Matrix oriented reduction of space-time Petrov-Galerkin variational problems.

Julian Henning, Davide Palitta, Valeria Simoncini, and Karsten Urban

Abstract Variational formulations of time-dependent PDEs in space and time yield $(d + 1)$-dimensional problems to be solved numerically. This increases the number of unknowns as well as the storage amount. On the other hand, this approach enables adaptivity in space and time as well as model reduction w.r.t. both type of variables. In this paper, we show that matrix oriented techniques can significantly reduce the computational timings for solving the arising linear systems outperforming both time-stepping schemes and other solvers.

1 Introduction

Time-stepping schemes based upon variational semi-discretizations are the standard approach for the numerical solution of time-dependent partial differential equations (PDEs). Using a variational formulation in space and a subsequent discretization e.g. in terms of finite elements, one is left with an evolution problem in time. Standard finite difference techniques then yield a time-marching scheme, where a spatial problem needs to be solved in each time step.

Even though theoretical investigations on space-time variational formulations of PDEs have been around for a long time, [10], it was seen prohibitive to treat the time

Julian Henning, Karsten Urban

Ulm University, Institute for Numerical Mathematics, Helmholtzstr. 20, 89081 Ulm (Germany), e-mail: `\protect\T1\textbraceleftjulian.henning,karsten.urban\protect\T1\textbraceright@uni-ulm.de`

Davide Palitta

Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg (Germany), e-mail: `palitta@mpi-magdeburg.mpg.de`

Valeria Simoncini

Università di Bologna, Centro AM², Dipartimento di Matematica, Piazza di Porta S. Donato 5, 40127 Bologna (Italy), and IMATI-CNR, Pavia, e-mail: `valeria.simoncini@unibo.it`

as an additional variable for numerical simulations. In fact, if $\Omega \subset \mathbb{R}^d$ denotes the spatial domain, adding the time $t \in (0, T) =: I$ as an additional unknown results in a PDE on $\Omega_I := I \times \Omega$ in Dimension $d + 1$, which is costly both w.r.t. the amount of storage and the required computation time.

Also due to the increasing computing power the point of view has changed over the past years. In fact, being able to simulate problems for d up to three until about 20 years ago, adding another dimension seemed to be impossible. Nowadays, where we face high-dimensional problems (e.g. from quantum physics or finance) with $d \gg 100$, adding another dimension seems almost negligible. Another aspect to use space-time variational problems arose from model reduction of parameterized time-dependent PDEs. In fact treating both time and space as variables allows one to perform model reduction for space *and* time, [14]. The time-stepping model reduction approach yields a time-marching scheme for a reduced spatial dimension but with the same number of time steps, [7].

In this paper, we address the question of how to efficiently solve the linear systems arising from a (full) Petrov-Galerkin discretization of space-time variational formulations of time-dependent PDEs. It turns out that the involved coefficient matrices, combining space and time discretizations, have a tensorproduct structure, which allows us to use more efficient matrix equations solvers than what can be done with the usual vector representation.

This paper is organized as follows: In §2 we review space-time variational formulations of some PDEs and describe corresponding Petrov-Galerkin discretizations as well as the arising linear systems in §3; §4 is devoted to the description of the numerical schemes and §5 to numerical experiments, in particular the comparison with time-stepping schemes.

2 Space-Time Variational Formulation of PDEs

The Heat Equation. Let $A : X \to X'$ be an elliptic operator on $X := H_0^1(\Omega)$ associated to a coercive bilinear form $a : X \times X \to \mathbb{R}$, and $f \in L_2(I; X')$. We look for $u \in U := H_{(0)}^1(I; X') \cap L_2(I; X)$ such that[1] $u_t + Au = f$, $u(0) = 0$, where homogeneous initial conditions are chosen only for convenience. The variational formulation then reads

$$\text{find } u \in U : \quad b(u, v) = \langle f, v \rangle \quad \text{for all } v \in V, \tag{1}$$

where $V := L_2(I; X)$, $b(u, v) := \int_0^T \int_\Omega u_t(t, x) v(t, x) \, dx \, dt + \int_0^T a(u(t), v(t)) \, dt$ and $\langle f, v \rangle := \int_0^T \int_\Omega f(t, x) v(t, x) \, dx \, dt$. The well-posedness is ensured by Nečas' conditions, namely boundedness, injectivity and inf-sup condition of $b(\cdot, \cdot)$, [5].

