$, and we now know also from Proposition 2.7, via (1.5), that Φ^* reduces to the Bolza functional $\mathcal{J}_{\widetilde{L},\widetilde{l}}$ on $\mathcal{S}^{\perp}$. The subgradients of Φ on $\mathcal{H}$ are given by

$$\begin{aligned}&(z'(\cdot),w'(\cdot),c_0',c_1')\in\partial\Phi(z(\cdot),w(\cdot),c_0,c_0)\quad\Longleftrightarrow\\ &\qquad(z'(t),w'(t))\in\partial L(z(t),w(t))\ \text{ a.e, and }\ (c_0',c_1')\in\partial l(c_0,c_1),\end{aligned} \tag{2.24}$$

where the part of the subgradient formula concerning the integral is known from [4]. But when $(z(\cdot),w(\cdot),c_0,c_1)$ belongs to $\mathcal{S}$ and $(z'(\cdot),w'(\cdot),c_0',c_1')$ belongs to $\mathcal{S}^{\perp}$, those subspaces being described by arcs $x(\cdot)$ and $p(\cdot)$ as in (2.20) and (2.22), the relations in (2.24) turn into the Euler-Lagrange condition and transversality condition in (b) of Theorem 2.6. □

3 Applying the progressive decoupling algorithm

In the framework furnished by Theorem 2.8, the progressive decoupling algorithm in (1.3) and (1.4) with parameter $r > 0$ can be applied to find solutions simultateously to $(P_{\rm arcs})$ and $(D_{\rm arcs})$. At first pass, that comes out as the following procedure.

Algorithm 3.1 (specialized PDA, preliminary version). *Iteration k begins with a pair of arcs $x^k(\cdot)$ and $p^k(\cdot)$ in $\mathcal{A}_n^2$. (For $k = 0$, they can be taken arbitrarily.) For the convex function Φ in (2.19) determine*

$$\begin{aligned}(z^k(\cdot), w^k(\cdot), c_0^k, c_1^k) = \;\; & \textit{unique minimizer over } (z(\cdot), w(\cdot), c_0, c_1) \in \mathcal{H} \textit{ of} \\ & \Phi(z(\cdot), w(\cdot), c_0, c_1) - \Big\langle (z(\cdot), w(\cdot), c_0, c_1), (\dot p^k(\cdot), p^k(\cdot), p^k(t_0), -p^k(t_1)) \Big\rangle \\ & \qquad + \frac{r}{2}\Big\| (z(\cdot), w(\cdot), c_0, c_1) - (x^k(\cdot), \dot x^k(\cdot), x^k(t_0), x^k(t_1)) \Big\|^2 .\end{aligned} \tag{3.1}$$

Then, with respect to the subspaces $\mathcal{S}$ in (2.20) and $\mathcal{S}^\perp$ in (2.22), update by

$$\begin{aligned}&(x^{k+1}(\cdot), \dot x^{k+1}, x^{k+1}(t_0), x^{k+1}(t_1)) = \mathrm{proj}_{\mathcal{S}}\,(z^k(\cdot), w^k(\cdot), c_0^k, c_1^k), \\ &(\dot p^{k+1}(\cdot), p^{k+1}(\cdot), p^{k+1}(t_0), -p^{k+1}(t_1)) = \\ &\qquad (\dot p^k(\cdot), p^k(\cdot), p^k(t_0), -p^k(t_1)) - r\, \mathrm{proj}_{\mathcal{S}^\perp}\,(z^k(\cdot), w^k(\cdot), c_0^k, c_1^k).\end{aligned} \tag{3.2}$$

The details in the steps of this algorithm must next examined to bring out their meaning and shed light on their practicality.

Theorem 3.2 (decoupling in time). *The minimization step (3.1) decomposes into calculating*

$$(c_0^k, c_1^k) = \underset{(c_0,c_1)}{\mathrm{argmin}} \Big\{ l(c_0, c_1) - c_0 \cdot p^k(t_0) + c_1 \cdot p^k(t_1) + \frac{r}{2}|c_0 - x^k(t_0)|^2 + \frac{r}{2}|c_1 - x^k(t_1)|^2 \Big\} \tag{3.3}$$

and for (almost every) $t \in [t_0, t_1]$

$$(z^k(t), w^k(t)) = \underset{(z,w)\in \boldsymbol{R}^n \times \boldsymbol{R}^n}{\mathrm{argmin}} \Big\{ L(z, w) - z \cdot p^k(t) - w \cdot \dot p^k(t) + \frac{r}{2}|z - x^k(t)|^2 + \frac{r}{2}|w - \dot x^k(t)|^2 \Big\}. \tag{3.4}$$

