Mathematisches Forschungsinstitut Oberwolfach

Report No. 4/2017

DOI: 10.4171/OWR/2017/4

Mini-Workshop: Adaptive Methods for Control Problems Constrained by Time-Dependent PDEs

Organised by
Max Gunzburger, Tallahassee
Karl Kunisch, Graz
Angela Kunoth, Köln

8 January – 14 January 2017

Abstract. Optimization problems constrained by time-dependent PDEs (Partial Differential Equations) are challenging from a computational point of view: even in the simplest case, one needs to solve a system of PDEs coupled *globally* in time and space for the unknown solutions (the *state*, the *costate* and the *control* of the system). Typical and practically relevant examples are the control of nonlinear heat equations as they appear in laser hardening or the thermic control of flow problems (Boussinesq equations). Specifically for PDEs with a *long time horizon*, conventional time-stepping methods require an enormous storage of the respective other variables. In contrast, *adaptive methods* aim at distributing the available degrees of freedom in an a-posteriori-fashion to capture singularities and are, therefore, most promising.

Mathematics Subject Classification (2010): 65xx, 49xx.

Introduction by the Organisers

In January 2004, two of us organized an Oberwolfach–Miniworkshop "Numerical Methods for Instationary Control Problems", see Report 03/2004. Citing from that report, we wrote "The topic for the current Miniworkshop organized by Karl Kunisch (Graz), Angela Kunoth (Bonn) and Rolf Rannacher (Heidelberg) emerged from the Oberwolfach-Workshop "Numerical Techniques for Optimization Problems with PDE Constraints" which was held February 16-22, 2003. It was realized that numerically solving control problems which are constrained by time-dependent nonlinear PDEs are particularly challenging with respect to the complexity of the problem. Mathematically one has to minimize a functional under

PDE constraints and possibly additional constraints on the state and the control. Standard discretizations on uniform grids in space and time will only yield solutions where the inherent structures of the problem (nonlinearity, constraints) are not sufficiently captured. Certain optimization problems for large coupled systems of partial differential equations are currently not yet numerically treatable or do not satisfy the time constraints required in practice. Overcoming this barrier can only be achieved by designing new mathematically founded algorithmic approaches. The road towards this goal leads to many interesting problems in optimization, linear algebra, numerics, analysis, and approximation theory." In that report, we also wrote "Different modern approaches to overcome the complexity issues in numerical simulations for PDE-constrained optimization have been presented and discussed. One of the approaches is to employ fast iterative solvers like multigrid on uniform grids. The methodology which conceptually provides the largest potential is to introduce adaptivity. This drastically reduces complexity but depending on the context may require solving an additional problem. Wavelet approaches particularly allow to resolve each of the variables separately and in addition provide a built-in preconditioning."

By now, robust preconditioners for the fast solution of the resulting linear systems of stationary elliptic PDEs for the (at least) three variables state, costate and control exist in a variety of forms for discretizations on uniform grids. Twelve years later, however, we find that several issues have still not yet been systematically investigated and that some promising approaches have not been exploited to its full potential. Specifically, the issue to reduce the complexity of the problem by introducing adaptivity in space and time is not well understood for finite element or other standard discretizations. Most importantly, the issues of discretizations for systems of PDEs with

(1) hierarchical spaces and adaptivity in space and time;
(2) error estimation, convergence and complexity estimates on different grids for the different variables state, costate and control;
(3) exchange of information from different grids while maintaining accuracy;
(4) stability issues for problems with long time horizons;

have not been understood. Although by-now full weak space-time formulations of the constraining PDEs have become increasing popular, they still lack a mathematical foundation and rigorous proofs except in the case when wavelet schemes are employed. Even in the case of a single parabolic PDE, there does not yet exist an adaptive finite-element based scheme for which convergence and optimal complexity (when compared to a reference best N-term approximation) has been proved. We think that some aspects from the wavelet methodology may be helpful, like when proving optimal complexity of adaptive finite element methods for elliptic PDEs ten years ago which introduced new theoretical paradigms. While one can fully exploit with wavelets the functional analytic framework, they are still difficult to construct and to implement which, therefore, leads us to search for concepts which are easier to realize.

Before addressing the issue of adaptivity in time and space, one needs to be aware in which situations this concept applies. This raises the question in which function spaces (particular, Besov spaces) one can expect the different solution variables to be. *Regularity* of solutions of PDEs in Besov or other function spaces which allow for isolated singularities in the data or the domain have become a major topic for the solution of a single elliptic or parabolic PDE in recent years. Simulations for elliptic control problems based on wavelets reveiled that the state may be in a smooth (Sobolev) space while the adjoint state inherits non-regularity from the way the optimization problem is posed. In addition, in the case of PDE-constrained control problems with additional state inquality constraints, the adjoint variable is no longer a function but a Borel measure. However, recent results for establishing L_∞ error estimates for finite element schemes lead to insights that a norm from a space like BMO might be more appropriate.

We have seen in the past years an abundance of manuscripts on "a priori-" and "a posteriori error estimates" for PDE-constrained control problems including varieties of additional control and state constraints which mainly follow the principle of employing one error estimator and one grid for all variables. In addition, another large amount of publications stems from introducing uncertainty quantification into the area and therefore, another level of complexity, i.e., PDE-constrained control problems with stochastic coefficients, see, e.g., Report 04/2013 for the Oberwolfach Workshop "Numerical Methods for PDE Constrained Optimization with Uncertain Data" organized by Matthias Heinkenschloss (Houston) and Volker Schulz (Trier). We explicitly did not want to thematize uncertain data this time.

The time was perfect for organizing a Miniworkshop which focussed on the issue of *adaptive methods* for PDE-constrained control problems with a possibly long time horizon with some leading mathematicians in an atmosphere of a small workshop with not too many participants, to attack in a systematical way the underlying theoretical and resulting practical issues sketched above. The workshop was well attended with experts from various backgrounds in PDE-constrained control problems, regularity of solutions of PDEs, finite elements and wavelet methods.

We invited Martin Gander who spent this week at the institute within a "Research in Pairs" program to also give a talk during our workshop and include his extended abstract as well.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows".

Mini-Workshop: Adaptive Methods for Control Problems Constrained by Time-Dependent PDEs

Table of Contents

Abstracts

Multigrid methods based on high order tensor product B-Splines for the valuation of American options with stochastic volatility and their greeks

Sandra Boschert

For the efficient numerical solution of elliptic variational inequalities on closed convex sets, multigrid methods based on piecewise linear basis functions have been investigated over the past decades [6]. Essential to their success is the appropriate approximation of the constraint set on coarser grids which is based on function values of piecewise linear basis functions. On the other hand, there are a number of problems which profit from higher order approximations. Among these are problems of pricing American options, formulated as a parabolic free boundary value problem involving for example the Black-Scholes equation or the Heston-equation. The Heston equation is a parabolic partial differential equation depending on the underlying price and the volatility with a convection and diffusion term. We formulate the free boundary problem arising from the Heston model as a parabolic variational inequality and show the existence and uniqueness of the weak solution by using the result of [5].

In addition to computing the apriori unknown free boundary (the optimal exercise price of the option), accurate pointwise derivatives of the value of the stock option or volatility up to order two (the so-called Greek letters) are of particular importance. We propose a monotone multigrid method for discretisations in terms of tensor product B-splines of arbitrary order and coincidental nodes in the interior of the spatial domain and Crank-Nicolson in time to solve the parabolic asymmetric variational inequality on a closed convex set [2]. To have a better approximation of the solution, which is only continuous for some specific points, by using high order B-splines, we let the nodes coincide in those specific interior points. Former works are about the monotone multigrid method with B-splines of high order without considering coincidental nodes and the application to the one dimensional Black-Scholes equation [3], or discretising the free boundary problem (American option) or boundary value problem (European option) containing the two dimensional Heston-equation by finite difference- or finite element methods [1, 4, 7, 9]. For our apllication, we construct restriction operators and monotone coarse grid approximations for tensor product B-splines of order k with $k-1$ coinciding nodes in the interior [2]. Additionally, a suitable smoother for the asymmetric discretised variational inequality is required. Therefore, an iterative method the so called projected Jacobi overrelaxation method, is presented [8]. Finally, the monotone multigrid method is applied to the discretised variational inequality, which is derived from the free boundary problem arising from the Heston model [2]. In particular, it is shown that a discretisation of the asymmetric variational inequality based on tensor product B-splines of order four enables us

to compute the second derivative of the value of the stock option pointwise to high precision.

References

[1] T. Apel, G. Winkler and U. Wystup. Valuation of Options in Heston's Stochastic Volatility Model using Finite Element Methods. Foreign Exchange Risk, Risk Publications, 2001.

[2] S. Boschert. Multigrid methods based on high order tensor product B-Splines for the valuation of American options with stochastic volatility and their Greeks (working title). PhD Thesis, in preparation.

[3] A. Kunoth and M. Holtz. B-Spline-based monotone Multigrid Methods. Society for Industrial and Applied Mathematics, SIAM J. Numer. Anal. Vol. 45, No. 3, pp. 1175–1199, 2007.

[4] S. Ikonen and J. Toivanen. Operator splitting methods for american options with stochastic volatility. Numer. Math.,2009, pp. 299–324.

[5] K. Ito and K. Kunisch. Parabolic variational inequalities: The Lagrange multiplier approach. Math. Pures Appl., pp. 415–449, 2006.

[6] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I. Numer. Math. 69, pp. 167–184, 1994.

[7] A. Kunoth, C. Schneiders and K. Wiechers. Multiscale Methods for the Valuation of American Options with Stochastic Volatility. International Journal of Computer Mathematics, pp. 1145–1163, 2012.

[8] J.S. Pang. Necessary and suffcient conditions for the convergence of iterative methods for the linear complementarity problem. Journal of optimization theory and applications, pp. 1–17, 1984.

[9] C.W. Oosterlee: On multigrid for linear complementary problems with application to American-style options. Electronics Transactions on Numerical Analysis, Vol. 15, S. 165-185, 2003.

Feedback control of large-scale dynamical systems

Tobias Breiten

(joint work with Peter Benner, Karl Kunisch, Laurent Pfeiffer, Martin Stoll)

Consider a linear control system of the form

$$\begin{aligned} \dot{y}(t) &= Ay(t) + Bu(t), \quad y(0) = y_0 \in \mathbb{R}^n, \\ y_{\text{obs}}(t) &= Cy(t), \end{aligned}$$

where $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{\ell\times n}$. With trajectories of the system, associate the cost functional

$$\mathcal{J}_T(y,u) = \frac{1}{2}\left(y(T)^T M y(T) + \int_0^T \|y_{\text{obs}}(t)\|^2 + \|u(t)\|^2 \mathrm{d}t \right),$$

where $M = M^T \succeq 0$. It is well-known, see e.g. [3], that the optimal control u_{opt} can be given in feedback form $u_{\text{opt}} = -B^T \Pi y$ where Π is the solution to the algebraic Riccati equation ($T = \infty$) or the differential Riccati equation ($T < \infty$), respectively. While feedback (or closed-loop) controls are more robust than open-loop controls, an explicit computation of the matrix valued unknown Π quickly

becomes infeasible, in particular for systems resulting from a spatial discretization of a partial differential equation (PDE). For parabolic problems with finite-dimensional control space, the singular values of Π are however known to decay very rapidly. As a result, classical iterative Krylov solvers (e.g. BiCG, MinRes) can be adapted to a low rank format (see [1],[5]) such that the solution of the Lyapunov equations arising within the Kleinman-Newton iteration can be solved efficiently for large-scale problems. Similar low rank concepts/phenomena will be discussed for efficiently solving saddle point problems arising from a full space-time discretization of open-loop finite horizon control problems (see [6] for more details).

We further investigate stabilization problems for bilinear control systems of the form

$$\begin{aligned} \dot{y}(t) &= Ay(t) + Ny(t)u(t) + Bu(t), \quad y(0) = y_0 \in \mathbb{R}^n, \\ y_{\text{obs}}(t) &= Cy(t), \end{aligned}$$

where $A, N \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times 1}$ and $C \in \mathbb{R}^{\ell\times n}$. These systems for example arise for the Fokker-Planck equation where the control variable u can be interpreted as a so-called optical tweezer. For more details on the precise functional analytic control setup, we refer to [2]. In contrast to the linear case, the optimal feedback law for these systems is nonlinear and relies on the gradient ∇V of the optimal value function V. Since for the computation of V a nonlinear PDE of dimension n has to be solved, this approach is only feasible for $n < 10$. For this reason, we discuss approximation techniques based on a Taylor series expansion of ∇V. It is shown that this approach leads to highly structured linear systems of tensor product structure. We provide some ideas on how to solve these equations by applicable tools from tensor calculus, see e.g. [4].

References

[1] P. Benner and T. Breiten. Low rank methods for a class of generalized Lyapunov equations and related issues. Numerische Mathematik, 124, 2013, 441–470.

[2] T. Breiten, K. Kunisch, and L. Pfeiffer. Control Strategies for the Fokker-Planck equation. SFB-Report 2016-003, University of Graz, 2016.

[3] R. Curtain and H. Zwart. An Introduction to Infinite-Dimensional Linear Systems Theory. Springer-Verlag, 1995.

[4] L. Grasedyck. Existence and Computation of Low Kronecker-Rank Approximations for Large Linear Systems of Tensor Product Structure. Computing 72, 2004, 247–265.

[5] D. Kressner and C. Tobler. Krylov subspace methods for linear systems with tensor product structure. SIAM Journal on Matrix Analysis and Its Applications, 31(4), 2010, 1688-1714.

[6] M. Stoll and T. Breiten. A low-rank in time approach to PDE-constrained optimization. SIAM Journal on Scientific Computing, 37(1), 2015, B1–B29.

Quarklet frames in adaptive numerical schemes

Stephan Dahlke
(joint work with Philipp Keding and Thorsten Raasch)

We are concerned with the numerical treatment of elliptic operator equations of the form

$$\mathcal{L}u = f, \tag{1}$$

where $\mathcal{L}$ is a linear operator that maps a Hilbert space H boundedly invertible onto its normed dual space H'. In our applications, H will always be a Sobolev space on a bounded domain contained in $\mathbb{R}^d$. For realistic problems, the numerical treatment of (1) usually requires the use of adaptive algorithms in order to increase efficiency. In recent years, in particular adaptive methods based on a wavelet basis $\{\psi_\lambda\}_{\lambda\in\Lambda}$ for H have been quite successful. Indeed, it has been shown that for huge classes of problems these algorithms are guaranteed to converge with optimal order, i.e., they asymptotically realize the convergence order of best N-term wavelet approximation. We refer, e.g., to [1] for details. Essentially, adaptive wavelet algorithms are based on adaptive space refinements, so that they can be interpreted as h-methods. However, it is well–known that hp-methods sometimes perform much better compared to h-methods. Therefore, it is our long-term goal to design adaptive hp-methods based on wavelet expansions. Our work was inpired by earlier results of H. Triebel on quarkonial decompositions for function spaces [5]. These kinds of decompositions indeed allow for a polynomial enrichment. The construction is based on a partition of unity, and therefore the resulting frame elements do not possess vanishing moments. Since in adaptive wavelet algorithms the vanishing moment property is essential for the design of the basic building blocks such as the **APPLY** routine, our first goal was do derive quarkonial decompositions that allow for this particular feature. We start with symmetrized cardinal B-splines of order $m > \gamma + 1/2$, i.e., $\varphi = N_m(\cdot + \lfloor \frac{m}{2} \rfloor)$ with $\mathrm{supp}\varphi = [-\lfloor \frac{m}{2} \rfloor, \lceil \frac{m}{2} \rceil]$. By a simple polynomial enrichment, we obtain *quarks*

$$\varphi_p(x) := (\tfrac{x}{\lceil m/2 \rceil})^p \varphi(x), \quad \text{for all } p \geq 0, x \in \mathbb{R}. \tag{2}$$

As shown in [2], for a given $\tilde{m} \in \mathbb{N}$ such that $\tilde{m} \geq m$ and $m + \tilde{m}$ is even, there exists a compactly supported wavelet ψ with

$$\psi(x) = \sum_{k\in\mathbb{Z}} b_k \varphi(2x - k), \quad \text{for all } x \in \mathbb{R} \tag{3}$$

and $\tilde{m}$ vanishing moments, $\langle \psi, P\rangle = 0$ for $\deg P < \tilde{m}$. Moreover, the collection

$$\Psi_R := \big\{\varphi(\cdot - k), 2^{j/2}\psi(2^j \cdot -k) : j \in \mathbb{N}_0, k \in \mathbb{Z}\big\} \tag{4}$$

is a Riesz basis for $L_2(\mathbb{R})$.