The Wave Equation. Next, we consider an equation of wave type. Here, for $H := L_2(\Omega)$, we view the operator A as a mapping $A : \text{Dom}(A) := \{\phi \in H : A\phi \in$

---

[1] $H_{(0)}^1(I; X') := \{w : I \to X' : w \in H^1(I; X'), w(0) = 0\}$, recall that $H^1(I; X') \hookrightarrow C(\bar{I}; X')$.

$H\} \to H$, or $A : H \to \mathrm{Dom}(A)'$. For $f \in L_2(I; H)$, we seek $u \in L_2(I; H)$ such that $u_{tt} + Au = f$, $u(0) = 0$, $u_t(0) = 0$, where we choose homogeneous initial conditions again only for convenience. In this case, it is not so obvious how to setup a well-posed variational form. It turns out that a very-weak setting is appropriate. We formulate the problem as in (1) by using $U := L_2(I; H)$ as trial and $V := \{v \in L_2(I; H) : v_{tt} + Av \in L_2(I; H), v(T) = v_t(T) = 0\}$ as test space. Then, one can show that (1) is well-posed for $b(u, v) := (u, v_{tt} + Av)_{L_2(I;H)}$ and $\langle f, v \rangle := (f, v)_{L_2(I;H)}$ for fixed $T < \infty$.

3 Petrov-Galerkin Discretizations

In order to determine a numerical approximation to the solution of a variational problem (1), one chooses finite-dimensional trial and test spaces, $U_\delta \subset U$, $V_\delta \subset V$, respectively. For convenience, we assume that their dimension is equal, i.e., $N_\delta := \dim U_\delta = \dim V_\delta$. The Petrov-Galerkin method then reads

$$\text{find } u_\delta \in U_\delta : \quad b(u_\delta, v_\delta) = \langle f, v_\delta \rangle \quad \text{for all } v_\delta \in V_\delta. \tag{2}$$

As opposed to the coercive case, the well-posedness of (2) is not inherited from that of (1). In fact, the spaces U_δ and V_δ need to be appropriately chosen in the sense that the discrete inf-sup (or LBB –Ladyshenskaja-Babuška-Brezzi) condition holds, i.e., there exists an $\beta > 0$ such that $\inf_{u_\delta \in U_\delta} \sup_{v_\delta \in V_\delta} \frac{b(u_\delta, v_\delta)}{\|u_\delta\|_U \|v_\delta\|_V} \geq \beta > 0$, where the crucial point is that $\beta \neq \beta_\delta$. The size of β is also crucial for the error analysis, since it holds that $\|u - u_\delta\|_U \leq \frac{1}{\beta} \inf_{w_\delta \in U_\delta} \|u - w_\delta\|_U$, [15].