Proof. This separability is immediately seen from inserting the formulas for the inner product and norm in (2.18) and Φ in (2.19) into (3.1). The minimization of the integral functional reduces to a minimization for each instant in time through the decomposability rule in [4]; see also [19, 14.60]. □

The subproblems of minimization in Theorem 3.2 are *proximal steps* in the language of numerical optimization, because they can be rewritten in the pattern of finding the unique minimizer with respect to u of $f(u) + \frac{1}{2}|u - v|^2$ for some convex function f. The associated mapping prox_f from v to u is *nonexpansive* (Lipschitz continuous with norm ≤ 1). That might help in coping with t parameterically in (3.4) in a numerical scheme of approximation.

Understanding how to execute the projections in (3.2) is another challenge, but one that turns out to be have a closed-form answer obtainable from the elementary theory of ordinary differential equations in terms of hyperbolic sines and cosines,

$$\cosh t = \tfrac{1}{2}[e^t + e^{-t}], \quad \sinh t = \tfrac{1}{2}[e^t - e^{-t}], \tag{3.5}$$

which are the derivatives of each other. The $2n \times 2n$ matrices

$$J = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}, \qquad e^{tJ} = \begin{bmatrix} (\cosh t)I & (\sinh t)I \\ (\sinh t)I & (\cosh t)I \end{bmatrix}, \tag{3.6}$$

have a central role in this.

Theorem 3.3 (arc projections). *The arcs $x(\cdot)$ and $p(\cdot)$ that correspond to the projections of a given $(z(\cdot), w(\cdot), c_0, c_1) \in \mathcal{H}$ onto the spaces $\mathcal{S}$ and $\mathcal{S}^\perp$ in (2.20) and (2.22) can be calculated as follows. Determine $\xi : [t_0, t_1] \to I\!R^n$ and $\pi : [t_0, t_1] \to I\!R^n$ by*

$$\begin{bmatrix} \xi(t) \\ \pi(t) \end{bmatrix} = \int_{t_0}^{t} e^{(s-t_0)J} \begin{bmatrix} w(s) \\ z(s) \end{bmatrix} dt. \tag{3.7}$$

Then take

$$\begin{gathered} \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} = e^{-(t-t_0)J} \begin{bmatrix} x_0 + \xi(t) \\ p_0 + \pi(t) \end{bmatrix} \quad \textit{where, for } \tau = t_1 - t_0, \\ x_0 = \tfrac{1}{2}[c_0 + e^{-\tau} c_1 - \xi(t_1) + \pi(t_1)], \qquad p_0 = \tfrac{1}{2}[c_0 - e^{-\tau} c_1 + \xi(t_1) - \pi(t_1)]. \end{gathered} \tag{3.8}$$

Proof. Any element of $\mathcal{H}$ can be represented uniquely as a sum of elements of $\mathcal{S}$ and $\mathcal{S}^\perp$, and those elements are its projections on those subspaces. Thus, we are looking for the unique arcs $x(\cdot)$ and $p(\cdot)$ in $\mathcal{A}^2_n$ such that

$$(z(\cdot), w(\cdot), c_0, c_1) = (x(\cdot), \dot{x}(\cdot), x(t_0), x(t_1)) + (\dot{p}(\cdot), p(\cdot), p(t_0), -p(t_1)),$$

or in other words, such that

$$\begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = -J \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} + \begin{bmatrix} w(t) \\ z(t) \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} x(t_0) + p(t_0) \\ x(t_1) - p(t_1) \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}. \tag{3.9}$$