In complete analogy to the wavelet ψ, let us consider the following *quarklets* ψ_p,

$$\psi_p(x) := \sum_{k\in\mathbb{Z}} b_k \varphi_p(2x - k), \quad \text{for all } p \in \mathbb{N}_0, x \in \mathbb{R}. \tag{5}$$

By assumption, $\psi_0 = \psi$ has $\tilde{m}$ vanishing moments. The following lemma shows that the other ψ_p have the same property.

Lemma 1. *For each $p \geq 0$, the quarklet ψ_p has $\tilde{m}$ vanishing moments.*

In order to obtain quarkonial decompositions for L_2 and associated Sobolev spaces, we consider the weighted system

$$\Psi_{Q,w} := \big\{ w_p \varphi_p(\cdot - k), w_p 2^{j/2} \psi_p(2^j \cdot - k) : p, j \in \mathbb{N}_0, k \in \mathbb{Z} \big\}. \tag{6}$$

Theorem 1. *Let $w_p \geq 0$ be chosen such that $w_p/(p+1)^{-1/2}$ is summable. Then $\Psi_{Q,w}$ is a frame for $L_2(\mathbb{R})$.*

By a suitable rescaling, we also get stable decompositions of Sobolev spaces.

Theorem 2. *For a given $\gamma > 0$, let $\varphi = N_m(\cdot + \lfloor \frac{m}{2} \rfloor), m > \gamma + 1/2$. Then, for the scaling factors $w_{p,j,s} := 2^{-js}(p+1)^{-2s-\delta}$, with $\delta > 1$, the system*

$$\Psi_{Q,\omega,s} = \{ \omega_{p,0,s} \varphi_p(\cdot - k), \omega_{p,j,s} 2^{j/2} \psi_p(2^j \cdot - k) \ : \ p, j \in \mathbb{N}_0, k \in \mathbb{Z} \}$$

has the frame property in $H^s, 0 < s < \gamma$.

As a first test problem, we studied the univariate Poisson equation with periodic boundary conditions. In this setting, the vanishing moment property yields the following decay estimate for the entries in the associated biinfinite stiffness matrix.

Proposition 1. *Let $m \geq 3$, $\varphi = N_m(\cdot + \lfloor m/2 \rfloor)$ and $\phi \in \{\varphi, \psi\}$. There exists $C = C(m, \psi)$, such that*

$$2^{-(j+j')} \big| \langle \phi'_{p,j,k}, \phi'_{p',j',k'} \rangle_{L_2(\mathbb{R})} \big| \leq C (p+1)^{m-1} (p'+1)^{m-1} 2^{-|j-j'|(m-3/2)}. \tag{7}$$

By means of this proposition, the follow compression result can be shown.

Theorem 3. *Let $m \geq 3$. For $J \in \mathbb{N}_0$, we define the biinfinite matrix A_J by dropping the entries $a_{\lambda,\lambda'}$ from A when*

$$a \, \log_2(1 + |p - p'|) + b \, |j' - j| > J, \tag{8}$$

with $a > 1$, $b \geq \frac{a}{a-1}$ and $-\tau + \frac{a(m-2)}{b} < -1$. Then the number of non-zero entries in each row and colum of A_J is of order 2^J, and

$$||A - A_J||_{\mathcal{L}(\ell_2(\Lambda))} \lesssim 2^{-J(m-2)/b}. \tag{9}$$

Therefore, a first important step towards the design of adaptive quarklet schemes has been taken. A suitable boundary adaptation to include Dirichlet boundary conditions can also be performed.

Further information can be found in [3, 4].

References

[1] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations – Convergence rates, Math. Comput. 70, 2001, 27–75.
[2] A. Cohen, I. Daubechies, and J.-C. Feauveau, Biorthogonal bases of compactly supported wavelets, Commun. Pure Appl. Math. 45, 1992, 485–560.
[3] S. Dahlke, P. Oswald, and T. Raasch, A note on quarkonial systems and multilevel partition of unity methods, Math. Nachr. 286, 2013, 600–613.
[4] S. Dahlke, P. Keding, and T. Raasch, Quarkonial frames with compression properties, Calcolo, 2016,DOI: 10.1007/s10092-016-0210-3.
[5] H. Triebel, The Structure of Functions, Birkhäuser, Basel/Boston/Berlin, 2001.

Five decades of time parallel time integration, and a note on the degradation of the performance of the parareal algorithm as a function of the Reynolds number

Martin J. Gander

Time parallel time integration has received renewed attention over the past decade, because, finally, in 2004, the visionary prediction of Nievergelt from 1964 became true [1]:

> "For the last 20 years, one has tried to speed up numerical computation mainly by providing ever faster computers. **Today, as it appears that one is getting closer to the maximal speed of electronic components**, emphasis is put on allowing operations to be performed in parallel. In the near future, much of numerical analysis will have to be recast in a more 'parallel' form."

There are nowadays so many processors on supercomputers that space parallelization often saturates before the available number of processors is employed, and thus speedup is not possible any more. This happens especially often for evolution problems, whose solution is usually computed by time-stepping, a process which is classically considered completely sequential.

Over the past five decades since Nievergelt, there have been many attempts to parallelize the classical time stepping approach, and the ideas employed can be classified into four different categories [2]:

Shooting Type Methods: These methods are based on a decomposition of the time interval into subintervals, and then a multiple shooting method is used, together with a Newton solver for the non-linear system of equations that needs to be solved for the shooting parameters. Nievergelt's method is a precursor of this, the complete algorithm was first proposed at the continuous level in [3]. Renewed interest was sparked by the presentation of the parareal algorithm [4], which is a multiple shooting method with an approximate Jacobian calculated on a coarse grid. A complete convergence analysis of parareal can be found in [5] for linear problems, and in [6] for non-linear problems. The parareal algorithm converges superlinearly, and works quite well for diffusive type problems, but is not effective for hyperbolic problems.

Waveform Relaxation: these methods go back to the existence proofs by Picard and Lindelöf for solutions of ordinary differential equations, and were adapted to become computational methods for circuits in [7]: the idea is to decompose the circuit into subcircuits, and then to solve iteratively only subcircuits, relaxing signals (waveforms) along wires that had to be cut in the decomposition to the previous iteration, which gave the method its name. Modern variants of waveform relaxation for PDEs use a domain decomposition of the spatial domain into subdomains, and then optimized transmission conditions, which leads to optimized Schwarz waveform relaxation methods [8]. These methods are well understood for parabolic problems [9, 10, 11], and also work for hyperbolic problems [12, 13]. Optimized transmission conditions are the key ingredient of many recently developed solvers for wave propagation problems, e.g. the sweeping preconditioner, source transfer and the method of polarized traces, for an overview, see [14, 15]. There are also waveform relaxation methods based on other domain decomposition methods, like Dirichlet-Neumann and Neumann-Neumann waveform relaxation methods [16, 18, 17].

Space-Time Multigrid Methods: The first space-time multigrid method was the parabolic multigrid method proposed by Hackbusch in [19]. The idea is to apply the smoother not only on one time level, but over several time levels, like advancing in time before having converged at the present time level. The parabolic multigrid method has multigrid performance, provided one is not coarsening in time. A multigrid waveform relaxation algorithm was proposed in [20], using waveform relaxation at the continuous level as a smoother. The first full space-time multigrid method was proposed and analyzed in [21], using special restriction and extension operators in time. A new space-time multigrid method for parabolic problems using optimized block Jacobi smoothers can be found in [22]. A complete convergence analysis exists for this method, and it has excellent weak and even strong scalability properties.

Direct time parallel methods: These methods solve the full space-time problem without iteration. The first such methods were special predictor corrector methods, where the prediction and correction step can be performed in parallel [23]. Such methods are ideal for small scale parallelism on multicore architectures, and the most modern variants are currently the RIDC methods [24]. These methods can solve for example a system of ordinary differential equations to 8th order accuracy at the cost of an Euler method using 8 cores; this factor can thus multiply the number of processors already used for the space parallelization. A very recent time parallel direct method is the ParaExp algorithm, which is based on a completely overlapping time domain decomposition and the rapid propagation of homogeneous problems using rational Krylov approximations [25]. This method is currently restricted to linear problems, but works very well also for hyperbolic problems.

Many more details about the historical development of time parallel time integration methods can be found in [2].

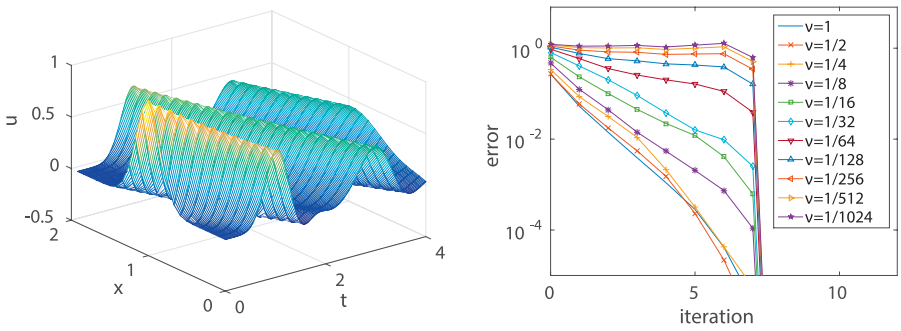


FIGURE 1. Solution for $\nu = 1/1024$ on the left, and degradation of the convergence of the parareal algorithm for decreasing ν (increasing Reynolds number) on the right.

After the presentation, Dörte Jando asked how the performance of the parareal algorithm degrades when one passes smoothly from the parabolic to a hyperbolic problem. I show here a numerical experiment to illustrate this, solving

$$u_t = \nu u_{xx} + u_x \quad \text{in } (0,2) \times (0,T), \quad T = 4, \tag{1}$$

with periodic boundary conditions[1] and the initial condition $u(0,x) = e^{-20(x-1)^2}$. I discretize the problem using centered finite differences with mesh size $h = 1/12$ and backward Euler in time with time step $\Delta t = 1/60$, which leads to the approximate solution shown on the left in Figure 1 for $\nu = 1/1024$, where ν represents the inverse of the Reynolds number. I then apply the parareal algorithm with 8 coarse time intervals, using one backward Euler step as the coarse solver. In Figure 1 on the right, I plot the decay of the maximum of the L^2 error in space over all coarse time points, as the parareal iteration progresses for a decreasing sequence of ν. One can clearly observe that when ν becomes small, the algorithm converges more and more slowly, and for very small ν, the only convergence mechanism left for the parareal algorithm is the convergence in a finite number of steps (here the 8th step), which however requires a number of iterations corresponding to the number of processors, and thus makes the method useless for parallelization, see also [5].

REFERENCES

[1] J. NIEVERGELT, *Parallel methods for integrating ordinary differential equations*, Comm. ACM, 7 (1964), pp. 731–733.

[2] M.J. GANDER, *50 Years of Time Parallel Time Integration*, in 'Multiple Shooting and Time Domain Decomposition', T. Carraro, M. Geiger, S. Körkel, R. Rannacher, editors, Springer Verlag (2015), pp. 69–114,

[3] P. CHARTIER AND B. PHILIPPE, *A parallel shooting technique for solving dissipative ODEs*, Computing, 51 (1993), pp. 209–236.

[1] otherwise the inflow boundary conditions would dominate the solution later in time, see [26]

[4] J.-L. Lions, Y. Maday, and G. Turinici, *A parareal in time discretization of PDEs*, C.R. Acad. Sci. Paris, Serie I, 332 (2001), pp. 661–668.
[5] M. J. Gander and S. Vandewalle, *Analysis of the parareal time-parallel time-integration method*, SIAM J. Sci.Comput., 29 (2007), pp. 556–578.
[6] M. J. Gander and E. Hairer, *Nonlinear convergence analysis for the parareal algorithm*, in Domain Decomposition Methods in Science and Engineering XVII, O. B. Widlund and D. E. Keyes, eds., vol. 60 of LNCS, Springer, 2008, pp. 45–56.
[7] E. Lelarasmee, A. E. Ruehli, and A. L. Sangiovanni-Vincentelli, *The waveform relaxation method for time-domain analysis of large scale integrated circuits*, IEEE Trans. on CAD of IC and Syst., 1 (1982), pp. 131–145.
[8] M. J. Gander, L. Halpern and F. Nataf, *Optimal Convergence for Overlapping and Non-Overlapping Schwarz Waveform Relaxation*, in Proceedings of the 11th International Conference on Domain Decomposition, ddm.org, (1999), pp. 27–36.
[9] M. J. Gander, and L. Halpern, *Optimized Schwarz Waveform Relaxation for Advection Reaction Diffusion Problems*, SIAM J. on Num. Anal., 45(2) (2007), pp. 666–697.
[10] D. Bennequin, M. J. Gander, and L. Halpern, *A homographic best approximation problem with application to optimized Schwarz waveform relaxation*, Math. of Comp., 78 (2009), pp. 185–223.
[11] D. Bennequin, M. J. Gander, L. Gouarin, and L. Halpern, *Optimized Schwarz Waveform Relaxation for Advection Reaction Diffusion Equations in Two Dimensions*, 134(3), (2016), pp. 513–567.
[12] M. J. Gander, L. Halpern, and F. Nataf, *Optimal Schwarz waveform relaxation for the one dimensional wave equation*, SIAM J. Num. Anal., 41 (2003), pp. 1643–1681.
[13] M. J. Gander and L. Halpern, *Absorbing boundary conditions for the wave equation and parallel computing*, Math. of Comp., 74 (2004), pp. 153–176.
[14] M.J. Gander, H. Zhang, *Iterative Solvers for the Helmholtz Equation: Factorizations, Sweeping Preconditioners, Source Transfer, Single Layer Potentials, Polarized Traces, and Optimized Schwarz Methods*, submitted to SIAM Review (2016).
[15] M.J. Gander, *On the limitations of sweeping type preconditioners for wave propagation*, Oberwolfach Report for the Mini-Workshop on Fast Solvers for Highly Oscillatory Problems (2016).
[16] F. Kwok, *Neumann-Neumann waveform relaxation for the time-dependent heat equation*, in Domain decomposition methods in science and engineering XVII, LNCS Springer (2014), pp. 189–198.
[17] M.J. Gander, F. Kwok and B. Mandal, *Dirichlet-Neumann and Neumann-Neumann waveform relaxation for the wave equation*, in Domain decomposition methods in science and engineering XVII, LNCS Springer (2014), pp. 501–509.
[18] M.J. Gander, F. Kwok and B. Mandal, *Dirichlet-Neumann and Neumann-Neumann Waveform Relaxation Algorithms for Parabolic Problems*, ETNA, 45 (2016), pp. 424–456.
[19] W. Hackbusch, *Parabolic multi-grid methods*, in Computing Methods in Applied Sciences and Engineering, VI, R. Glowinski and J.-L. Lions, eds., North-Holland, 1984, pp. 189–197.
[20] C. Lubich and A. Ostermann, *Multi-grid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216–234.
[21] S. Vandewalle and E. Van de Velde, *Space-time concurrent multigrid waveform relaxation*, Annals of Numer. Math, 1 (1994), pp. 347–363.
[22] M. J. Gander and M. Neumüller, *Analysis of a new space-time parallel multigrid algorithm for parabolic problems*, SIAM J. Sci. Comp., 38(4) (2016), pp. A2173–A2208.
[23] W. L. Miranker and W. Liniger, *Parallel methods for the numerical integration of ordinary differential equations*, Math. Comp., 91 (1967), pp. 303–320.
[24] A. J. Christlieb, C. B. Macdonald, and B. W. Ong, *Parallel high-order integrators*, SIAM Journal on Scientific Computing, 32 (2010), pp. 818–835.

