Sparse discretization of sparse control problems

Evelyn Herberg1,*, Michael Hinze1,**, and Henrik Schumacher2,***

1 Mathematisches Institut, Universität Koblenz-Landau, Campus Koblenz, Universitätstraße 1, 56070 Koblenz, Germany.
2 Institut für Mathematik, RWTH Aachen University, Templergraben 55, 52062 Aachen, Germany.

We consider optimal control problems that inherit a sparsity structure, especially we look at problems governed by measure controls. Our goal is to achieve maximal sparsity on the discrete level. We use variational discretization of the control problems utilizing a Petrov-Galerkin approximation of the state, which induces controls that are composed of Dirac measures. In the parabolic case this allows us to achieve sparsity on the discrete level in space and time. Numerical experiments show the differences of this approach to a full discretization approach.

© 2019 The Authors Proceedings in Applied Mathematics & Mechanics published by Wiley-VCH Verlag GmbH & Co. KGaA Weinheim

1 Problem formulation

Inspired by [2] we consider the continuous minimization problem

\[
\min_{(u_0,u) \in M(\bar{\Omega}_c) \times M(\hat{Q}_c)} J(u_0, u) := \frac{1}{q} \| y - y_d \|_{L^q(Q)}^q + \alpha \| u \|_{M(\bar{Q}_c)} + \beta \| u_0 \|_{M(\bar{\Omega}_c)},
\]

where the state \( y \in L^q(Q) \) solves a heat equation with right hand side \( u \), initial data \( u_0 \) and zero boundary values in a very weak sense, i.e.

\[
\int_Q \left( \frac{\partial w}{\partial t} + \Delta w \right) y \, dx \, dt = \int_{\bar{\Omega}_c} w \, du + \int_{\bar{\Omega}_c} w(0) \, du_0,
\]

holds for all \( w \in W \), where \((u_0, u) \in M(\bar{\Omega}_c) \times M(\hat{Q}_c) \) and

\[
W = \left\{ w \in L^2(0,T;H^1(\Omega)) : \left( \frac{\partial w}{\partial t} + \Delta w \right) \in L^\infty(Q) \text{ and } w(x,T) = 0 \text{ in } \Omega \right\}.
\]

For more details we refer to [2, Definition 2.1]. Here, \( \Omega \) is an open bounded domain in \( \mathbb{R}^d \), \( d \in \{1,2,3\} \) with Lipschitz boundary \( \Gamma := \partial \Omega \), and \( \bar{\Omega}_c \subseteq \Omega \). For given \( T > 0 \) we consider the time interval \((0,T)\) and the space-time domain \( \hat{Q}_c := \Omega \times (0,T) \). Let \( \alpha > 0 \), \( \beta > 0 \) be suitable parameters and \( q \in (1,\min\{2,(d+2)/d\}) \). As control spaces we consider the real and regular Borel measures \( M(\bar{\Omega}_c) := C(\bar{\Omega}_c)^* \) and \( M(\hat{Q}_c) := C(\hat{Q}_c)^* \), respectively. In this setting, (P) has a unique solution \((\hat{u}_0, \hat{u})\) with associated state \( \hat{y} \), which can be proven similar to [2, Theorem 2.7]. For the controls we observe the following sparsity (see [2, Corollary 3.2.3]):

\[
\text{supp}(\hat{u}_0^+) \subset \{ x \in \bar{\Omega}_c : \hat{w}(x,0) = -\beta \}, \quad \text{supp}(\hat{u}_0^-) \subset \{ x \in \bar{\Omega}_c : \hat{w}(x,0) = +\beta \},
\]

\[
\text{supp}(\hat{u}^+) \subset \{ (x,t) \in \hat{Q}_c : \hat{w}(x,t) = -\alpha \}, \quad \text{supp}(\hat{u}^-) \subset \{ (x,t) \in \hat{Q}_c : \hat{w}(x,t) = +\alpha \},
\]

where \( \hat{w} \in L^2(0,T;H^1(\bar{\Omega})) \cap C(\bar{\Omega}) \) is the adjoint state (see [2, Theorem 3.1.]) and \( \hat{u}_0 = \hat{u}_0^+ - \hat{u}_0^- \) and \( \hat{u} = \hat{u}^+ - \hat{u}^- \) are the respective Jordan decompositions. Considering the generic case that \( \hat{w} \) is not constant on sets of measure greater than zero, this implies that the support sets of the controls are of measure zero. This motivates us to suggest a discretization strategy, which preserves this sparsity structure on the discrete level in space and time.

2 Variational discretization

We want to achieve the desired maximal discrete sparsity, i.e. Dirac-measures in space-time, by choosing the Petrov-Galerkin ansatz and -test spaces that will induce this structure in combination with the variational discretization concept introduced in [5]. This concept, via the discretization of the test space and the optimality conditions, induces an implicit discretization of the controls \((u_0, u) \in M(\bar{\Omega}_c) \times M(\hat{Q}_c) \). This is how we control the discrete structure of the controls.