The solution to this linear differential equation is known to be given by

$$\begin{bmatrix} x(t) \\ p(t) \end{bmatrix} = e^{-(t-t_0)J} \left(\begin{bmatrix} x_0 \\ p_0 \end{bmatrix} + \int_{t_0}^{t} e^{(s-t_0)J} \begin{bmatrix} w(t) \\ z(t) \end{bmatrix} \right), \qquad \begin{bmatrix} x(t_0) \\ p(t_0) \end{bmatrix} = \begin{bmatrix} x_0 \\ p_0 \end{bmatrix}, \tag{3.10}$$

where x_0 and p_0 must be gleaned from (3.9) along with $x_1 = x(t_1)$ and $p_1 = p(t_1)$. In the notation (3.7), we have from (3.10) with $\tau = t_1 - t_0$ that

$$\begin{bmatrix} x(t_1) \\ p(t_1) \end{bmatrix} = e^{-\tau J} \left(\begin{bmatrix} x_0 \\ p_0 \end{bmatrix} + \begin{bmatrix} \xi(t_1) \\ \pi(t_1) \end{bmatrix} \right) = \begin{bmatrix} \cosh\tau & -\sinh\tau \\ -\sinh\tau & \cosh\tau \end{bmatrix} \begin{bmatrix} x_0 + \xi(t_1) \\ p_0 + \pi(t_1) \end{bmatrix}$$

and therefore

$$\begin{bmatrix} x_1 \\ p_1 \end{bmatrix} = \begin{bmatrix} \cosh\tau[x_0 + \xi(t_1)] - \sinh\tau[p_0 + \pi(t_1)] \\ -\sinh\tau[x_0 + \xi(t_1)] + \cosh\tau[p_0 + \pi(t_1)] \end{bmatrix}.$$

Because $\cosh\tau + \sinh\tau = e^\tau$, if follows that $e^{-\tau}[x_1 - p_1] = x_0 - p_0 + \xi(t_1) - \pi(t_1)$. From the second part of (3.9) we then have

$$x_0 + p_0 = c_0, \qquad x_0 - p_0 = e^{-\tau} c_1 - \xi(t_1) + \pi(t_1),$$

and this leads at once to the designation of x_0 and p_0 in (3.8). □

Theorem 3.3 reveals that the challenge in the projections comes down entirely to the challenge of integrating expressions of the form $\theta(t)e^{\pm t}$ over bounded t intervals and handling that parametrically.

These details about the implementing the steps in Algorithm 3.1 enable us to redescribe the procedure the following manner.

Algorithm 3.4 (specialized PDA, ultimate version). *Initiate with any arcs $x^0(\cdot)$ and $p^0(\cdot)$ in $\mathcal{A}^2_n$. In iteration k, having arcs $x^k(\cdot)$ and $p^k(\cdot)$ in $\mathcal{A}^2_n$,*

(a) *calculate (c^k_0, c^k_1) from (3.3) and $(z^k(t), w^k(t))$ for $t \in [t_0, t_1]$ from (3.4),*
(b) *determine $\xi^k(\cdot)$ and $\pi^k(\cdot)$ next from (3.7),*
(c) *get arcs $\widehat{x}^k(\cdot)$ and $\widehat{p}^k(\cdot)$ from (3.8),*
(c) *then update by taking $x^{k+1}(\cdot) = \widehat{x}^k(\cdot)$ and $p^{k+1}(\cdot) = p^k(\cdot) - r\widehat{p}^k(\cdot)$.*

Theorem 3.5 (convergence of the algorithm). *Suppose solutions to both $(P_{\rm arcs})$ and $(D_{\rm arcs})$ exist under the criteria in Theorem 2.4, or Theorem 2.5. Then Algorithm 3.4 will produce sequences of arcs $x^k(\cdot)$ and $p^k(\cdot)$ in $\mathcal{A}^2_n$ that converge to particular solutions $\bar{x}(\cdot)$ and $\bar{p}(\cdot)$ to these problems.*

The convergence is in the sense that the derivative functions $\dot{x}^k(\cdot)$ and $\dot{p}^k(\cdot)$ converge weakly in $\mathcal{L}^2_n$ to $\dot{\bar{x}}(\cdot)$ and $\dot{\bar{p}}(\cdot)$, while $x^k(t) \to \bar{x}(t)$ and $p^k(t) \to \bar{p}(t)$ for every $t \in [t_0, t_1]$.