[25] M. J. Gander and S. Güttel, *ParaExp: A parallel integrator for linear initial-value problems*, SIAM Journal on Scientific Computing, 35 (2013), pp. C123–C142.
[26] M. J. Gander, *Analysis of the Parareal Algorithm Applied to Hyperbolic Problems Using Characteristics*, Bol. Soc. Esp. Mat. Apl., 42 (2008), pp. 21–35.

Optimal multifidelity model management for UQ and control problems

Max Gunzburger
(joint work with Benjamin Peherstorfer and Karen Willcox)

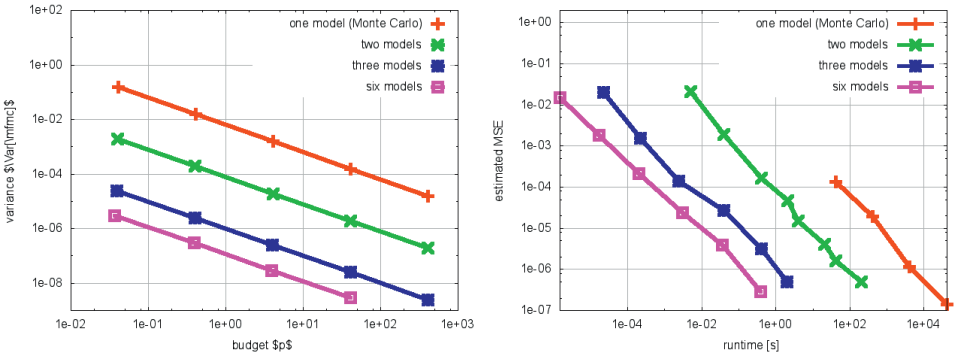
In many situations across computational science and engineering, multiple computational models are available that describe a system of interest. These different models have varying evaluation costs and varying fidelities. Typically, a computationally expensive high-fidelity model describes the system with the accuracy required by the application at hand whereas lower-fidelity models are less accurate but computationally cheaper than the high-fidelity model. Settings such as optimization and uncertainty quantification require multiple model evaluations at many different inputs which often leads to computational demands that exceed available resources if only the high-fidelity model is used.

The overall premise of the multifidelity methods we develop is that low-fidelity models are leveraged for speedup whereas the high-fidelity model is kept in the loop to establish accuracy and/or meet convergence guarantees. Existing acceleration methods typically exploit a multilevel hierarchy of surrogate models that follow a known rate of error decay and computational costs; however, a general collection of surrogate models, which may include projection-based reduced models, data-fitted models, support vector machines, and simplified physics models, does not necessarily give rise to such a hierarchy. Our multifidelity approach has the following features.

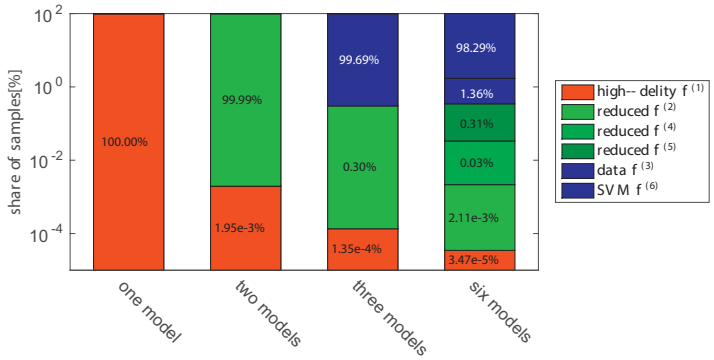
- It provides a framework to combine an arbitrary number of surrogate models of any type;
- Instead of relying on error and cost rates, an optimization problem balances the number of model evaluations across the high-fidelity and surrogate models with respect to error and costs.
- We show that a unique analytic solution of the model management optimization problem exists under mild conditions on the models
- Our multifidelity method makes occasional recourse to the high-fidelity model to estimate intra-model correlations and model costs;
- In doing so it provides an unbiased estimator of the statistics of the high-fidelity model, even in the absence of error bounds and error estimators for the surrogate models.

Numerical experiments with linear and nonlinear examples show that speedups by orders of magnitude are obtained compared to Monte Carlo estimation that only invokes a single model. As an illustration, consider the example of a locally damaged plate in bending. Four uniformly distributed parameters appear in the model: the load and the thickness and position and area of the damage. The

quantity of interest is the expected value of the maximum deflection of plate. The high-fidelity model of interest if a finite element model with 300 degrees of freedom. The cheaper and putatively less accurate surrogate models we use are 3 proper orthogonal decomposition models having 10, 5, and 2 degrees of freedom, a data-fit model in the form of a 256 point piecewise-linear interpolant of high-fidelity model solutions, and a support vector machine (SVN) with 256 points. Variance, correlation, and costs estimates are determined from 100 samples of each model. The first figure provides information about the variance and mean-squared error (MSE) for several model combinations. One sees that the largest improvements occur in going from one to two and two to three models; adding further reduced/SVM models only slightly reduces the MSE. Theoretical and computational MSE estimates match well.



The next figure provides the distribution of number of times each model is evaluated is automatically determined using our multi-fidelity Monte Carlo method. The number of samples changes exponentially between models with the highest number of samples for the data-fitted and SVM models; note that cost ratio between the high-fidelity and most used SVN model is approximately 10^6.



References

[1] B. Peherstorfer, K. Willcox, and M. Gunzburger, Optimal model management for multifidelity Monte Carlo estimation. SIAM J. Sci. Comp., 38, 2016, A3163-3194.

[2] B. Peherstorfer, K. Willcox, and M. Gunzburger, Survey of multifidelity methods in uncertainty propagation, inference, and optimization, Technical Report, MIT Aerospace Computational Design Laboratory TR-16-1, 2016.

On shape optimization with parabolic state equation

Helmut Harbrecht

(joint work with Johannes Tausch)

Shape optimization is a well established mathematical and computational tool in case of an elliptic state equation, see, e.g., [1] and the references therein. In contrast, in case of a parabolic state equation, the literature on shape optimization is quite limited.

Theoretical results on shape optimization with parabolic state equation can be found, for instance, in [3, 2, 7, 8, 9] and the references therein. Nonetheless, the development of efficient numerical methods for shape optimization problems with parabolic state equation is still in its beginning stages, especially for three-dimensional geometries.

We aimed at the development of such efficient methods in [4, 5, 6], where the focus was on shape identification problems for the heat equation. Namely, it is intended to determine inclusions or voids from measurements of the temperature and the heat flux at the outer boundary of the domain under consideration. The particular shape identification problem is reformulated as a shape optimization problem. Then, the Hadamard representation of the shape gradient is computed by means of the adjoint method. By identifying the sought boundary with its parametrization, a gradient based nonlinear Ritz-Galerkin scheme can be applied to discretize the shape optimization problem.

Since the adjoint equation is reversal in time, space-time discretization schemes are quite attractive for the determination of the states and their adjoints. To that end, we cast the parabolic boundary value problems into parabolic boundary integral equations. These boundary integral equations are discretized by using a Nyström discretization in space. For the time discretization, appropriate singularity corrected trapezoidal quadrature rules are applied to handle the singularities of the heat kernel and the solution. A space-time fast multipole method for the rapid evaluation of thermal potentials, developed in [10, 11], is employed to solve the discretized boundary integral equations.

Numerical experiments are carried out to demonstrate the feasibility and scope of the present approach. In particular, we are able to reconstruct unknown shapes in three dimensions on a laptop in less than half an hour computation time even though up to 1200 design parameters and about 120 000 boundary elements have been used for the discretization of the shape optimization problem.

References

[1] M. Delfour and J.-P. Zolesio. *Shapes and geometries. Metrics, Analysis, Differential Calculus, and Optimization*. 2nd edition, SIAM, Philadelphia, 2011.

[2] R. Dziri and J.-P. Zolesio. Shape derivative for the heat-conduction equation with Lipschitz continuous coefficients. *Boll. Un. Mat. Ital. B* **10**(3) (1996) 569–594.

[3] S. El Yacoubi and J. Sokolowski. Domain optimization problems for parabolic control systems. *Appl. Math. Comput. Sci.* **6** (1996) 277–289.

[4] H. Harbrecht and J. Tausch. An efficient numerical method for a shape identification problem arising from the heat equation. *Inverse Problems* **27**(6) (2011) 065013.

[5] H. Harbrecht and J. Tausch. On the numerical solution of a shape optimization problem for the heat equation. *SIAM J. Sci. Comput.* **35**(1) (2013) A104–A121.

[6] H. Harbrecht and J. Tausch. On shape optimization with parabolic state equation. In G. Leugering et al., editors, *Trends in PDE Constrained Optimization*, pages 213–229. Volume 165 of *International Series of Numerical Mathematics*, Birkhäuser, Basel, 2014.

[7] A. Henrot and J. Sokolowski. A shape optimization problem for the heat equation. In *Optimal control (Gainesville, FL, 1997)*, pages 204–223. Volume 15 of *Applied Optimization*, Kluwer Acad. Publ., Dordrecht, 1998.

[8] K.-H. Hoffmann and J. Sokolowski. Interface optimization problems for parabolic equations. Shape design and optimization. *Control Cybernet.* **23** (1994) 445–451.

[9] J. Sokolowski. Shape sensitivity analysis of boundary optimal control problems for parabolic systems. *SIAM J. Control Optim.* **26** (1988) 763–787.

[10] J. Tausch. A fast method for solving the heat equation by layer potentials. *J. Comp. Phys.* **224** (2007) 956–969.

[11] J. Tausch. Nyström discretization of parabolic boundary integral equations. *Appl. Numer. Math.* **59** (2009) 2843–2856.

Iterative solution of optimality systems in optimal control

Roland Herzog

(joint work with Kirk M. Soodhalter)

In this talk we address the iterative solution of self-adjoint saddle-point systems in Hilbert spaces, as well as the construction of preconditioners. Suppose that V and Q are two (real) Hilbert spaces and V^* and Q^* are their duals. Suppose further that $A \in \mathcal{L}(V, V^*)$, $B \in \mathcal{L}(V, Q^*)$ and $C \in \mathcal{L}(Q, Q^*)$ are bounded linear operators. We consider self-adjoint saddle-point systems, which are of the form

$$\begin{bmatrix} A & B^\star \\ B & -C \end{bmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \quad \text{with } A = A^\star,\ C = C^\star.$$

These systems arise as optimality conditions in optimal control problems, but also, for instance, in mixed finite element discretizations of numerous problems. In the context of partial differential equations, and especially time-dependent PDEs, these systems are usually of large scale after discretization, which calls for the use of iterative solvers.

In the first part of the talk, we discuss the role of block-diagonal preconditioners $P = \mathrm{blkdiag}(P_V, P_Q)$ composed of self-adjoint, positive definite blocks $P_V : V \to V^*$ and $P_Q : Q \to Q^*$. On the one hand, the obvious purpose of preconditioning is to render the spectrum of the preconditioned operator favorable so that iterative solvers, such as MINRES [4], exhibit sufficiently fast convergence. It is also well

known that the preconditioner takes exactly the role of the inner product in $V \times Q$; see for instance [1].

In this talk we emphasize yet another role of preconditioning. Apparently, the two residual components

$$\begin{aligned} r_1 &= f - A\,u - B^\star p \quad &\in V^* \\ r_2 &= g - B\,u + C\,p \quad &\in Q^* \end{aligned}$$

belong to different spaces and they also carry different meanings. In the context of optimal control, for instance, one of the residuals would be associated with feasibility w.r.t. the PDE contraint, while the other residual is related to the optimality. By contrast, for the stationary Stokes system one of the residuals measures the satisfaction of the balance of forces, while the other one is associated with the condition of mass conservation (or incompressibility).

Iterative solvers, and particularly MINRES, keep track of the *total* residual norm squared

$$\|r_1\|_{V^*}^2 + \|r_2\|_{Q^*}^2$$

during the iterations, where the norms in V^* and Q^* are induced by the preconditioner components P_V and P_Q; that is, $\|r_1\|_{V^*}^2 = \langle r_1, P_V^{-1} r_1 \rangle$ holds and similarly for the second term. From this it becomes apparent that the preconditioner must render the squared norms of the two residual subvectors physically compatible so that they can be added. This reasoning reveals some further conditions which can be used to find the proper dependence of the preconditioner building blocks on the problem's parameters.

Having said that the two contributions to the total residual norm squared originate from different equations with different physical meanings, it would be useful if an iterative solver, such as MINRES, would be able to keep track of $\|r_1\|_{V^*}^2$ and $\|r_2\|_{Q^*}^2$ *separately*. This issue is addressed in [2] and an extended implementation of MINRES is developed with very little computational overhead compared to traditional implementations. The Matlab code is available from [3]. Numerical examples are presented during the talk which reveal interesting convergence behavior of the two residual subvector norms. In particular, elliptic and parabolic optimal control problems are included.

References

[1] A. Günnel and R. Herzog and E. Sachs. A note on preconditioners and scalar products in Krylov subspace methods for self-adjoint problems in Hilbert space. Electronic Transactions on Numerical Analysis, 41, 2014, 13–20.

[2] R. Herzog and K. Soodhalter. A modified implementation of MINRES to monitor residual subvector norms for block systems. Preprint, 2016, TU Chemnitz.

[3] R. Herzog and K. Soodhalter. SUBMINRES. A modified implementation of MINRES to monitor residual subvector norms for block systems. DOI 10.5281/zenodo.47393, 2016.

[4] C. Paige and M. Saunders. Solution of sparse indefinite systems of linear equations. SIAM Journal on Numerical Analysis, 12, 1975, 617–629.