The Heat Equation. Starting with the temporal discretization, choose some integer $N_t > 1$ and set $\Delta t := T/N_t$ resulting in a temporal triangulation $\mathcal{T}_{\Delta t}^{\mathrm{time}} \equiv \{t^{k-1} \equiv (k - 1)\Delta t < t \leq k\,\Delta t \equiv t^k, 1 \leq k \leq N_t\}$ in time. Denote by $S_{\Delta t} = \mathrm{span}\{\sigma^1, \ldots, \sigma^{N_t}\}$ piecewise linear finite elements on I, where σ^k is the (interpolatory) hat-function with the nodes t^{k-1}, t^k and t^{k+1} (resp. truncated for $k \in \{0, N_t\}$) and $Q_{\Delta t} = \mathrm{span}\{\tau^1, \ldots, \tau^{N_t}\}$ piecewise constant finite elements, where $\tau^k := \chi_{I^k}$, the characteristic function on the temporal element $I^k := (t^{k-1}, t^k)$. For the spatial discretization, we choose any conformal $X_h = \mathrm{span}\{\phi_1, \ldots, \phi_{N_h}\} \subset X$, e.g. piecewise linear finite elements. Then, we set $U_\delta := S_{\Delta t} \otimes X_h$, $V_\delta = Q_{\Delta t} \otimes X_h$, $\delta = (\Delta t, h)$. It can be shown that this yields LBB. Moreover, for $A = -\Delta$ and choosing the energy norm on X as well as a slightly modified norm on U, one can even prove that $\beta = 1$, [14]. Finally, we remark that this specific discretization coincides with the Crank–Nicolson (CN) scheme if a trapezoidal approximation of the right-hand side temporal integration is used. Hence, we can later compare space-time Petrov-Galerkin numerical schemes with a CN time-stepping scheme.

Finally, we detail the linear system of equations $B_\delta^T u_\delta = f_\delta$, where

$$[B_\delta]_{(k,i),(\ell,j)} = (\dot{\sigma}^k, \tau^\ell)_{L_2(I)} (\phi_i, \phi_j)_{L_2(\Omega)} + (\sigma^k, \tau^\ell)_{L_2(I)} a(\phi_i, \phi_j), \qquad (3)$$

$$[f_\delta]_{(\ell,j)} = (f, \tau^\ell \otimes \phi_j)_{L_2(I;H)}, \qquad (4)$$

which means that we get a tensorproduct structure for the stiffness matrix $B_\delta = D_{\Delta t} \otimes M_h + C_{\Delta t} \otimes A_h$, where the matrices are defined in an obvious manner. The right-hand side is not yet in a tensorproduct structure. However, we can achieve that by determining an approximation

$$f(t, x) \approx \sum_{p=1}^{P} \vartheta_p(t) f_p(x) =: f^P(t, x), \qquad (5)$$

e.g. by the *Empirical Interpolation Method* (EIM), [2]. By choosing P sufficiently large, we can achieve any desired accuracy. Then, we get $[f_\delta^P]_{(\ell,j)} = \sum_{p=1}^{P}(\vartheta_p, \tau^\ell)_{L_2(I)} (f_p, \phi_j)_{L_2(\Omega)}$, i.e., $f_\delta^P = \sum_{p=1}^{P} h_p \otimes g_p$.

The Wave Equation. Constructing a stable pair of trial and test spaces for the wave equation is again a nontrivial task. Following an idea from [3], we first define the test space and construct the trial space in a second step in order to guarantee LBB, which, however, deteriorates with increasing T. Doing so, we set $R_{\Delta t} := \text{span}\{\varrho^1, \ldots, \varrho^{N_t}\} \subset H_T^2(I) := \{\rho \in H^2(I) : \rho(T) = \dot{\rho}(T) = 0\}$, e.g. piecewise quadratic splines on $\mathcal{T}_{\Delta t}^{\text{time}}$. For space, we choose any conformal $Z_h = \text{span}\{\psi_1, \ldots, \psi_{N_h}\} \subset H^2(\Omega) \cap H_0^1(\Omega)$, e.g. piecewise quadratic finite elements. Then, we define $V_\delta := R_{\Delta t} \otimes Z_h$, a tensor product space. The trial space U_δ is constructed by applying the adjoint PDE operator to each test basis function, i.e. $v_{k,i} := \frac{d^2}{dt^2} \varrho^k(t) \psi_i(x) + A(\varrho^k(t)\psi_i(x)) = \ddot{\varrho}^k(t)\psi_i(x) + \varrho^k(t)A\psi_i(x)$. We detail the arising linear system of equations starting with the stiffness matrix