Proof. The convergence is assured by the general theory of the progressive decoupling algorithm in [15], but that's weak convergence in the Hilbert space $\mathcal{H}$. According to the structure of $\mathcal{H}$ in (2.17), that would seem to mean just that $x^k(\cdot)$ and $\dot{x}^k(\cdot)$ converge weakly to $\bar{x}(\cdot)$ and $\dot{\bar{x}}(\cdot)$ in $\mathcal{L}^2_n$ while the endpoints $x^k(t_0)$ and $x^k(t_1)$ converge in $I\!R^n$ to $\bar{x}(t_0)$ and $\bar{x}(t_1)$, and likewise for the dual arcs. But for any $t \in (t_0, t_1)$,

$$x^k(t) = x^k(t_0) + \langle \dot{x}^k(\cdot), \chi_{[t_0,t]}(\cdot)\rangle \quad \text{for the characteristic function} \quad \chi_{[t_0,t]}(\cdot) \in \mathcal{L}^2_n,$$

so the weak convergence of $\dot{x}^k(\cdot)$ to $\dot{\bar{x}}(\cdot)$ along with the convergence of $x^k(t_0)$ to $\bar{x}(t_0)$ implies the convergence of $x^k(t)$ to $\bar{x}(t)$. $\square$

4 Refinements and Extensions

The dynamical decoupling procedure that has been presented has to be seen more as a template for designing computations than a numerical algorithm already implementable on problems in practice. That's true at a fundamental level because it deals with elements of infinite-dimensional spaces that generally can't be represented by finite arrays of numbers. But it's also true because the procedure asks for its steps to be carried out with precision. Instead there should be stopping critera which allow errors to some extent, keeping them adequately in check, and theory to guarantee that a solution will nevertheless be attained — in a limit. There should also be information about when the rate of such convergence might be linear with respect to some norm, and so forth.

Such a framework for inexact computation does already exist for the progressive decoupling algorithm itself, but so far only in finite-dimensional implementations. It has recently been worked out in [17, Sec. 4] along with results about how linear convergence is tied to whether a property of metric subregularity holds at the solution being approached by the generated sequence. That property is generic in a sense, and its modulus is revealed in [17] to dictate the rate of linear convergence as a function of the proximal parameter r in the algorithm. The main insight is that there's no ideal choice of r, but rather a trade-off: higher r makes the dual elements converge faster at the expense of the primal elements, while lower r has the opposite influence.

But again all that theory is finite-dimensional, although some aspects of the primal-dual trade-off with respect to the choice or r were earlier revealed in [15, Theorem 1]. Perhaps an infinite-dimensional extension is in the offing, but that has not, so far, been pursued.

Of course, problems in continuous time are often approximated effectively by problems in discrete time. There's no barrier to doing that with convex problems of Bolza, and such discretizations can even be seen for instance in [12] and [18]. The template here might serve, in adaptation, for finding solutions to such finite-dimensional problems, and then all features of the algorithm's theory would be available.

Many problems of such kind may anyway arise directly in discrete time because they model control by way of "management decisions," which naturally are consecutive in stages. There is not yet, however, any worked-out theory that relates the solutions to discretized versions of continuous-time convex problems of Bolza to those of the undiscretized problems. But variational analysis offers much that could be helpful along those lines.

A topic of interest in connection with these considerations is whether the decoupling procedure in its specialization laid out in Algorithm 3.4 can implemented successfully without having to resort to general arcs $x^k(\cdot)$ and $p^k(\cdot)$ in $\mathcal{A}^2_n$. The optimal arcs $\bar{x}(\cdot)$ and $\bar{p}(\cdot)$ in the limit will satisfy, according to Theorem 2.8, the Hamiltonian condition in (2.16) as a differential inclusion, so they will actually be in $\mathcal{A}^\infty_n$, not just $\mathcal{A}^2_n$. That's because the set-valued mapping from $I\!R^n \times I\!R^n$ to $I\!R^n \times I\!R^n$ in the differential inclusion is locally bounded, due to the Hamiltonian function being locally Lipschitz continuous. Could arcs in $\mathcal{A}^\infty_n$ instead of $\mathcal{A}^2$ be utilized from the start? The answer is yes. From (3.4), it can be seen that if $x^k(\cdot)$ and $p^k(\cdot)$ are in $\mathcal{A}^\infty_n$, then $z^k(\cdot)$ and $w^k(\cdot)$ will be in $\mathcal{L}^\infty_n$. In the projection operation in Theorem 3.3, which rests on solving the differential equation (3.9) in the case of $z^k(\cdot)$ and $w^k(\cdot)$, the arcs in the solution will again have their derivatives in $\mathcal{L}^\infty_n$, and when those projections enter the update, the resulting arcs $x^{k+1}(\cdot)$ and $p^{k+1}(\cdot)$ will be once more in $\mathcal{A}^\infty_n$. Thus,

$$\begin{array}{l}\text{if the algorithm is initiated with arcs } x^0(\cdot) \text{ and } p^0(\cdot) \text{ in } \mathcal{A}^\infty_n,\\ \text{then all subsequent arcs } x^k(\cdot) \text{ and } p^k(\cdot) \text{ will also be in } \mathcal{A}^\infty_n.\end{array} \tag{4.1}$$