Reduced order modeling for time-dependent optimization problems with initial value controls & goal-oriented global error estimation for BDF methods

Dörte Jando

In this talk I treat optimal control problems governed by instationary partial differential equations which, after semi-discretization in space, result in large-scale time-dependent optimization problems. A powerful solution approach is provided for example by multiple shooting. I am motivated by the need to construct faster solvers for this kind of problems and the research presented in this talk should be understood as first steps in this direction.

In the first part, which is based on [1], I present a new reduced order model (ROM) approach to efficiently compute approximate Hessian information for the following large-scale linear quadratic optimal control problems where the optimal control is the initial value.

$$\text{(1a)}\quad \min_{\mathbf{s}\in\mathbb{R}^n} \int_0^T \frac{1}{2}\mathbf{y}(t)^T\mathbf{Q}\mathbf{y}(t) + \mathbf{c}(t)^T\mathbf{y}(t)dt + \frac{1}{2}\mathbf{y}(T)^T\mathbf{Q}_T\mathbf{y}(T) + \mathbf{c}_T^T\mathbf{y}(T) + \frac{1}{2}\mathbf{s}^T\mathbf{R}\mathbf{s},$$

where for given $\mathbf{s}$ the state $\mathbf{y} \in H^1(0,T;\mathbb{R}^n)$ solves

$$\text{(1b)}\quad \mathbf{M}\frac{d}{dt}\mathbf{y}(t) + \mathbf{A}\mathbf{y}(t) = \mathbf{f}(t), \qquad t \in (0,T)$$

$$\text{(1c)}\quad \mathbf{M}\mathbf{y}(0) = \mathbf{M}\mathbf{s}.$$

Problems of this type arise in source inversion problems, as subproblems in data assimilation or as subproblems in multiple shooting formulations.

A major difficulty in applying ROMs to the solution of (1) arises from the fact that the initial data $\mathbf{s}$ are variable. It is numerically expensive to identify a small subspace that contains the solution of (1). Therefore, it would be expensive to compute a small ROM that allows to replace the underlying large-scale differential equation (1b,c) by this ROM. To avoid this, we solve (1) using a Newton-type method for optimization where objective and gradient are computed using the original large-scale problem (1), but ROMs are used to approximate Hessian information.

Problem (1) can be solved using the conjugate gradient (CG) method on the optimality condition. Each iteration of the CG method requires the computation of the Hessian times a vector of initial data, which in turn requires the expensive sequential solution of two large-scale differential equations related to (1). To reduce the computational expense, we replace these two linear differential equations with projection based ROMs leading to approximate Hessians.

In general, existing ROM approaches in simulation and optimization address problems with a small number of inputs in the right hand side and fixed initial data. In the considered case, the situation is different: the difficulties are that the initial data vary and that this number of inputs is large.

The new approach, after having selected a basic ROM, augments this basic ROM by one vector. This vector is either the right hand side or the vector of

initial data to which the Hessian is applied to. The basic ROM $\mathbf{V}, \mathbf{W} \in \mathbb{R}^{n\times k}$ can be obtained e.g. by taking the first k eigenvectors of $(\mathbf{A}, \mathbf{M})$ corresponding to the smallest eigenvalues. The new ROM augmentation now augments the basic ROM by a vector $\mathbf{a}$ such that

$$\text{range}(\widehat{\mathbf{V}}) = \text{range}([\mathbf{V}, \mathbf{a}]), \quad \text{range}(\widehat{\mathbf{W}}) = \text{range}([\mathbf{W}, \mathbf{a}]), \quad \widehat{\mathbf{W}}^T\mathbf{M}\widehat{\mathbf{V}} = \mathbf{I}.$$

Although the augmentation is computationally inexpensive and the size of the ROM increases only by one, this new augmented ROM produces substantially better approximations than the basic ROM. We prove this for important special cases and observe a related behavior numerically for a more general advection diffusion reaction equation.

I also show how to use these ROMs in a CG method to solve (1). When the basic ROM is augmented by the vector of initial data to which the Hessian is applied to, then the resulting augmented ROM Hessian vector product is nonlinear. While in general, this can negatively impact the CG method, in important special cases only one augmentation is needed, and hence the linear CG method can be applied.

Overall, the new ROM Hessians provide substantially better approximations than the underlying basic ROM, and thus they can substantially reduce the computing time needed to solve these optimal control problems. Further information can be found in [1].

In the second part of the talk I sketch new goal-oriented global error estimators for variable multistep backward differentiation formulae (BDF) methods. BDF methods are particularly suitable to solve large-scale, nonlinear and stiff initial value problems. I first derive a Petrov-Galerkin finite element (FE) formulation of the BDF method and its discrete adjoint scheme obtained by reverse internal numerical differentiation (IND). Since Hilbert spaces are not sufficient for multistep methods the formulation is based on a Banach space setting using the duality pairing between continuous functions and normalized functions of bounded variation. In this setting differentiation and discretization commute [2].

The new global error estimators are based on the above formulation and make use of the dual weighted residual methodology. Defect integrals or local truncation errors are used as local error quantities and they are weighted by the discrete IND adjoints $\{\boldsymbol{\lambda}_n\}_{n=0}^N$

$$I(\mathbf{y}(T)) - I(\mathbf{y}_N) \approx E = \sum_{n=0}^{N-1} \boldsymbol{\lambda}_{n+1}^T(\text{local error quantity at } t_{n+1}).$$

Both estimators are asymptotically correct and converge optimally for one-step BDF method. For multistep BDF methods with selfstarting procedure the estimator with defect integrals converges suboptimal while the estimator with local truncation errors is again asymptotically correct and of optimal order. Moreover, good estimation quality in variable BDF-type methods is observed in practice [3].

References

[1] M. Heinkenschloss and D. Jando. Reduced Order Modeling for Time-Dependent Optimization Problems with Initial Value Controls. Submitted 2016.
[2] D. Beigel, M.S. Mommer, L. Wirsching and H.G. Bock. Approximation of weak adjoints by reverse automatic differentiation of BDF methods. Numerische Mathematik, 126(3), pp. 383–412, 2014.
[3] D. Jando. Efficient goal-oriented global error estimators for BDF methods using discrete adjoints. Journal of Computational and Applied Mathematics, 316, pp. 195–212, 2017.

Parallel multigrid on adaptive meshes

Guido Kanschat
(joint work with Conrad Clevenger, Timo Heister, Bärbel Holm, Martin Kronbichler)

When considering multigrid with local smoothing on adaptively refined meshes, we consider on each level a finite element space V_ℓ corresponding to the meshes in Figure 1. The mesh for the next coarser level is obtained by removing all of the

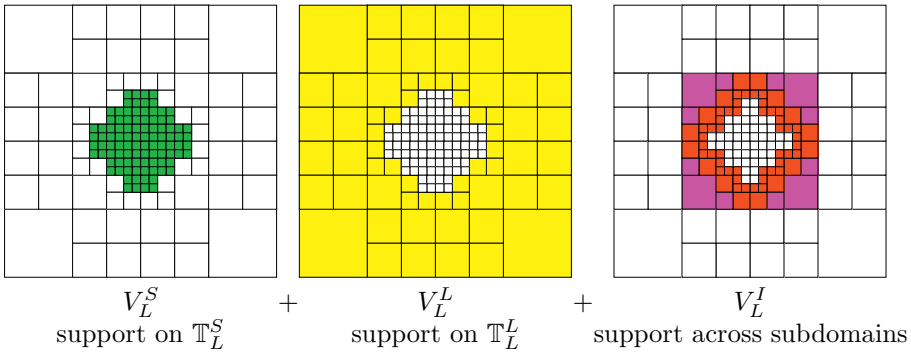


Figure 1. Subspace splitting for local smoothing

finest cells (in green on the left). Due to possibly very local refinement, standard smoothing on such a mesh does not lead to an algorithm with optimal complexity. Local smoothing, where we constrain smoothing in V_ℓ to those finest cells, touches each degree of freedom only once and thus restores optimality. Therefore, we are dealing on each level with a subspace structure as in Figure 1, namely the space V_ℓ^S of functions with support on cells strictly on level ℓ, that is, related to the conforming coarse mesh by ℓ levels of refinement, and the space V_ℓ^L of functions with support on cells of lower level. In the case of conforming elements with degrees of freedom on the interfaces between cells, the interface spaces V_ℓ^I must be added. Our goal is a method which performs as well as multigrid on uniformly refined meshes, but with local smoothing. It was shown in [1], that this is possible for discontinuous Galerkin methods by simply adding two matrices on each level

to the algorithm, corresponding to the outgoing and incoming numerical fluxes on the edge of refinement.

In [2], this was extended to conforming methods by the following rationale: in order achieve conformity, the trace (full, normal or tangential, depending on space of conformity) of the finite element solution in the refined region must be constrained to coincide with the trace of the solution in the coarse region. Thus, we have to eliminate additional degrees of freedom on the finer cells and end up with a trace of a function in $V_{\ell-1}$; therefore, smoothing on the coarser level can take care of these modes. Accordingly, we constrain local smoothing to the interior of the refined domain, or in terms of Figure 1, reducing the splitting to V_ℓ^S and $V_\ell^L \oplus V_\ell^I$. More precisely, the smoothing is done in the refined part of the domain with an essential boundary condition taken from the surrounding coarser level on the edge of refinement. Since matrix entries have to be removed in order to emulate essential boundary conditions, two additional matrices must be added in order to compute residuals across the refinement edge, similar to the discontinuous case.

In the second part, the `deal.II` implementation of the algorithms above on highly parallel systems is discussed as in [3]. Distribution of the leaf mesh, that is, the top cells of the hierarchy, uses the space filling curves of `p4est` (see [4]). The leaf mesh is then perfectly balanced. For distributing coarser meshes, we propose the simple heuristic of associating every parent cell to the same process as its first child, thus being able to determine the owning process of a cell without the necessity of communicating or storing additional data. We report strong scaling experiments up to around 10,000 cores on the SuperMUC machine in München exhibiting almost linear speedup down to about 25,000 degrees of freedom or 3,000 cells on the leaf mesh. While this is sufficient for most applications, the experiments also exhibit solution times per degree of freedom higher by a factor two to four for adaptive meshes compared to global refinement. A closer study attributes this loss of performance to imbalance of work between computer nodes on the finest and coarsest meshes, combined with global synchronization for grid transfers. In particular, there may be idle nodes on the finest and coarsest levels. Summarizing, a very simple heuristic for hierarchical load balancing based on uniform distribution of the leaf mesh yields very satisfactory parallelization results. We hope to close the performance gap between uniform and adaptive refinement by a restructuring of communication patterns in the multigrid method.

Finally, we pointed out the penalties paid for adaptively refined meshes. First, they produce irregular mesh structures for which more data has to be stored. Data is also accessed in a more irregular pattern, such that prefetch operations are less effective and more cache misses are encountered. Algorithms on irregular data structures must be implemented in a special way such that the vector units of modern hardware can be exploited. These issues are already encountered on single modern computers, and they will become more pronounced as processors become more integrated and offer more parallelism, and may cause a loss of several orders of magnitude compared to the peak performance of a machine. Even mathematically, local mesh refinement usually perturbs superconvergence. Thus, while only

adaptive algorithms are asymptotically optimal in the presence of singularities, they require a more and more involved implementation such that the break even point is not far below the desired accuracy for a computation.

References

[1] G. Kanschat. Multi-level methods for discontinuous Galerkin FEM on locally refined meshes. Computers & Structures, 82(28), 2004, pp. 2437–2445.
[2] B. Janssen and G. Kanschat. Adaptive multilevel methods with local smoothing for H^1- and H^{curl}-conforming high order finite element methods SIAM J. Sci. Comput., 33(4), 2011, pp. 2095–2114.
[3] T. C. Clevenger, T. J. Heister, G. Kanschat, and M. Kronbichler. A Flexible, Parallel, Adaptive Geometric Multigrid method for FEM. In preparation. 2017.
[4] C. Burstedde, L. C. Wilcox, and O. Ghattas. p4est: Scalable Algorithms for Parallel Adaptive Mesh Refinement on Forests of Octrees. SIAM J. on Sci. Comput. 33(3), 2011, pp. 1103–1133

Quasi-optimality of Galerkin approximations of parabolic problems

Christian Kreuzer
(joint work with Andreas Veeser and Francesca Tantardini)

A Galerkin approximation U of u is called quasi-optimal when there exists a constant C such that

$$\|u - U\| \leq C \inf_V \|u - V\|,$$

where V varies over the discrete trial space. Galerkin finite element schemes are very popular in the numerical solution of parabolic problems like the heat equation

$$\begin{aligned} \partial_t u - \Delta u &= f \quad \text{in } \Omega \times (0,T) \\ u &= 0 \quad \text{on } \partial\Omega \times (0,T) \\ u(\cdot,0) &= u_0 \quad \text{in } \Omega. \end{aligned}$$

In contrast to stationary problems, however, most of the error analysis for parabolic problems aims in proving optimal convergence rates without invoking quasi-optimality properties of the Galerkin approximation; for an overview compare e.g. with the monograph [9]. Quasi-optimality together with a proper interpolation operator implies optimal order convergence but not vice versa, i.e., quasi-optimality is stronger than optimal order estimates. Moreover, optimal order bounds typically depend on the exact solution, require stronger regularity assumptions and the bound does not necessarily vanish whenever the error does so.

This was recognised already in 1970 by Douglas and Dupont [5] who derived quasi-optimal estimates for a fixed semi-discretisation in space invoking norms with fractional time derivatives; compare also with [1, 8]. Related results concerning $L^2(H^1)$ norms or $H^1(H^{-1}) \cap L^2(H^1)$ norms can e.g. be found in [2, 6]. In [4] Dupont considered methods with dynamically changing spatial discretisations and discovered a 'counterexample' for which the Galerkin approximation

does not converge to the true solution when $h^4/\tau \to 0$, where h and τ denote the mesh respectively time-step sizes.

A common assumption of such results is that the L^2-projection onto the underlying discrete space is H^1-stable; see [2, 6]. Recently, Tantardini and Veeser [7] proved that this assumption is indeed necessary for fixed semi-discretisation in space. We extend this result in three steps.

First, we generalise [7] to semi-discretisations in space with dynamically changing spatial discretisations. To be more precise, we consider a time-partition $0 = t_0 < t_1 < \cdots < t_N = T$, $I_n = (t_{n-1}, t_n]$ and discrete spaces $\mathbb{V}_n$ such that the corresponding L^2-projections Π_n are uniformly H^1-stable. The Galerkin solution U is then given by

$$\begin{aligned} \partial_t U - \Delta_n U &= \Pi_n f, \quad \text{in } \Omega \times (t_{n-1}, t_n] \\ U(t_{n-1}^+) &= \Pi_n U(t_{n-1}). \end{aligned}$$

Here $\Delta_n : \mathbb{V}_n \to \mathbb{V}_n$ denotes the discrete Laplacian and $U(t_{n-1}^+) = \lim_{t \searrow t_{n-1}} U(t)$. It actually turns out that the method is inf-sup stable and therefore quasi-optimal among all trial functions V satisfying $V(t_{n-1}^+) = \Pi_n V(t_{n-1})$, $n = 1, \dots, N$. We replace the best-approximation by a suitable interpolation in order to avoid the coupling across time-steps and to obtain a priori estimates consisting of local approximation errors. However, the decoupling introduces an additional term, which can be bounded by

$$\sum_{n=0}^{N-1} \|\Pi_n^+ (\mathrm{id} - \Pi_n) u(t_n)\|_{L^2(\Omega)}^2, \tag{1}$$

where Π_n^+ denotes the L^2-projection onto $\mathbb{V}_n \oplus \mathbb{V}_{n+1}$. Note that a similar term appears also in the $L^\infty(L^2) \cap L^2(H^1)$ estimates of Chrysafinos and Walkington [3]. We emphasise that (1) vanishes when $\mathbb{V}_{n+1} \subset \mathbb{V}_n$ and that it is consistent with Dupont's counterexample since in this case it behaves like h^4/τ.