$$\begin{aligned}[B_\delta]_{(k,i),(\ell,j)} &= b(v_{k,i}, \varrho^\ell \otimes \psi_j) = (\ddot{\varrho}^k \otimes \psi_i + \varrho^k \otimes A\psi_i, \ddot{\varrho}^\ell \otimes \psi_j + \varrho^\ell \otimes A\psi_j)_{L_2(I;H)} \\ &= (\ddot{\varrho}^k, \ddot{\varrho}^\ell)_{L_2(I)} (\psi_i, \psi_j)_{L_2(\Omega)} + (\ddot{\varrho}^k, \varrho^\ell)_{L_2(I)} (\psi_i, A\psi_j)_{L_2(\Omega)} \\ &\quad + (\varrho^k, \ddot{\varrho}^\ell)_{L_2(I)} (A\psi_i, \psi_j)_{L_2(\Omega)} + (\varrho^k, \varrho^\ell)_{L_2(I)} (A\psi_i, A\psi_j)_{L_2(\Omega)}, \end{aligned}$$

so that $B_\delta = Q_{\Delta t} \otimes M_h + (D_{\Delta t} + D_{\Delta t}^T) \otimes A_h + M_{\Delta t} \otimes Q_h$, again with obvious definitions of the matrices. For the right-hand side, we perform again an EIM-type approximation $f^P(t, x)$. Then, $[f_\delta^P]_{(\ell,j)} = \sum_{p=1}^{P}(\vartheta_p \otimes f_p, \varrho^\ell \otimes \psi_j)_{L_2(I;H)} = \sum_{p=1}^{P}(\vartheta_p, \varrho^\ell)_{L_2(I)} (f_p, \psi_j)_{L_2(\Omega)}$, so that the right-hand side has the same structure as in the first example. Due to the asymptotic behavior of the inf-sup-constant, we expect stability problems as $\Delta t \to 0$, i.e., $N_t \to \infty$.

4 Efficient Numerical Methods for Tensorproduct Systems

In both cases described above (and in fact also in space-time variational formulations of other PDEs), we obtain a (regular) linear system of the form

$$B_\delta u_\delta = f_\delta \quad \text{with} \quad B_\delta = \sum_{p=1}^{P_B} D_p \otimes A_p, \quad f_\delta = \sum_{\ell=1}^{P_f} h_\ell \otimes q_\ell, \qquad (6)$$

where all involved matrices are sparse and (at least some of) the A_q are s.p.d. Recall that $(D_p \otimes A_p)x = \text{vec}(A_p X D_p^T)$, where vec stacks the columns of a given matrix one after the other, and $x = \text{vec}(X)$. We can thus rewrite the system $B_\delta u_\delta = f_\delta$ in (6) as the linear *matrix* equation $\sum_{p=1}^{P_B} A_p U_\delta D_P^T = \sum_{\ell=1}^{P_f} q_\ell h_\ell^T$, with $u_\delta = \text{vec}(U_\delta)$. Matrix equations are receiving significant attention in the PDE context, due to the possibility of maintaining the structural properties of the discretized problem, while limiting memory consumptions; see [13]. Under certain hypotheses, a large variety of discretization methodologies such as finite differences, isogeometric analysis, spectral (element) methods, certain finite element methods as well as various parametric numerical schemes rely on tensor product spaces; see, e.g., [1, 4, 8, 9]. More recently, all-at-once time discretizations have shown an additional setting where tensor product approximations naturally arise; see, e.g., [11] and references therein. Among the various computational strategies discussed in the literature [13], here we focus on projection methods that reduce the original equation to a similar one, but of much smaller dimension.