Whether that would affect the description of convergence, however, is unclear. The key would be showing that the $\mathcal{L}^\infty_n$ norms of the derivatives stay bounded in the process. If so, because arcs in $\mathcal{A}^\infty_n$ are Lipschitz continuous, the pointwise convergence of $x^k(t)$ to $\bar{x}(t)$ and $p^k(t)$ to $\bar{p}(t)$ would turn into uniform convergence of these functions over the interval $[t_0, t_1]$.

Problems (P_{arcs}) and (D_{arcs}) could be extended to have time-dependent $L(t, x, y)$ and correspondingly a time-dependent Hamiltonian function $H(t, x, y)$. There is plenty of theory to support that already in the original papers [5] and [7]. We have kept from it to avoid the extra complication of the many considerations of measurability and measurable selections that are necessarily involved in that (although all are fully manageable), and also for the convenience of working with the growth conditions in Assumption 2.1, which come from a paper [20] where L is time-independent. That had the advantage on the side of providing us with Theorem 2.5 as well.

A notable limitation of all the Bolza problem theory so far discussed is hidden in the finiteness of the Hamiltonian H, which our growth conditions push even further. This finiteness is effectively equivalent to the exclusion of state constraints, either in the primal problem or the dual problem. Could an extension be made to allow them? It can, but a serious technical threshold has to be crossed. State constraints in the primal problem require admitting in the dual problem arcs $p(\cdot)$ that aren't even in $\mathcal{A}^1_n$, but are only of bounded variation; roughly this means replacing $\dot{p}(t)dt$ by a more general $I\!R^n$-valued signed measure on the interval $[t_0, t_1]$. Symmetry then suggests allowing the arcs $x(\cdot)$ in the primal problem likewise the freedom to be only of bounded variation; see [9], [10]. All surely interesting, but daunting for an extension of the Hilbert-space-based methodology of progressive decoupling.

What more might be done when control functions are made explicit, as they are in (1.7) through the structure of L in (1.6)? In the algorithm, as it stands, the consequence of this is that in each iteration the minimization in (3.4) is replaced by obtaining $(z^k(t), u^k(t))$ from the minimization with respect to z and u of

$$K(Cz,u) - z{\cdot}p^k(t) - w{\cdot}\dot{p}^k(t) + \frac{r}{2}|z - x^k(t)|^2 + \frac{r}{2}|Az + Bu - \dot{x}^k(t)|^2,$$

followed by setting

$$w^k(t) = Az^k(t) + Bu^k(t).$$

The procedure would go forward with $z^k(\cdot)$ and $w^k(\cdot)$ as before, but a sequence of control functions $u^k(\cdot)$ producing the arcs $x^k(\cdot)$ would be maintained on the side. Dual control functions could be generated along the way through the dual expression in (1.9).

Another approach to explicit controls might be to pass to a different linkage formulation with a different space $\mathcal{H}$ and subspace $\mathcal{S}$ that incorporates the A, B and C matrices and functions $u(\cdot)$ alongside of the others. That would help to guide the sequence of control functions $u^k(\cdot)$, posing it as weak L^2-type convergence, say, but there would be a drawback. Projections onto the different subspace $\mathcal{S}$ and its complement $\mathcal{S}^\perp$ would need to be executed, and they wouldn't be as simple as the projections in Theorem 3.3. The matrices J and e^{tJ} in (3.6) would be replaced by matrices that take A, B and C into account and are far more complex to handle.

To conclude this discussion, it's important to note that the decoupling of dynamics proposed here isn't necessarily limited to *convex* problems of optimal control. The theory of the progressive decoupling algorithm allows application to nonconvex objective functions Φ in the linkage problem through "elicitation" of convexity. This amounts to supposing that a newly formulated sufficient condition for local optimality, called *strong variational sufficiency*, holds with respect to the primal-dual solution pair to which the sequence generated by the basic algorithm in (1.3) and (1.4) converges, and choosing an "elicitation parameter" along side r that is high enough in relation to that parameter . Convergence of the sequence is guaranteed then, despite the nonconvexity in underlying the problem, as long as the algorithm is initiated near enough to local optimality and the minimization steps are executed locally in that respect, too.