Second, we consider the backward Euler dG semi-discretisation in time with variable time-step sizes, i.e., $U_0 = u_0$ and $U_n \in H^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$, $n = 1, \dots, N$, with

$$U_n - U_{n-1} - \tau_n \Delta U_n = \int_{I_n} f \, \mathrm{d}t \quad \text{in } H^{-1}(\Omega).$$

Since the method is non-conforming, we obtain quasi-optimality up to a consistency error, which can be bounded again by a best-error. In order to prove a priori estimates, we involve stability of a Scott-Zhang type interpolation in time, which requires that $\frac{\tau_{n-1}}{\tau_n}$ is uniformly bounded. In the case $u \in H^1(H^1) \hookrightarrow C(H^1)$, this restriction can be avoided by using a Lagrange type interpolation.

Finally, we combine the two previous steps in order to obtain quasi-optimal bounds for the fully space-time discrete backward Euler dG method. The a priori estimates involve then both, the uniform boundedness of $\frac{\tau_{n-1}}{\tau_n}$ resulting from the

time discretisation as well as the term

$$\sum_{n=0}^{N-1} \|\Pi_n^+(\mathrm{id}-\Pi_n)u(t_n)\|^2_{L^2(\Omega)}$$

accounting for dynamically changing spaces.

The above results generalise in several directions:

- The Laplace operator can be replaced by a time-dependent possibly non-symmetric elliptic operator $A : (0,T) \to \mathcal{L}(H^1, H^{-1})$ which is measurable in $(0,T)$ and uniformly bounded and coercive.
- The techniques can be extended to higher order dG schemes in time.
- Similar ideas apply to the ultra weak solution of the heat equation, where $u \in L^2(H^1)$ is sought.

References

[1] C. Baiocchi and F. Brezzi. Optimal error estimates for linear parabolic problems under minimal regularity assumptions. Calcolo, 20(2), 1983, 143–176.

[2] K. Chrysafinos and L. S. Hou. Error estimates for semidiscrete finite element approximations of linear and semilinear parabolic equations under minimal regularity assumptions. SIAM J. Numer. Anal., 40(1), 2002, 282–306.

[3] K. Chrysafinos and N. J. Walkington. Error estimates for the discontinuous Galerkin methods for parabolic equations. SIAM J. Numer. Anal., 44, 2006, 349–366.

[4] T. Dupont. Mesh modification for evolution equations. Math. Comp., 39(159), 1982, 85–107.

[5] J. Douglas and T. Dupont. Galerkin methods for parabolic equations. SIAM J. Numer. Anal., 7, 1970, 575–626.

[6] W. Hackbusch Optimal $H^{p,p/2}$ error estimates for a parabolic Galerkin method. SIAM J. Numer. Anal., 18(4), 1981, 681–692.

[7] F. Tantardini and A. Veeser. The L^2-projection and quasi-optimality of Galerkin methods for parabolic equations. SIAM J. Numer. Anal., 54, 2016, pp. 317–340.

[8] F. Tomarelli. Regularity theorems and optimal error estimates for linear parabolic Cauchy problems. Numer. Math., 45(1), 1984, 23–50.

[9] V. Thomée. Galerkin finite element methods for parabolic problems. vol. 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 2006.

Adaptive wavelet methods for parabolic PDE-constrained control problems

Angela Kunoth

(joint work with Max Gunzburger)

Optimization problems constrained by PDEs are challenging from a computational point of view: one needs to solve a system of PDEs coupled globally in space and time if the underlying PDE is time-dependent. This global coupling is an unavoidable feature of such control problems as specified next: typically an adjoint PDE comes into place.

PDE-constrained control problems. Let Y, U be Hilbert spaces over $\mathbb{R}$ which shall host the state y of a system and a control by which the state can be influenced. Let $J : Y \times U \to \mathbb{R}$ be a twice differentiable functional, and

let $K : Y \times U \to Q'$ be a (in y, u Fréchet-) differentiable function where Q' denotes the topological dual of another Hilbert space Q. Consider the constrained minimization problem

$$\inf_{(y,u)\in Y\times U} J(y,u) \quad \text{subject to } K(y,u) = 0. \tag{1}$$

For the constraints $K(y,u) = 0$ (the PDE), we assume that there exists a unique solution $y \in Y$ when $u \in U$ is given. A typical way to solve (1) is to compute the zeroes of the first order Fréchet derivatives of the Lagrangian functional which is defined by introducing the co-(or adjoint) state p by which the constraints are appended to the functional J, i.e., $L(y,u,p) := J(y,u) + \langle K(y,u), p\rangle_{Q'\times Q}$ with $L : Y \times U \times Q \to \mathbb{R}$. Denoting by $L_z(y,u,p) := \frac{\partial}{\partial z}L(y,u,p)$ and $L_{zz}(y,u,p) := \frac{\partial^2}{\partial z^2}L(y,u,p)$ the first and second variation, of L with respect to $z = y, u, p$, and assuming that J is quadratic in both y, u and that K is linear in y, u, the necessary conditions for optimality yield the linear system of equations

$$\begin{pmatrix} L_{yy} & L_{yu} & K_y^* \\ L_{uy} & L_{uu} & K_u^* \\ K_y & K_u & 0 \end{pmatrix} \begin{pmatrix} y \\ u \\ p \end{pmatrix} = g \iff: \begin{pmatrix} \mathcal{A} & \mathcal{B}^* \\ \mathcal{B} & 0 \end{pmatrix} \begin{pmatrix} (y,u)^\top \\ p \end{pmatrix} = g \iff: Gq = g \tag{2}$$

with some right hand side g and C^* the dual of C. The Hessian of L or the Karush-Kuhn-Tucker (KKT) operator G has for such linear-quadratic problems constant entries, and the necessary conditions are also sufficient. Moreover, if J or K do not contain products yu, one has $L_{yu} = L_{uy} = 0$ so that $\mathcal{A}$ is a block diagonal operator. Typically, the quadratic functional (1) contains inner products so that the resulting Riesz operators L_{yy}, L_{uu} are symmetric which implies that $\mathcal{A}$ and, thus, G is symmetric. In all the cases we consider, the operators $\mathcal{A} : \mathcal{V} \to \mathcal{V}$, $\mathcal{B} : \mathcal{V} \to Q'$ for some Hilbert space $\mathcal{V}$ are continuous; $\operatorname{Im}\mathcal{B} = Q'$; and $\mathcal{A}$ is invertible on $\operatorname{Ker}\mathcal{B}$ so that the saddle point problem (2) has for $g \in \mathcal{V}' \times Q'$ a unique solution $q \in \mathcal{V} \times Q$ by the Brezzi-Fortin theory. Thus, we can consider constrained linear-quadratic minimization problems (1) as symmetric saddle point problems (2) with a boundedly invertible linear mapping $G : \mathcal{V} \times Q \to \mathcal{V} \times Q'$ where $\mathcal{V} := Y \times U$. Two standard examples to which this scenario applies are the following.

Dirichlet problem with distributed control. Consider the standard weak formulation of a second order elliptic PDE with homogeneous Dirichlet boundary conditions. Choosing $Y := H_0^1(\Omega)$ and $U := Y'$, we consider for given $f \in Y'$ the linear operator equation

$$K(y,u) := Ay - f - u = 0 \tag{3}$$

and the quadratic objective functional

$$J(y,u) := \frac{1}{2}\|y - y_*\|_Y^2 + \frac{\omega}{2}\|u\|_{Y'}^2 \tag{4}$$

for a given target state $y_* \in Y$ and any fixed weight parameter $\omega > 0$. We assume that $A : Y \to Y'$ is a linear (not necessarily symmetric) boundedly invertible

operator. The norms in (4) can more generally be norms on Hilbert spaces as long as the constrained optimization system (1) possesses a unique solution.

Denote by $R : Y \to Y'$ the Riesz operator defined by the inner product $(\cdot,\cdot)_Y$ inducing $\|\cdot\|_Y$, $\langle v, Rw\rangle_{Y\times Y'} := (v,w)_Y$, $v, w \in Y$. Since $(\cdot,\cdot)_Y$ is symmetric, R is also. Consequently, the Lagrangian is

$$L(y,u,p) = \frac{1}{2}\langle y-y_*, R(y-y_*)\rangle_{Y\times Y'} + \frac{\omega}{2}\langle u, R^{-1}u\rangle_{Y\times Y'} + \langle Ay-f-u, p\rangle_{Y'\times Y} \tag{5}$$

implying that the system (2) becomes

$$G\begin{pmatrix} y\\ u\\ p\end{pmatrix} := \begin{pmatrix} R & 0 & A^*\\ 0 & \omega R^{-1} & -I\\ A & -I & 0\end{pmatrix}\begin{pmatrix} y\\ u\\ p\end{pmatrix} = \begin{pmatrix} Ry_*\\ 0\\ f\end{pmatrix}, \tag{6}$$

i.e., $\mathcal{A} = \mathrm{diag}(R, \omega R^{-1})$ and $\mathcal{B} = (A, -I)$. The system matrix G defined in (6) is symmetric since R is and is boundedly invertible by the properties of $\mathcal{A}$, $\mathcal{B}$.

Parabolic PDE with distributed control. The constraint $K(y,u) = 0$ in (1) is here a linear parabolic evolution PDE in a variation of the full space-time weak formulation from [5]. The parabolic operator equation is formulated such that the resulting operator B is boundedly invertible from $\mathcal{X} := L^2(I)\otimes Y$ to $\mathcal{Y}' := ((L^2(I)\otimes Y)\cap(H^1_T(I)\otimes Y'))'$ where $H^1_T(I)$ is the closure of the functions in $H^1(I)$ which vanish at end time T and $I := (0,T)$ denotes the time interval. The constraints are of the form (3) with the parabolic evolution operator $B = \partial_t + A$ in full weak space-time form in place of A, see [4] for details. Choosing the objective function then as in (4) with the obvious changes for the norms, i.e., using the norms for $\mathcal{X}$, $\mathcal{Y}$, we arrive at a system very similar to (6) with symmetric $\mathcal{A} = \mathrm{diag}(R_1, \omega R_2)$ with the respectively defined Riesz operators. Finally, the resulting operator G is a boundedly invertible mapping from $\mathcal{Z} := \mathcal{X}\times\mathcal{Y}'\times\mathcal{X}$ onto $\mathcal{Z}'$.

Adaptive wavelet methods for the parabolic PDE-constrained control problem. In view of the fully in space and time coupled system (2), conventional time-stepping methods require an enormous storage. In contrast, adaptive methods in both space and time which aim at distributing the available degrees of freedom in an a-posteriori-fashion to capture singularities are most promising.

Employing adaptive wavelet schemes for full weak space-time formulations of the parabolic PDEs, we can prove convergence and optimal complexity for control problems constrained by a linear parabolic PDE [4], generalizing the ideas from [3] for control problems constrained by an elliptic PDE.

Our method of attack is based on the solution paradigm developed for a single linear elliptic PDE in [1, 2]. Firstly, the constraints are represented by means of a full weak space–time formulation as a linear system in ℓ_2 in wavelet coordinates, following the approach in [5]. Secondly, a quadratic cost functional involving a tracking–type term for the state and a regularization term for the distributed control is also formulated in terms of ℓ_2 sequence norms of wavelet coordinates. This functional serves as a representer for a functional involving different Sobolev norms with possibly non-integral smoothness parameter. Standard techniques

from optimization are then used to derive the first order necessary conditions as a coupled system in ℓ_2–coordinates.

Our proposed adaptive method can be interpreted as an inexact gradient method for the control. In each iteration step, the primal and the adjoint system are solved up to a prescribed accuracy by the adaptive algorithm. It is shown that the adaptive algorithm converges. Moreover, the algorithm is proved to be asymptotically optimal: the convergence rate achieved for computing each of the components of the solution (state, adjoint state and control) up to a desired target tolerance is asymptotically the same as the wavelet–best N–term approximation of each solution component, and the total computational work is proportional to the number of computational unknowns.

References

[1] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods for elliptic operator equations – Convergence rates, Math. Comp., 70 (2001), 27–75.

[2] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods II – Beyond the elliptic case, Found. Comput. Math., 2 (2002), 203–245.

[3] W. Dahmen, A. Kunoth, *Adaptive wavelet methods for linear–quadratic elliptic control problems: Convergence rates*, SIAM J. Contr. Optim. **43**(5) (2005), 1640–1675.

[4] M. Gunzburger, A. Kunoth, *Space-time adaptive wavelet methods for control problems constrained by parabolic evolution equations*, SIAM J. Contr. Optim. **49**(3) (2011), 1150–1170.

[5] Chr. Schwab, R. Stevenson, *Space–time adaptive wavelet methods for parabolic evolution equations*, Math. Comp. **78**, (2009) 1293–1318.

Numerical approximations of linear controllability and inverse problems using space-time variational methods

Arnaud Münch

We address the numerical approximation of null controls for linear partial differential equations and emphasize the interest of space-time variational methods, well suited for adaptivity. Seminal contributions in this topic are due to Roland Glowinski and co-workers in the nineties [3] using time marching variational methods based on duality arguments. As an example, let us consider the boundary controllability of the following hyperbolic equation

$$
(1)\qquad \begin{cases} Ly := y_{tt} - \nabla\cdot(c(x)\nabla y) + d(x,t)y = f, & (x,t)\in Q_T := \Omega\times(0,T),\\ (y(\cdot,0), y_t(\cdot,0)) = (y_0,y_1)\in \boldsymbol{H}, & x\in\Omega,\\ y = v1_{\Gamma_0}, & (x,t)\in\Sigma_T := \partial\Omega\times(0,T), \end{cases}
$$

posed over a bounded domain $\Omega\subset\mathbb{R}^N$ ($N\geq 1$), with C^2 boundary. Here, $\boldsymbol{H} := L^2(\Omega)\times H^{-1}(\Omega)$, $f\in L^2(Q_T)$, $c\in C^1(\overline{\Omega};\mathbb{R}_+^*)$, $d\in L^\infty(Q_T)$. v is a control function in $L^2(\Sigma_T)$ acting on the non empty subset Γ_0 of the boundary $\partial\Omega$. Let $\mathcal{C}(y,v)$ be defined as the space of controls v in $L^2(\Sigma_T)$ and states y such that $y=y(v)$ solves (1) and satisfies the null controllability condition $(y(\cdot,T), y_t(\cdot,T)) = (0,0)$ in Ω. Under conditions on the triplet (Ω,T,Γ_0) given in [1], $\mathcal{C}(y,v)$ is non empty.