In the following we will indicate discretization in space-time by index \( \sigma := (\tau,h) \). We define the discrete state space consisting of piecewise linear and continuous finite elements in space and piecewise constant functions with respect to time...
and the discrete test space consisting of continuous and piecewise linear functions in space and time. This setting yields a Crank-Nicholson scheme with smoothing step (see [3] and [4]). Since the space-time discrete adjoint state \( \tilde{w}_\sigma \) is piecewise linear in space and time, we deduce from the sparsity pattern in the generic case that the support of \( \tilde{u} \) is a subset of the space-time grid points and the support of \( \tilde{u}_0 \) is a subset of the space grid points. Consequently the induced discrete control spaces consists of dirac measures in space and space-time concentrated in grid points, respectively. Similar to the results in [1] we can prove that the projections of all continuous solutions \( (\tilde{u}_0, \tilde{u}) \) onto the induced discrete control spaces coincide with the unique discrete solution. Therefore it is meaningful to work with this discrete control space in numerical examples.

We will compare our approach to the discontinuous Galerkin discretization from [2, Chapter 4.1.], where the discrete state and test space are identical to the discrete state space in our approach and an implicit Euler time stepping scheme is used to constitute the discrete state equation.

### 3 Computational results

We numerically solve the variationally discretized and fully discretized problems, both by a semi smooth Newton’s method, using the respective optimality systems. To simplify, we fix \( u_0 = 0 \) and assume \( Q_t = Q \). We choose \( \Omega = (0, 1) \) and \( T = 1.5 \) and set \( q = \frac{1}{3} \). To visualize the occurring differences, we use a coarse equidistant \( 10 \times 15 \) space-time grid.

We generate a numerical example by calculating the associated state \( y(u) \) for \( u = \delta_{0,5,0.5} \) and then use the interpolation of \( y(u) \) as the desired state \( y_{\delta} \) in the problem (P). Consequently this example problem is a source identification that inherits sparsity. If the parameter \( \alpha \) equals zero, the only term remaining in the target functional is \( \frac{1}{2} \| y_{\sigma} - y_{\delta} \|_{L^2(Q)} \) and consequently the discrete controls \( u_{\sigma, VD} \) and \( u_{\sigma, DG} \) can be calculated by the inverse of the discrete solution operator of the heat equation. Here the indices ”VD” and ”DG” indicate the two discretization approaches - variational discretization and discontinuous Galerkin - respectively. Due to the different discretization approaches these differ, which can be observed in figure 1. As a consequence of the discretization error in \( y_{\delta} \), we are not able to reproduce \( u \) in either case.

![Fig. 1: From left to right: \( u = \delta_{0,5,0.5} \), \( y(u) \), \( y_{\delta}, u_{\sigma, VD}, u_{\sigma, DG} \)](image)

Now we choose \( \alpha = 0.65 \). Note that the chosen \( u \) is not the true solution for problem (P) with \( \alpha = 0.65 \), but rather the true source in the heat equation. Hence we do not expect to find \( u \) rather the location of \( u \), when we solve for the optimal controls \( \tilde{u}_{\sigma, VD} \) and \( \tilde{u}_{\sigma, DG} \) with associated states \( y_{\sigma, VD} \) and \( y_{\sigma, DG} \). See figure 2. Since \( u \) was located on the grid, we are able to find the exact location \((0.5, 0.5)\) in the variational discrete case. The pertubations in the state \( y_{\sigma, VD} \) are caused by the ratio of time and space gridizes in the Crank-Nicholson scheme. For the discontinuous Galerkin scheme, we observe that due to choice of the control space the discrete control \( u_{\sigma, DG} \) is piecewise constant on the time interval \([0.5, 0.6]\). As a consequence of the implicit Euler time stepping scheme the discrete state \( y_{\sigma, DG} \) is smoother.

![Fig. 2: From left to right: \( \tilde{u}_{\sigma, VD}, \tilde{y}_{\sigma, VD}, \tilde{u}_{\sigma, DG}, \tilde{y}_{\sigma, DG} \) (with \( \alpha = 0.65 \))](image)

### References

[1] E. Casas, C. Clason, and K. Kunisch, SIAM Journal on Control and Optimization 51(1), 28-63 (2013).
[2] E. Casas and K. Kunisch, ESAIM: Control, Optimisation and Calculus of Variations 22(2), 355-370 (2016).
[3] C. Goll, R. Rannacher, and W. Wollner (2015). Journal of Computational Finance 18(4), 1-37 (2015).
[4] N. von Daniels, M. Hinz, and M. Vierling, SIAM Journal on Control and Optimization 53(3), 1182-1198 (2015).
[5] M. Hinz, Computational Optimization and Applications 30(1), 45-61 (2005).