The locally sufficient condition in question has been shown in [16] to exactly replicate classical sufficient conditions in nonlinear programming and at least some of its generalizations. But what counterpart might it have for nonconvex problems of Bolza? All that lies in unexplored territory, as of this writing. Where first-order conditions for optimality are concerned, though, there is already a solid platform; see [1], [2], [3], [14].

References

[1] P.D. Loewen and R.T. Rockafellar, "The adjoint arc in nonsmooth optimization," *Transactions American Math. Soc.* 325 (1991), 39–72.

[2] P.D. Loewen and R.T. Rockafellar, "New necessary conditions for the generalized problem of Bolza," *SIAM J. Control Opt.* 34 (1996), 1496–1511.

[3] P.D. Loewen and R.T. Rockafellar, "Bolza problems with general time constraints," *SIAM J. Control Opt.* 35 (1997), 2050–2069.

[4] R.T. Rockafellar, "Integrals which are convex functionals," *Pacific J. Math.* 24 (1968), 525–539.

[5] R.T. Rockafellar, "Conjugate convex functions in optimal control and the calculus of variations," *J. Math. Analysis Appl.* 32 (1970), 174–222.

[6] R.T. Rockafellar, "Generalized Hamiltonian equations for convex problems of Lagrange," *Pacific J. Math.* 33 (1970), 411–427.

[7] R.T. Rockafellar, "Existence and duality theorems for convex problems of Bolza," *Transactions American Math. Soc.* 159 (1971), 1–40.

[8] R.T. Rockafellar, *Conjugate Duality and Optimization*, No. 16 in Conference Board of Math. Sciences Series, SIAM Publications, 1974.

[9] R.T. Rockafellar, "Dual problems of Lagrange for arcs of bounded variation," in *Calculus of Variations and Control Theory*, D. L. Russell (ed.), Academic Press, 1976, 155–19

[10] R.T. Rockafellar, "Optimality conditions for convex control problems with nonnegative states and the possibility of jumps," in *Game Theory and Math. Economics*, O. Moeschlin (ed.), North-Holland, 1981, 339–349.

[11] R.T. Rockafellar, "Linear-quadratic programming and optimal control," *SIAM J. Control and Opt.* 25 (1987), 781–814.

[12] R.T. Rockafellar, "Multistage convex programming and discrete-time optimal control," *Control and Cybernetics* 17 (1988), 225–246.

[13] R.T. Rockafellar, "Hamiltonian trajectories and duality in the optimal control of linear systems with convex costs," *SIAM J. Control Opt.* 27 (1989), 1007–1025.

[14] R.T. Rockafellar, "Equivalent subgradient versions of Hamiltonian and Euler-Lagrange equations in variational analysis," *SIAM J. Control Opt.* 34 (1996), 1300–1315.

[15] R.T. Rockafellar, "Progressive decoupling of linkages in optimization and variational inequalities with elicitable convexity or monotonicity," *Set-Valued and Variational Analysis* 27 (2019), 863–893.

[16] R.T. Rockafellar, "Augmented Lagrangians and hidden convexity in sufficient conditions for local optimality," *Mathematical Programming* 198 (2023), 159-194.

[17] R.T. Rockafellar, "Generalizations of the proximal method of multipliers in convex optimization" *Computational Optimization and Applications*, accepted August 2023

[18] R.T. Rockafellar and R.J-B Wets, "Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete time," *SIAM J. Control Opt.* 28 (1990), 810–822.

[19] R.T. Rockafellar and R.J-B Wets, *Variational Analysis*, No. 317 in the series *Grundlehren der Mathematischen Wissenschaften*, Springer-Verlag, 1997.

[20] R.T. Rockafellar and P.R. Wolenski, "Convexity in Hamilton-Jacobi theory 1: dynamics and duality," *SIAM J. Control Opt.* 40 (2001), 1323–1350.

[21] R.T. Rockafellar and P.R. Wolenski, "Convexity in Hamilton-Jacobi theory 2: envelope representations," *SIAM J. Control Opt.* 40 (2001), 1351–1372.