Duality arguments then allow to show that the control of minimal L^2-norm is given by $v = c(x)\nabla\varphi \cdot \boldsymbol{\nu}\, 1_{\Gamma_0}$ where φ solves the adjoint equation

$$L^\star\varphi := 0 \text{ in } Q_T, \quad (\varphi(\cdot,0), \varphi_t(\cdot,0)) = (\varphi_0, \varphi_1) \in \boldsymbol{H'} \text{ in } \Omega, \quad \varphi = 0 \text{ on } \Sigma_T \tag{2}$$

associated to (φ_0, φ_1) solution of the optimal control problem:

$$\inf_{(\varphi_0,\varphi_1)\in \boldsymbol{H'}} J^\star(\varphi_0,\varphi_1) := \frac{1}{2}\|c(x)\nabla\varphi\cdot\boldsymbol{\nu}\|^2_{L^2(\Gamma_0\times(0,T))} + ((\varphi_0,\varphi_1),(-y_1,y_0))_{\boldsymbol{H'},\boldsymbol{H}}. \tag{3}$$

The main tool to assert the well-posedness of (3) (and therefore the null controllability of (1)) is the observability property for (2): there exists a constant $C_{obs} > 0$ such that

$$\|\varphi_0,\varphi_1\|^2_{\boldsymbol{H'}} \le C_{obs}\|c(x)\nabla\varphi\cdot\boldsymbol{\nu}\|^2_{L^2(\Gamma_0\times(0,T))}, \quad \forall(\varphi_0,\varphi_1)\in\boldsymbol{H'}. \tag{4}$$

The minimization of $J^\star$ can be done using a gradient method which requires at each iteration the resolution of (2) by a time marching method. This raises the issue of the convergence of the discrete controls with respect to the approximation parameters. Uniformly controllable schemes are available in the literature only for simple situations (see [6] for the one dimensional case and constant coefficients). On the other hand, the conjugate functional $J^\star$ can be minimized directly with respect to the adjoint variable φ in the space-time functional space $\boldsymbol{W} := \{\varphi : \varphi \in C^0(0,T;H^1_0(\Omega)) \cap C^1(0,T;L^2(\Omega)); L^\star\varphi = 0 \,\text{in}\, Q_T\}$. The constraint equality $L^\star\varphi = 0$ in $L^2(Q_T)$ is then taken into account using a Lagrange multiplier $\lambda \in L^2(Q_T)$ leading to the saddle point formulation :

$$\sup_{\lambda\in L^2(Q_T)} \inf_{\varphi\in\boldsymbol{\Phi}} \mathcal{L}(\varphi,\lambda) := J^\star(\varphi) + < L^\star\varphi, \lambda >_{L^2(Q_T)} \tag{5}$$

where $\boldsymbol{\Phi} := \{\varphi : \varphi \in C^0(0,T;H^1_0(\Omega)) \cap C^1(0,T;L^2(\Omega)); L^\star\varphi \in L^2(Q_T)\} \supset \boldsymbol{W}$. The main tool to prove the well-posededness of (5) is a global estimate (deduced from (4)): there exists a constant $C > 0$ such that

$$\|\varphi(\cdot,0),\varphi_t(\cdot,0)\|^2_{\boldsymbol{H'}} \le C\left(\|c(x)\nabla\varphi\cdot\boldsymbol{\nu}\|^2_{L^2(\Gamma_0\times(0,T))} + \|L^\star\varphi\|^2_{L^2(Q_T)}\right), \quad \forall\varphi\in\boldsymbol{\Phi}. \tag{6}$$

At the discretization level, the main interest of the formulation (5) with respect to (3) is that the estimate (6) still holds for any finite dimensional space $\boldsymbol{\Phi}_h \subset \boldsymbol{\Phi}$ with the same constant C independent of a discretization parameter h (contrarily to (4)). The first order optimality system associated to $\mathcal{L}$ can be solved over $\boldsymbol{\Phi_h}\times\boldsymbol{\Lambda_h} \subset \boldsymbol{\Phi}\times L^2(Q_T)$ leading, for appropriate choices of the spaces $\boldsymbol{\Phi}_h$ and $\boldsymbol{\Lambda_h}$, to strong convergent properties as $h \to 0^+$ of the form

$$\|L^\star(\varphi-\varphi_h)\|_{L^2(Q_T)} + \|c(x)\nabla(\varphi-\varphi_h)\cdot\boldsymbol{\nu}\|_{L^2(\Gamma_0\times(0,T))} + \|\lambda-\lambda_h\|_{L^2(Q_T)} \to 0.$$

In particular, the function $v_h := c(x)\nabla\varphi_h\cdot\boldsymbol{\nu}\, 1_{\Gamma_0}$ defines a strong convergent approximation of the control v. Moreover, this variational approach allows the adaptivity of the discrete space-time mesh (of $\overline{Q_T}$) in order to capture singularities of the solution and/or to reduce the computational cost. As an illustration, Figure 1 displays the approximation λ_h of the multiplier λ (which is actually the

controlled solution y) over an adapted mesh for $\Omega = (0,1)$, $T = 2.4$, $\Gamma_0 = \{1\}$, $c := 1$, $d := 0$ and $(y_0(x), y_1(x)) = (4x\,1_{(0,1/2)}(x), 0)$ for which the optimal control is $v(t) = 2(1-t)1_{(1/2,3/2)}(t)$. The refinement occurs along the line of singularity of the controlled solution generated by the discontinuity of y_0 at $x = 1/2$ (we refer to [4] for details).

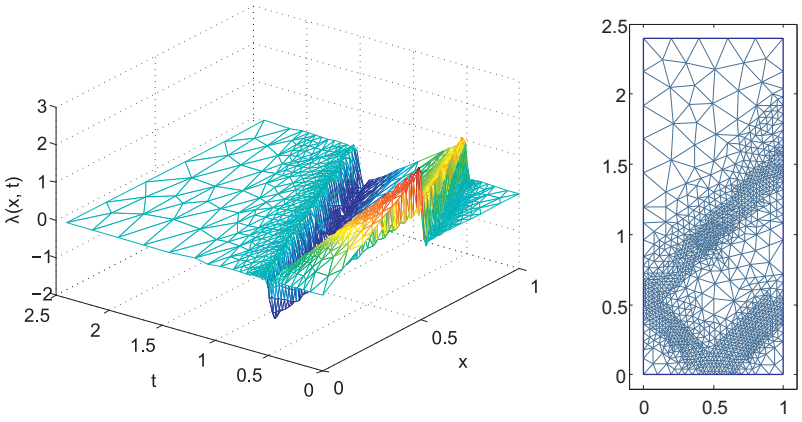


FIGURE 1. Approximated solution $\lambda_h \subset \Lambda$ in Q_T (left) using an adapted mesh (right).

This space-time variational approach may also be employed to approximate control problems for parabolic equations

$$\text{(7)} \qquad \begin{cases} Ly := y_t - \nabla \cdot (c(x)\nabla y) + d(x,t)y = v\,1_\omega, & (x,t) \in Q_T := \Omega \times (0,T), \\ y(\cdot,0) = y_0,\ x \in \Omega, \qquad y = 0,\ (x,t) \in \Sigma_T, \end{cases}$$

with y_0 in $L^2(\Omega)$. v is the control acting on ω, non empty subset of Ω. From [2], the set $\mathcal{C}(y,v)$ is non empty for any $y_0 \in L^2(\Omega)$, $T > 0$ and $\omega \subset \Omega$. The key tool to apply the method is a Carleman global estimate (similar to (6)) proved in [2]: there exists $C > 0$ such that

$$\text{(8)} \qquad \|\varphi(\cdot,0)\|^2_{H_0^1(\Omega)} \le C\Big(\|\rho_0^{-2}\varphi\|^2_{L^2(\omega\times(0,T))} + \|\rho^{-2}L^\star\varphi\|^2_{L^2(Q_T)} \Big), \quad \forall \varphi \in \mathbf{\Phi}$$

with weights $\rho, \rho_0 \in L^\infty(Q_T, \mathbb{R}_\star^+)$ vanishing exponentially as $t \to T^-$. The space $\mathbf{\Phi}$ is defined as the completion of $\mathbf{\Phi}_0 := \{\varphi \in C^2(\overline{Q_T}); \varphi = 0 \,\text{on}\, \Sigma_T\}$ with respect to the scalar product defined by the right hand side term of (8). $\mathbf{\Phi}$ is included in $C([0, T-\delta], H_0^1(\Omega))$ for any $\delta > 0$ so that the solution φ of (5) has a singular behavior as $t \to T^-$. In view of this property (due to the regularization effect of the heat kernel), the refinement of the space-time mesh is mainly necessary in the neighborhood of $\Omega \times \{t = 0\}$ and allows the strong convergence of the discrete controls defined here by $v_h := \varphi_h\,1_\omega$ as $h \to 0$. Remark that such strong

convergence is still an open issue when a time marching discretization is employed. Eventually, such variational methods may be used as well for inverse problems and allow to reconstruct with robustness the whole solution of a partial differential equations from a partial observation (see [5]).

References

[1] C. Bardos, G. Lebeau and J. Rauch. Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary. SIAM J. Control Optim., 30, 1992, 1024–1065.

[2] A.V. Fursikov and O.Yu Imanuvilov. Controllability of evolution equations. Lecture Notes Series, 34. Seoul National University, Research Institute of Mathematics. 1996. iv+163.

[3] R. Glowinski, J.-L. Lions and H. Jiwen. Exact and approximate controllability for distributed parameter systems. Encyclopedia of Mathematics and its Applications,117. Cambridge University Press, Cambridge, 2008. pages xii+458.

[4] N. Cîndea and A. Münch. A mixed formulation for the direct approximation of the control of minimal L^2-norm for linear type wave equations. Calcolo, 52, 2015, 245–288.

[5] N. Cîndea and A. Münch. Inverse problems for linear hyperbolic equations using mixed formulations. Inverse Problems, 31, 2015, 075001, 38.

[6] A. Münch. A uniformly controllable and implicit scheme for the 1-D wave equation. Math. Model. Numer. Anal. 39, 2005, 377-418.

Error estimation for discrete POD models of parabolic control problems

Ira Neitzel

(joint work with Martin Gubisch and Stefan Volkwein)

We discuss a-posteriori error estimates for approximate solutions of linear-quadratic optimal control problems with control u and state y, governed by parabolic equations with pointwise control bounds, of the form

$$\text{Minimize } J(y,u) := \frac{1}{2}\int_0^T \int_\Omega (y(t,x) - y_d(t,x))^2 dx\, dt + \frac{\nu}{2}\sum_{i=1}^{N_u}\int_0^T u_i(t)^2\,\mathrm{d}t$$

subject to

$$\partial_t y - \Delta y = \sum_{i=1}^{N_u} u_i\chi_i + f \text{ in } (0,T)\times\Omega, \qquad y(0,\cdot) = y_0 \text{ in } \Omega,$$

$$y = 0 \text{ in } (0,T)\times\partial\Omega, \qquad u_a \le u(t) \le u_b \text{ almost everywhere in } (0,T),$$

where the last inequality is to be understood componentwise. In this setting, $\Omega \subset \mathbb{R}^{2,3}$ is a polygonally or polyhedrally bounded domain, and the real numbers $T > 0$ and $\nu > 0$ denote a final time and cost parameter, respectively. Moreover, the initial state $y_0 \in H_0^1(\Omega)$, the desired state $y_d \in L^2(\Omega)$, and the control bounds $u_a, u_b \in \mathbb{R}^{N_u}$ are given. Note that the functions $\chi_i : \Omega \to \mathbb{R}$, $i = 1,\dots,N_u$, are also fixed data, and the finitely many controls depend on time, only.

We focus on estimating the error of a suboptimal solution obtained by applying the model-order reduction method of proper orthogonal decomposition (POD).

This question has been subject to earlier research in [7]. Therein, the authors interpret the obtained approximate solution as a solution of a perturbed full-order problem, cf. also Malanowski et. al., [4]. An a posteriori error estimator is obtained with the help of first order necessary optimality conditions. This approach, however, does not take into account the fact that the snapshots for the reduced order model are obtained from discretized PDEs. The error estimator contains (continuous) solutions of partial differential equations, and is therefore not computable in the strict sense.

In our work, we aim at closing this gap by combining error estimation techniques for finite element discretization and model-order reduction. We extend the approach of [7] to estimate the error between the computed solution $\bar{u}_{kh}^p$ of the lower order model and the solution $\bar{u}_{kh}$ of a finite-element-discretization of the full-order model. The error between the latter and the solution $\bar{u}$ of the continuous full-order problem is known a priori to be of order $k + h^2$, if k denotes the time-discretization parameter for discontinuous Galerkin discretization in time of order zero, and h the spatial discretization parameter of a usual H^1-conforming continuous Galerkin discretization in space, cf. [6]. We use this to design an updating algorithm for the reduced-order models, motivated by the following two observations:

(1) The reduced-order model already includes the finite element discretization errors and it does not seem reasonable to decrease the POD residual below the order of the finite element discretization error.
(2) If, however, the error of the POD approximation does not reflect the order of the discretization error even when increasing the size of the reduced order model, the current POD basis likely does not reflect the dynamics of the optimal POD basis, which would be obtained from the (unknown) optimal state. Then, an update of the POD basis may be required to improve the results

At this point, the order of the finite-element discretization error is determined by estimating numerically the constants in the a priori error estimate from [6]. Relying completely on computable a posteriori error estimates would be a desirable result of future research.

Further information can be found in [3] and [1], where in the latter the reader can also find results for problems with regularized pointwise state constraints. In [2], detailed error analysis for the model reduction error is carried out. Concerning a posteriori discretization error estimates and adaptive finite element discretization for parabolic problems without inequality constraints we refer for instance to [5] and the references therein.

References

[1] M. Gubisch Model order reduction techniques for the optimal control of parabolic partial differential equations with control and state constraints. Phd thesis, University of Konstanz, 2016.

[2] M. Gubisch, S. Volkwein. Proper orthogonal decomposition for linear-quadratic optimal control. To appear in P. Benner, A. Cohen, M. Ohlberger, and K. Willcox (eds), Model Reduction and Approximation: Theory and Algorithms. SIAM, Philadelphia, PA, 2017.
[3] M. Gubisch, I. Neitzel, S. Volkwein. A-posteriori error estimation of discrete POD models for PDE-constrained optimal controls. INS Preprint No. 1608, 2016. Accepted for publication.
[4] K. Malanowski, C. Büskens, and H. Maurer. Convergence of approximations to nonlinear optimal control problems. in A.V. Fiacco (ed.): Mathematical programming with data perturbations. Pure Appl. Math. 195, 253–284, 1998.
[5] D. Meidner and B. Vexler. Adaptive space-time finite element methods for parabolic optimization problems. SIAM J.Contr. Optim. 46 (1), 116–142, 2007.
[6] D. Meidner and B. Vexler. A priori error estimates for space-time finite element discretization of parabolic optimal control problems Part II: Problems with control constraints. SIAM J.Contr. Optim. 47 (3), 1301–1329, 2008.
[7] F. Tröltzsch and S. Volkwein. POD a-posteriori error estimates for linear-quadratic optimal control problems. Comp. Optim. Appl.44(1), 83–115, 2009.