Discretized Heat Equation. The problem $B_\delta u_\delta = f_\delta$ stemming from (3,4) yields the following generalized Sylvester equation

$$M_h U_\delta D_{\Delta t} + A_h U_\delta C_{\Delta t} = F_\delta, \qquad \text{with} \quad F_\delta := [g_1, \ldots, g_P][h_1, \ldots, h_P]^T. \quad (7)$$

The spatial stiffness and mass matrices A_h and M_h typically have significantly larger dimensions N_h than the time discretization matrices $D_{\Delta t}$, $N_{\Delta t}$, i.e., $N_t \ll N_h$. We therefore use a reduction method only for the space variables by projecting the problem onto an appropriate space. A matrix Galerkin orthogonality condition is then applied to obtain the solution: given $V_m \in \mathbb{R}^{N_h \times k_m}$, $k_m \ll N_h$, with orthonormal columns, we consider the approximation space range(V_m) and seek $Y_m \in \mathbb{R}^{k_m \times N_t}$ such that $U_{\delta,m} := V_m Y_m \approx U_\delta$ and the residual $R_m := F_\delta - (M_h U_{\delta,m} D_{\Delta t} + A_h U_{\delta,m} C_{\Delta t})$ satisfies the Galerkin condition $R_m \perp$ range(V_m). Imposing this orthogonality yields that $V_m^T R_m = 0$ is equivalent to $V_m^T F_\delta V_m - (V_m^T M_h V_m)Y_m D_{\Delta t} - (V_m^T A_h V_m)Y_m C_{\Delta t} = 0$. The resulting problem is again a generalized Sylvester equation, but of much smaller size, therefore Schur-decomposition oriented methods can cheaply be used, [13, sec.4.2], see [13] for a discussion on projection methods as well as their matrix and convergence properties.

For selecting V_m, let $F = F_1 F_2^T$ with F_1 having full column rank. Given the properties of A_h, M_h, we propose to employ the rational Krylov subspace $\mathcal{RK}_m :=$ range($[F_1, (A_h - \sigma_2 M_h)^{-1} M_h F_1, (A_h - \sigma_3 M_h)^{-1} M_h F_1, \ldots, (A_h - \sigma_m M_h)^{-1} M_h F_1]$), where the shifts σ_s can be determined adaptively while the space is being generated; see [13] for a description and references. The obtained spaces are nested, $\mathcal{RK}_m \subseteq \mathcal{RK}_{m+1}$, therefore the space can be expanded if the approximation is not sufficiently good. To include a residual-based stopping criterion, the residual norm can be computed in a cheap manner, see, e.g., [6, 11] for the technical details.

Discretized Wave problem. The problem $B_\delta u_\delta = f_\delta$ now takes the matrix form

$$M_h U_\delta Q_{\Delta t}^T + A_h U_\delta (D_{\Delta t} + D_{\Delta t}^T) + Q_h U_\delta M_{\Delta t} = F_\delta. \tag{8}$$

This three-term equation cannot be solved directly as before, therefore we opt for using preconditioned GMRES on the vectorized equation. The preconditioner is given by the functional $\mathcal{P} : U \rightarrow M_h U Q_{\Delta t}^T + Q_h U M_{\Delta t}$, corresponding to the discretized forth order operators, and exploits the matrix structure. Hence, at the kth GMRES iteration we solve the generalized Sylvester equation $M_h W Q_{\Delta t}^T + Q_h W M_{\Delta t} = V_k$ where V_k is such that $v_k = \text{vec}(V_k)$ is the previous basis vector. Since in this one-dimensional problem dimensions are limited, this matrix equation is solved by explicitly diagonalizing the pairs (Q_h, M_h) and $(Q_{\Delta t}, M_{\Delta t})$ [13].

5 Numerical Experiments

In this section we show that the numerical solution of the linear system $B_\delta u_\delta = f_\delta$ can largely benefit from the exploitation of its Kronecker sum structure (6). The performance of the all-at-once methods is compared in terms of both computational time and memory requirements. For the heat equation, We also document the performances of CN in terms of computational time. We are not aware of any variant of CN that is able to exploit the low-rank structure of the underlying problem and we thus employ the classical CN scheme. Such implementation leads to large running times and a storage demand that is always equal to $N_t \cdot N_h$ as the full U_δ is allocated.

The tolerance of the final relative residual norm is set to 10^{-8} and in the following tables we also report the number of iterations needed to achieve such accuracy and the numerical rank of the computed solution. All results were obtained with Matlab R2017b on a machine with 2.1 GHz processors and 192GB of RAM.

Example 5.1 (The heat equation) We consider the equation on the cube $\Omega = (-1, 1)^3$ with homogeneous Dirichlet boundary conditions and the time interval $I := (0, 10)$ with initial conditions $u(0, x, y, z) \equiv 0$. The right-hand side is $f(t, x, y, z) := 10 \sin(t) t \cos(\frac{\pi}{2} x) \cos(\frac{\pi}{2} y) \cos(\frac{\pi}{2} z)$ and its discretized version is thus low rank. For discretization in space, linear finite elements were chosen, leading to the discretized generalized Sylvester equation in (7). We compare the performance of the Galerkin projection method based upon rational Krylov spaces described in §4 (denoted RKSM) with that of a low-rank version of preconditioned GMRES (denoted LR-FGMRES-RKSM). For the latter, the preconditioner is chosen as a fixed (five) number of iterations of the rational Krylov Galerkin projection method. Since the preconditioner is a non-linear operator, a flexible variant of GMRES is used. We refer the reader to [11] for more details. The results are displayed in Table 1.

The CN method leads to an excessive workload compared with the all-at-once approaches for all considered values of N_h and N_t, with the computational time growing linearly with the number of time steps N_t. The performance of the other methods is independent of the time discretization, and it only depends on the spatial component of the overall discrete operator. In fact, spatial mesh independence seems to also be achieved.

Table 1: Results for Example 5.1: different values of N_h and N_t. Memory allocations for RKSM and LR-FGMRES+RKSM are given by $\mu_{\text{mem}} \cdot (N_h + N_t)$.

			RKSM			LR-FGMRES+RKSM				CN
N_h	N_t	Its	μ_{mem}	rank(U_δ)	Time (s)	Its	μ_{mem}	rank(U_δ)	Time (s)	Time (s)
41300	100	16	17	10	88.74	4	83	13	143.13	296.16
	300	13	14	9	68.86	4	74	10	130.27	871.38
	500	13	14	9	65.88	4	75	11	134.73	1468.40
347361	100	16	17	9	3144.02	4	71	10	4793.63	13805.09
	300	14	15	9	2673.83	4	78	9	4780.46	41701.10
	500	14	15	9	2699.98	4	80	9	4827.13	70044.52

Example 5.2 (The wave equation) We consider the wave problem with $A = -\Delta$ on $\Omega = (0, 1)$ with homogeneous Dirichlet boundary conditions and $I := (0, 1)$. Setting $f(t, x) := \sin(2\pi x) + 4\pi^2 t^2 \sin(2\pi x)$ yields the analytical solution $u(t, x) = t^2 \sin(2\pi x)$. We choose cubic B-Splines for the discretization in space and time. The discretized problem thus leads to the matrix equation in (8). In Figure 1 we report some our preliminary results. Note, that our above discretization does not yield an equivalent time-stepping scheme with which we could do comparisons.

Fig. 1: Example 5.2. Left: Results for different values of N_h and N_t. Right: Relative residual norm history for some values of N_h and N_t.