A posteriori error analysis for optimal control problems

Arnd Rösch
(joint work with Kristina Kohls, Christian Kreuzer, Kunibert G. Siebert)

We discuss an elliptic optimal control problem

$$\begin{aligned} &\min_{(u,y)\in\mathbb{U}^{\mathrm{ad}}\times\mathbb{Y}} \mathcal{J}[u,y] = \psi(y) + \frac{\alpha}{2}\|u\|_{\mathbb{U}}^2 \\ &\text{subject to} \qquad y\in\mathbb{Y}: \quad \mathcal{B}[y,\,v] = \langle f+u,\,v\rangle_{\mathbb{Y}\times\mathbb{Y}^*} \qquad \forall v\in\mathbb{Y}, \end{aligned} \tag{1}$$

where ψ is assumed to be a convex quadratic functional. For the bilinear form $\mathcal{B}$ we require an inf-sup condition. The set of admissible controls is formulated as

$$\mathbb{U}^{\mathrm{ad}} = \{u\in\mathbb{U}, u\in\mathcal{C} \text{ a.e. in } \Gamma\} \tag{2}$$

with a given closed convex set $\mathcal{C}\subset\mathbb{R}^m$.

We analyze the discretization of such a problem by an adaptive finite element method (AFEM). We study the variational discretization as well as a full discretized problem. In the case of the variational discretization we consider

$$\begin{aligned} &\min_{(U,Y)\in\mathbb{U}^{\mathrm{ad}}\times\mathbb{Y}(\mathcal{G})} \mathcal{J}[U,Y] = \psi(Y) + \frac{\alpha}{2}\|U\|_{\mathbb{U}}^2 \\ &\text{subject to} \qquad Y\in\mathbb{Y}(\mathcal{G}): \quad \mathcal{B}[Y,\,V] = \langle f+U,\,V\rangle \qquad \forall V\in\mathbb{Y}(\mathcal{G}). \end{aligned} \tag{3}$$

The full discretized problem is given by

$$\begin{aligned} &\min_{(U,Y)\in\mathbb{U}^{\mathrm{ad}}(\mathcal{G})\times\mathbb{Y}(\mathcal{G})} \mathcal{J}[U,Y] = \psi(Y) + \frac{\alpha}{2}\|U\|_{\mathbb{U}}^2 \\ &\text{subject to} \qquad Y\in\mathbb{Y}(\mathcal{G}): \quad \mathcal{B}[Y,\,V] = \langle f+U,\,V\rangle \qquad \forall V\in\mathbb{Y}(\mathcal{G}). \end{aligned} \tag{4}$$

where the discrete admissible set is defined by

$$\mathbb{U}^{\mathrm{ad}}(\mathcal{G}) := \mathbb{U}^{\mathrm{ad}}\cap\mathbb{U}(\mathcal{G}).$$

In both cases we study the loop of an adaptive finite element discretization

$$\mathsf{SOLVE} \longrightarrow \mathsf{ESTIMATE} \longrightarrow \mathsf{MARK} \longrightarrow \mathsf{REFINE}. \tag{5}$$

For each of these parts we formulate certain requirements. These requirements are satisfied for a large class of control discretizations, state equations, error estimators, and all standard marking strategies. The framework covers boundary control or/and with boundary observation.

Our main result states convergence of the AFEM-solutions to the solution of the optimal control problem and the convergence of the error indicator to zero.

References

[1] K. Kohls, A. Rösch, and K. G. Siebert. A posteriori error analysis of optimal control problems with control constraints. SIAM Journal Control and Optimization 52(3): 1832–1861, 2014.

[2] K. Kohls, A. Rösch, and K. G. Siebert. Convergence of Adaptive Finite Elements for Optimal Control Problems with Control Constrains. Trends in PDE Constrained Optimization, vol. 165 of Internat. Ser. Numer. Math., Birkhäuser/Springer, Basel, 2015, pp. 403–419.

[3] P. Morin, K. G. Siebert, and A. Veeser. A basic convergence result for conforming adaptive finite elements. Math. Models Methods Appl. Sci., 18(5) (2008), pp. 707–737.

[4] K. G. Siebert. A convergence proof for adaptive finite elements without lower bound. IMA J. Numer. Anal., 31(3) (2011), pp. 947–970.

Space–time finite element methods and distributed control problems

Olaf Steinbach

(joint work with Martin Neumüller, Huidong Yang, Markus Fleischhacker)

As model problem we consider the Dirichlet problem for the heat equation,

$$\begin{aligned} \alpha\partial_t u(x,t) - \Delta_x u(x,t) &= f(x,t) && \text{for } (x,t) \in Q := \Omega \times (0,T), \\ u(x,t) &= 0 && \text{for } (x,t) \in \Sigma := \Gamma \times (0,T), \\ u(x,0) &= 0 && \text{for } x \in \Omega, \end{aligned}$$

where $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, is a bounded domain with Lipschitz boundary $\Gamma = \partial\Omega$, $\alpha > 0$ is the heat capacity, and f is a given source term. The related variational formulation is to find $u \in X := L^2(0,T; H^1_0(\Omega)) \cap H^1_{0,}(0,T; H^{-1}(\Omega))$ satisfying

$$\int_0^T \int_\Omega \Big[\alpha\partial_t u(x,t)v(x,t) + \nabla_x u(x,t) \cdot \nabla_x v(x,t)\Big] dx\, dt = \int_0^T \int_\Omega f(x,t)v(x,t)\, dx\, dt$$

for all $v \in Y := L^2(0,T; H^1_0(\Omega))$. Unique solvability of this Galerkin–Petrov variational formulation then follows from the stability condition for the associated bilinear form, see, e.g., [4],

$$\|\alpha\partial_t u - \Delta_x u\|_{L^2(0,T;H^{-1}(\Omega))} \le \sup_{0 \neq v \in L^2(0,T;H^1_0(\Omega))} \frac{a_\Omega(u,v)}{\|v\|_{L^2(0,T;H^1_0(\Omega))}} \quad \text{for all } u \in X.$$

For the discretization we may use either a discontinuous Galerkin approach, [2], or a conforming finite element method, [4]. For the latter we introduce finite element spaces $X_h \subset X$ and $Y_h \subset Y$, where we need to ensure $X_h \subset Y_h$. For example, we can chose $X_h = Y_h$ as the finite element space of piecewise linear and continuous functions with vanishing Dirichlet and initial conditions. Then, and by using a discrete norm in X_h, we can establish a discrete inf–sup stability condition

from which we conclude unique solvability of the Galerkin–Petrov finite element method, and quasi–optimality, i.e. Cea's lemma, see [4].

The definition of the finite element spaces X_h and Y_h is done with respect to an admissible and locally quasi–uniform decomposition of the space–time cylinder $Q = \Omega \times (0,T) \subset \mathbb{R}^{n+1}$ into simplicial finite elements, e.g. tetrahedra for $n = 2$, and pentatops for $n = 3$, see [3]. In particular we can use unstructured and locally refined meshes to resolve the discrete solution. Any appropriate algorithm for adaptive mesh refinement is based on suitable a posteriori error estimators, see, e.g., [5]. Here we use residual type techniques to estimate the local residua and the jump of the flux in the spatial direction. Numerical examples include not only the linear heat equation, but also the heat equation with nonlinear reaction terms as they appear, for example, in the Schlögl model, and in the Nagumo equation.

This adaptive space–time finite element approach is then used for the numerical approximation of distributed control problems with the linear heat equation as constraint. As model problem we consider the minimization of the cost functional

$$\mathcal{J}(u,z) := \frac{1}{2}\int_\Omega [u(x,T) - \overline{u}(x)]^2 dx + \frac{1}{2}\varrho \|z\|^2_{L^2(0,T;H^{-1}(\Omega))}$$

subject to the Dirichlet problem for the heat equation,

$$\begin{aligned} \partial_t u(x,t) - \Delta_x u(x,t) &= z(x,t) && \text{for } (x,t) \in Q := \Omega \times (0,T), \\ u(x,t) &= g(x,t) && \text{for } (x,t) \in \Sigma := \Gamma \times (0,T), \\ u(x,0) &= u_0(x) && \text{for } x \in \Omega. \end{aligned}$$

Since we consider the control in the energy space $L^2(0,T;H^{-1}(\Omega))$, the associated norm is realized by solving a quasi–static Dirichlet boundary value problem for the Laplacian. In addition to the primal Dirichlet problem for the heat equation we have to solve the adjoint problem

$$\begin{aligned} \partial_t p(x,t) + \Delta_x p(x,t) &= 0 && \text{for } (x,t) \in Q, \\ p(x,t) &= 0 && \text{for } (x,t) \in \Sigma, \\ p(x,T) &= u(x,T) - \overline{u}(x) && \text{for } x \in \Omega, \end{aligned}$$

the optimality or gradient condition

$$p(x,t) + \varrho w(x,t) = 0 \quad \text{for } (x,t) \in Q,$$

and the quasi–static Dirichlet problem

$$-\Delta_x w(x,t) = z(x,t) \quad \text{for } (x,t) \in Q, \quad w(x,t) = 0 \quad \text{for } (x,t) \in \Sigma.$$

Since the adjoint Dirichlet problem for the heat equation includes a termination condition, a space–time finite element method seems to be an appropriate choice for the overall discretization of the coupled system. The stability and error analysis of the finite element approximation can be done as in the case of the primal problem. First numerical examples confirm the theoretical estimates.

More challenging is the construction of suitable preconditioning strategies for an efficient iterative solution of the resulting linear system of algebraic equations. When using time slabs one can design a geometric space–time multigrid method

[1, 2] with optimal coarsening in space and time. In fact, this also allows for an optimal parallelization of the finite element solver for the heat equation, with almost perfect scaling. For completely unstructured space–time finite element meshes one has to use algebraic multigrid methods instead of a geometric approach. Preliminary numerical results are promising.

From an application point of view we need to consider the heat equation with nonlinear reaction terms also for the optimal control problem. In addition we may consider control constraints in $L^2(0,T;H^{-1}(\Omega))$, which result in a first kind variational inequality. After discretization, the nonlinear discrete system can be solved by a semi–smooth Newton method.

Acknowledgement: This work has been supported by the Austrian Science Fund (FWF) under the grant SFB Mathematical Optimization and Applications in Biomedical Sciences.

References

[1] M. Gander, M. Neumüller. Analysis of a new space–time parallel multigrid algorithm for parabolic problems. SIAM J. Sci. Comput., 38, 2016) no. 4, A2173–A2208.
[2] M. Neumüller. Space–time methods: Fast solvers and applications. PhD thesis, TU Graz, 2013.
[3] M. Neumüller, O. Steinbach. Refinement of flexible space–time finite element meshes and discontinuous Galerkin methods. Comput. Visual. Sci., 14, 2011, 189–205.
[4] O. Steinbach. Space–time finite element methods for parabolic problems. Comput. Meth. Appl. Math., 15, 2015, 551–566.
[5] R. Verfürth: A posteriori error estimation techniques for finite element methods. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

Sparse optimal control of FitzHugh-Nagumo equations and strong local minima

Fredi Tröltzsch
(joint work with Eduardo Casas and Christopher Ryll)

We consider the following class of optimal control problems:

$$\text{Minimize} \quad \frac{1}{2}\int_Q (y(x,t)-y_Q(x,t))^2\,dxdt + \frac{\nu}{2}\int_Q u^2(x,t)\,dxdt + \mu\int_Q |u(x,t)|\,dxdt \tag{1}$$

subject to the parabolic state equation

$$\begin{aligned} \frac{\partial y}{\partial t}(x,t) - \Delta y(x,t) + R(x,t,y(x,t)) &= u(x,t) && \text{in } Q \\ \partial_n y(x,t) &= 0 && \text{on } \Sigma \\ y(x,0) &= y_0(x) && \text{in } \Omega \end{aligned} \tag{2}$$

and to the box constraints

$$u \in \mathcal{U}_{\rm ad} := \{u \in L^\infty(Q) : a \le u(x,t) \le b \quad \text{for almost all } (x,t) \in Q\}. \tag{3}$$

Here, $\Omega \subset \mathbb{R}^N$, $N \in \mathbb{N}$, is a bounded Lipschitz domain with boundary Γ; $T > 0$ is a fixed time, $Q := \Omega \times (0,T)$, and $\Sigma := \Gamma \times (0,T)$. Moreover, constants $\nu \geq 0$, $a < 0 < b$, initial data $y_0 \in L^\infty(\Omega)$, and a desired state $y_Q \in L^p(Q)$ with $p > N/2 + 1$ are given. By ∂_n, the normal derivative at Σ is denoted.

The nonlinearity ("reaction term") $R : \mathbb{R} \to \mathbb{R}$ has the form

$$R(y) = \rho\,(y - y_1)(y - y_2)(y - y_3),$$

with fixed real numbers $\rho > 0$ and $y_1 < y_2 < y_3$.

In Theoretical Physics, equation (2) is known as Schlögl model; it is also called Nagumo or Allen Cahn equation and, for a special choice of R, Chafee-Infante equation. We also consider the more general FitzHugh-Nagumo equations

$$\begin{aligned}
\frac{\partial y}{\partial t}(x,t) - \Delta y(x,t) + R(y(x,t)) + \alpha\, z(x,t) &= u(x,t) && \text{in } Q \\
\partial_n y(x,t) &= 0 && \text{on } \Sigma \\
y(x,0) &= y_0(x) && \text{in } \Omega \\
\frac{\partial z}{\partial t}(x,t) + \beta\, z(x,t) - \gamma\, y(x,t) + \delta &= 0 && \text{in } Q \\
z(x,0) &= z_0(x) && \text{in } \Omega
\end{aligned}$$

and associated control problems with more general objective functionals. Here, the *activator* y is complemented by the *inhibitor* z. For convenience, we present the optimality conditions only for the simpler Schlögl model. The FitzHugh-Nagumo equations can be discussed analogously by some obvious modifications.

The theory relies on the following basic facts that can be shown in a standard way, cf. [2, 3]:

- For each $u \in L^p(Q)$, equation (2) has a unique weak solution $y \in Y := L^\infty(Q) \cap W(0,T)$ denoted by y_u; moreover, y belongs to $C(\Omega \times (0,T])$.
- The control-to-state mapping $G : L^p(Q) \to Y$ is of class C^∞.
- For all $\nu \geq 0$, $\mu \geq 0$, problem (1)-(3) has at least one (globally) optimal control that is denoted by $\bar{u}_\nu$.

The subscript ν indicates that $\bar{u}_\nu$ is associated with the Tikhonov regularization parameter ν. We mention ν, since we are particularly interested in passing to the limit $\nu \to 0$, while μ is fixed. For $\nu = 0$, we write $\bar{u}_0 =: \bar{u}$.

The parameter μ accounts for the sparsity of locally optimal controls. The larger μ is, the larger is the set, where $\bar{u}$ is vanishing, cf. Theorem 2 below.

First- and second-order optimality conditions need the adjoint state φ_u that, for the Schlögl model, is the unique solution to the adjoint equation

$$\begin{aligned}
-\frac{\partial \varphi_u}{\partial t} - \Delta \varphi_u + R'(y_u)\,\varphi_u &= y_u - y_Q && \text{in } Q \\
\partial_n \varphi_u &= 0 && \text{on } \Sigma \\
\varphi_u(x,T) &= 0 && \text{in } \Omega.
\end{aligned} \tag{4}$$

Let us introduce the L^1-norm functional $j : u \mapsto \|u\|_{L^1(Q)}$, $j : L^\infty(Q) \to \mathbb{R}$. By $\partial j(u)$, we denote its subdifferential at u.