		GMRES+LYAP			backslash
N_h	N_t	Its.	rank(U_δ)	Time (s)	Time (s)
256	256	16	13	0.21	1.17
	512	36	35	1.51	2.39
	1024	81	74	20.97	12.61
512	256	26	31	0.61	2.30
	512	40	43	2.64	5.09
	1024	81	74	20.97	12.61
1024	256	50	59	3.55	4.82
	512	68	72	10.12	11.13
	1024	102	92	54.15	24.28
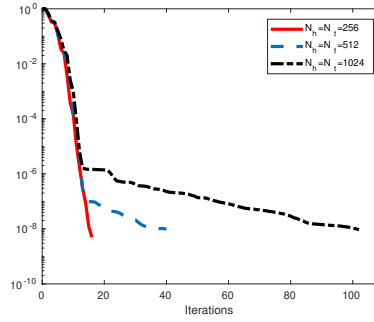
The table on the left shows that the performances of our preconditioned scheme are quite good for small values of N_h and N_t. Indeed, in this case, the preconditioner manages to drastically reduce the number of iterations needed to converge so that GMRES+LYAP turns out to be faster than the Matlab solver `backslash` applied to the solution of the linear system $B_\delta u_\delta = f_\delta$, in spite of the 1D nature (in space) of the problem. However, the effectiveness of the adopted preconditioner worsens by increasing the number of degrees of freedom. This is due to a dramatic increment in the condition number of the coefficient matrices (see the discussion at the end of §3) that causes an abrupt very slow decrement (almost stagnation) in the GMRES residual at the level that seems to be related to the conditioning of the involved matrices, see Fig. 1 (right). As it is, the problem associated with handling this ill-conditioning in the algebraic equation is crucial for the overall solver performance, and will be the topic of future works. Alternatively, one may try to directly address the solution of the multiterm matrix equation (8) as it is done in [12] for certain stochastic PDEs.

Projection-based solvers may be very well-suited for our framework as the solution U_δ turns out to be low-rank. However, such an approach needs the design of ad hoc approximation spaces and we plan to pursue this challenging research direction in the near future.

References

1. M. Bachmayr, A. Cohen, and W. Dahmen. Parametric PDEs: sparse or low-rank approximations? *IMA J. Numer. Anal.*, 38:1661–1708, 2018.
2. M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris*, 339(9):667–672, 2004.
3. J. Brunken, K. Smetana, and K. Urban. (Parametrized) First Order Transport Equations: Realization of Optimally Stable Petrov-Galerkin Methods. *SIAM J. Sci. Comput.*, 41(1):A592–A621, 2019.
4. W. Dahmen, R. DeVore, L. Grasedyck, and E. Süli. Tensor-sparsity of solutions to high-dimensional elliptic partial differential equations. *Found. Comput. Math.*, 16:813–874, 2016.
5. R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 5*. Springer-Verlag, Berlin, 1992. Evolution problems. I.
6. V. Druskin and V. Simoncini. Adaptive rational Krylov subspaces for large-scale dynamical systems. *Systems and Control Letters*, 60:546–560, 2011.
7. S. Glas, A. Mayerhofer, and K. Urban. Two ways to treat time in reduced basis methods. In *Model reduction of parametrized systems*, volume 17 of *MS&A. Model. Simul. Appl.*, pages 1–16. Springer, Cham, 2017.
8. M. Griebel and H. Harbrecht. On the construction of sparse tensor product spaces. *Math. Comp.*, 82(282):975–994, 2013.
9. B. N. Khoromskij and C. Schwab. Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs. *SIAM J. Scient. Comput.*, 33(1):364–385, 2011.
10. J.-L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications. Vol. 2*. Travaux et Recherches Mathématiques, No. 18. Dunod, Paris, 1968.
11. D. Palitta. Matrix equation techniques for certain evolutionary partial differential equations. arXiv math.NA, no. 1908.11851, 2019.
12. C. E. Powell, D. Silvester, and V. Simoncini. An efficient reduced basis solver for stochastic Galerkin matrix equations. *SIAM J. Sci. Comput.*, 39(1):A141–A163, 2017.
13. V. Simoncini. Computational methods for linear matrix equations. *SIAM Review*, 58(3):377–441, 2016.
14. K. Urban and A. T. Patera. An improved error bound for reduced basis approximation of linear parabolic problems. *Math. Comp.*, 83(288):1599–1615, 2014.
15. J. Xu and L. Zikatanov. Some observations on Babuška and Brezzi theories. *Numer. Math.*, 94(1):195–202, 2003.