Theorem 1 (Necessary optimality conditions)**.** *If $\bar{u}_\nu$ is optimal for* (1)-(3)*, then there exist a unique adjoint state $\varphi_\nu := \varphi_{\bar{u}_\nu}$ solving* (4) *for $u := \bar{u}_\nu$, and a function $\bar{\lambda}_\nu \in \partial j(\bar{u}_\nu)$ such that*

$$\int_Q (\bar{\varphi}_\nu + \nu\bar{u}_\nu + \mu\bar{\lambda}_\nu)(u - \bar{u}_\nu)\, dx\, dt \geq 0 \quad \forall u \in \mathcal{U}_{\rm ad}. \tag{5}$$

By a pointwise discussion of the variational inequality (5), we obtain

Theorem 2 (Sparsity)**.** *Let $\bar{u}_\nu$, $\bar{\varphi}_\nu$, and $\bar{\lambda}_\nu$ be as in Theorem 1, and assume that $\nu > 0$. Then the following relations hold a.e. in Q,*

$$\begin{aligned}
&\bar{u}_\nu(x,t) = 0 \Leftrightarrow |\bar{\varphi}_\nu(x,t)| \leq \mu,\\
&\bar{u}_\nu(x,t) = \mathrm{Proj}_{[a,b]}\left(-\frac{1}{\nu}[\bar{\varphi}_\nu(x,t) + \mu\bar{\lambda}_\nu(x,t)]\right),\\
&\bar{\lambda}_\nu(x,t) = \mathrm{Proj}_{[-1,+1]}\left(-\frac{1}{\mu}\bar{\varphi}_\nu(x,t)\right);
\end{aligned}$$

in particular, $\bar{\lambda}_\nu \in \partial j(\bar{u}_\nu)$ is unique.

This theorem reveals two important facts: The larger μ is, the more points $(x,t) \in Q$ satisfy the first inequality. Hence, the set of points where $\bar{u}_\nu$ vanishes, increases with μ. This is the sparsity property. Moreover, it is not difficult to prove that there exists $\mu_0 > 0$ such that $\bar{u}_\nu \equiv 0$ if $\mu \geq \mu_0$. The uniqueness of $\bar{\lambda} \in \partial j(\bar{u}_\nu)$ is decisive for gradient computations in the conjugate gradient method or the semismooth Newton method.

In our numerical tests for the optimal control of the FitzHugh-Nagumo equations we observed a stable behavior of the optimal controls $\bar{u}_\nu$ as $\nu \downarrow 0$. Numerically, we saw a Lipschitz behavior of the controls, i.e. $\|\bar{u}_\nu - \bar{u}\| \leq c\,\nu$, cf. [2]. This indicates that some kind of second-order sufficient optimality condition should be satisfied at $\bar{u}$. What form should such a condition have? While second-order sufficient optimality conditions are almost standard for fixed $\nu > 0$, it is more delicate to formulate them with vanishing Tikhonov parameter ν.

Define

$$F(u) := \frac{1}{2}\|y_u - y_Q\|^2_{L^2(Q)}, \qquad J(u) := F(u) + \mu\,\|u\|_{L^1(Q)}.$$

Associated with $\bar{u}$ and fixed $\tau > 0$, we introduce the extended critical cone $E^\tau_{\bar{u}}$ that consists of all directions $v \in L^2(Q)$ satisfying

$$F'(\bar{u})v + \mu\, j'(\bar{u}; v) \leq \tau\|\eta_v\|_{L^2(Q)} \tag{6}$$

and the sign conditions

$$v(x,t)\begin{cases} \geq 0 & \text{if} \quad \bar{u}(x,t) = a\\ \leq 0 & \text{if} \quad \bar{u}(x,t) = b, \end{cases}$$

where $\eta_v := G'(\bar{u})\, v$.

Theorem 3 (Second order sufficient condition)**.** *Let $\bar{u} \in \mathcal{U}_{\rm ad}$ along with the state $\bar{y}$ and the adjoint state $\bar{\varphi}$ satisfy the first order necessary optimality conditions of Theorem 1. Assume that*

$$\exists \tau > 0 \quad \text{and} \quad \exists \sigma > 0 \quad \text{such that} \quad F''(\bar{u})v^2 \geq \sigma \, \|\eta_v\|^2_{L^2(Q)} \quad \forall v \in E^\tau_{\bar{u}}. \tag{7}$$

Then there exist $\varepsilon > 0$ and $\delta > 0$ such that for all $u \in \mathcal{U}_{\rm ad}$ with $\|y_u - \bar{y}\|_{L^\infty(Q)} < \varepsilon$,

$$J(\bar{u}) + \frac{\delta}{2}\|y_u - \bar{y}\|^2_{L^2(Q)} \leq J(u). \tag{8}$$

If $\bar{u}$ satisfies the conditions of this theorem, then it is a local minimizer that is strong in the sense of calculus of variations. The first result on strong local solutions in the control of PDEs was published in [1] for semilinear elliptic equations.

Remark. *We can also work with the well known cone $C^\tau_{\bar{u}}$ of critical directions. It is defined similarly as $E^\tau_{\bar{u}}$, but we have to substitute* (6) *by the condition $v(x,t) = 0$ if $|\bar{\varphi}(x,t)| \geq \tau$. Then Theorem 3 holds true for all $u \in \mathcal{U}_{\rm ad}$ with $\|u - \bar{u}\|_{L^2(Q)} < \varepsilon$; hence, $\bar{u}$ is a weak local minimizer.*

Theorem 4 (Hölder stability as $\nu \downarrow 0$)**.** *Let $\bar{u}$ be a strict locally optimal control of our control problem for $\nu = 0$ and $\{\bar{u}_\nu\}$ be a sequence of local solutions of problem* (1)-(3) *with $\nu > 0$ such that $\|\bar{u}_\nu - \bar{u}\|_{L^2(Q)} \to 0$ for $\nu \to 0$ (such a sequence can be proven to exist). Assume that $\bar{u}$ satisfies the second-order sufficient conditions of Theorem 3. Then*

$$\lim_{\nu \to 0} \frac{1}{\sqrt{\nu}} \|\bar{y}_\nu - \bar{y}\|_{L^2(Q)} = 0,$$

where $\bar{y}_\nu = G(\bar{u}_\nu)$ and $\bar{y} = G(\bar{u})$.

If we assume the second-order condition (7) for all $u \in C^\tau_{\bar{u}}$, then Theorem 4 remains true, [2, 3].

References

[1] T. Bayen, J. Bonnans, and F. Silva, Characterization of local quadratic growth for strong minima in the optimal control of semi-linear elliptic equations. Trans. Amer. Math. Soc., 366 (2014), 2063–2087.

[2] E. Casas, C. Ryll, and F. Tröltzsch, Second order and stability analysis for optimal sparse control of the FitzHugh-Nagumo equation. SIAM J. Control Optim. 53 (2015), 2168–2202.

[3] E. Casas and F. Tröltzsch, Second order optimality conditions for weak and strong local solutions of parabolic optimal control problems. Vietnam Journal of Mathematics, 44 (2016), 181–202.

Robust a posteriori error estimates for stabilized finite element discretizations of non-stationary convection-diffusion problems

Rüdiger Verfürth
(joint work with Lutz Tobiska)

There is a wide range of stabilized finite element methods for non-stationary convection-diffusion problems such as streamline diffusion methods, local projection schemes, subgrid-scale techniques, and continuous interior penalty methods to name only a few (cf. [2] and the references given there). In this talk, which is based on [1], we show that all these schemes give rise to the same robust a posteriori error estimates. Here, as usual, robustness means that the upper and lower bounds for the error are uniform with respect to the size of convection or reaction terms relative to the diffusion. Such robust a posteriori error estimates are an indispensable building block for efficient adaptive schemes that compute an approximate solution to the differential equation with a prescribed tolerance and (nearly) optimal complexity (cf. [3] and the references given there).

Our analysis is based on the general approach of [3] which gives the generic robust equivalence of error and residual and provides robust global upper and local lower bounds for the residual up to a consistency error in the upper bound. The latter depends on the particular stabilization method. Consequently, our main task consists in deriving explicit and computable upper bounds for the consistency errors of the various schemes. It turns out that all stabilization methods mentioned above yield cosistency errors that do not spoil the upper bound on the error.

References

[1] L. Tobiska, R. Verfürth. Robust a posteriori error estimates for stabilized finite element methods. IMA J. Numer. Anal. **35** (2015), no. 4, 1652–1671.
[2] H.-G. Roos, M. Stynes, and L. Tobiska. Robust Numerical Methods for Singularly Perturbed Differential Equations. Springer (2008).
[3] R. Verfürth. A Posteriori Error Estimation Techniques for Finite Element Methods. Oxford University Press (2013).

Discrete maximal parabolic regularity and best approximation results for Galerkin finite element solutions of parabolic problems

Boris Vexler
(joint work with Dmitriy Leykekhman)

This talk is devoted to discrete maximal parabolic regularity results for Galerkin finite element discretization of parabolic problems. Let Ω be a Lipschitz domain in $\mathbb{R}^d$, $d = 2, 3$ and $I = (0, T)$. We consider the heat equation as a model of a parabolic second order partial differential equation,

$$\begin{aligned} \partial_t u(t,x) - \Delta u(t,x) &= f(t,x), && (t,x) \in I \times \Omega, \\ u(t,x) &= 0, && (t,x) \in I \times \partial\Omega, \\ u(0,x) &= 0, && x \in \Omega \end{aligned}$$

with a right-hand side $f \in L^s(I; L^p(\Omega))$ for some $1 \le p, s \le \infty$.

The maximal parabolic regularity on the continuous level says that for every $1 < p, s < \infty$ there exists a constant C such that,

$$\|\partial_t u\|_{L^s(I;L^p(\Omega))} + \|\Delta u\|_{L^s(I;L^p(\Omega))} \le C \|f\|_{L^s(I;L^p(\Omega))}, \quad \text{for all} \quad f \in L^s(I; L^p(\Omega)).$$

The maximal parabolic regularity is an important analytical tool and has a number of applications, especially nonlinear problems and/or optimal control problems when sharp regularity results are required. In this talk we present similar maximal parabolic regularity results for time discrete discontinuous Galerkin solutions as well as for the fully discrete Galerkin approximations, see [4] for details. Such results allow to prove best approximation type estimates with respect to various norms, especially with respect to the $L^\infty(I \times \Omega)$ norm, see [3]. Moreover, they play a crucial role in establishing optimal order a priori error estimates for optimal control problems, e.g., in the context of pointwise control [1, 2].

The main result of [4] can be summarized as follows. Let u_k be the time semidiscrete discontinuous Galerkin solution of the above heat equation. Then there holds

$$\left(\sum_{m=1}^{M} \|\partial_t u_k\|^s_{L^s(I_m;L^p(\Omega))} \right)^{\frac{1}{s}} + \|\Delta u_k\|_{L^s(I;L^p(\Omega))} + \left(\sum_{m=1}^{M} k_m \left\| \frac{[u_k]_{m-1}}{k_m} \right\|^s_{L^p(\Omega)} \right)^{\frac{1}{s}} \le C \ln \frac{T}{k} \|f\|_{L^s(I;L^p(\Omega))},$$

for $1 \le s \le \infty$ and $1 \le p \le \infty$, with obvious notation changes in the case of $s = \infty$. Here, I_m with $k_m = |I_m|$ $(m = 1, 2, \ldots, M)$ are the time intervals for the temporal discretization and $[u_k]_{m-1}$ denotes the jump of the piecewise continuous (in time) solution. The appearance of the logarithmic term is natural for the critical values $s = 1$, $p = 1$, $s = \infty$, or $p = \infty$, since the corresponding maximal parabolic regularity results for the continuous problem hold only for $1 < s, p < \infty$. In [4] we provide also a fully discrete counterpart of the above result.

References

[1] D. Leykekhman and B. Vexler. Optimal a priori error estimates of parabolic optimal control problems with pointwise control. SIAM J. Numer. Anal., 51(5), 2013, 2797–2821.

[2] D. Leykekhman and B. Vexler. A priori error estimates for three dimensional parabolic optimal control problems with pointwise control. SIAM J. Control Optim., 54(5), 2016, 2403–2435.

[3] D. Leykekhman and B. Vexler. Pointwise best approximation results for Galerkin finite element solutions of parabolic problems. SIAM J. Numer. Anal., 54(3), 2016, 1365–1384.

[4] D. Leykekhman and B. Vexler. Discrete maximal parabolic regularity for Galerkin finite element methods. Numer. Math., published online, 2016.

Reporter: Angela Kunoth

Participants

Sandra Boschert
Mathematisches Institut
Universität zu Köln
Weyertal 86-90
50931 Köln
GERMANY

Dr. Tobias Breiten
Institute for Mathematics
and Scientific Computing
University of Graz
Heinrichstrasse 36/III
8010 Graz
AUSTRIA

Prof. Dr. Stephan Dahlke
Fachbereich Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Strasse (Lahnbg.)
35032 Marburg
GERMANY

Prof. Dr. Max D. Gunzburger
Department of Scientific Computing
Florida State University
Tallahassee FL 32306-4120
UNITED STATES

Prof. Dr. Helmut Harbrecht
Departement Mathematik und
Informatik
Universität Basel
Spiegelgasse 1
4051 Basel
SWITZERLAND

Prof. Dr. Roland Herzog
Fakultät für Mathematik
Technische Universität Chemnitz
Reichenhainer Strasse 41
09126 Chemnitz
GERMANY

Dr. Dörte Jando
I W R
Universität Heidelberg
Im Neuenheimer Feld 205
69120 Heidelberg
GERMANY

Prof. Dr. Guido Kanschat
I W R
Universität Heidelberg
Im Neuenheimer Feld 205
69120 Heidelberg
GERMANY

Prof. Dr. Christian Kreuzer
Fakultät für Mathematik
Ruhr-Universität Bochum
44801 Bochum
GERMANY

Prof. Dr. Angela Kunoth
Mathematisches Institut
Universität zu Köln
Weyertal 86-90
50931 Köln
GERMANY

Prof. Dr. Arnaud Diego Münch
Laboratoire de Mathématiques
Campus Universitaire des Cezeaux
3, place Vasarely
63178 Aubière Cedex
FRANCE

Prof. Dr. Ira Neitzel
Institut für Numerische Simulation
Universität Bonn
Wegelerstrasse 6
53115 Bonn
GERMANY

Prof. Dr. Arnd Rösch
Fakultät für Mathematik
Universität Duisburg-Essen
Thea-Leymann-Strasse 9
45127 Essen
GERMANY

Prof. Dr. Olaf Steinbach
Institut für Numerische Mathematik
Technische Universität Graz
Steyrergasse 30
8010 Graz
AUSTRIA

Prof. Dr. Fredi Tröltzsch
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 4-5
Strasse des 17. Juni 136
10623 Berlin
GERMANY

Prof. Dr. Rüdiger Verfürth
Fakultät für Mathematik
Ruhr-Universität Bochum
44780 Bochum
GERMANY

Prof. Dr. Boris Vexler
Zentrum Mathematik
Technische Universität München
Boltzmannstrasse 3
85748 Garching bei München
GERMANY