A COMPARATIVE STUDY OF NONLINEAR GALERKIN FINITE ELEMENT METHODS FOR DISSIPATIVE EVOLUTION PROBLEMS

GUIDO NABH† AND ROLF RANNACHER†

Abstract. Nonlinear Galerkin methods have been proposed in the literature for the long-time solution of semi-linear dissipative evolution problems, e.g., the nonstationary Navier-Stokes equations. In the context of finite element (FE) discretization these methods result in two-level linearization schemes, where on the fine-grid level reduced models for the nonlinearity are used. This paper aims at a better understanding of the mechanisms active in this approach in order to estimate its potential for improving numerical computations. To this end, asymptotic error estimates are given for various nonlinear Galerkin FE methods applied to a one-dimensional model problem containing for example the Burgers and Kuramoto-Sivashinsky equations. This analysis is supplemented by test computations. The result is that, in this context, the modeling aspect of the *nonlinear* Galerkin method is questionable and that its potential for practical computations cannot be predicted on the basis of mere asymptotic error analysis. The dominant mechanism governing its error behavior seems to be more the quality in the approximation of the linear main part of the operator, i.e., its critical eigenvalues, rather than the particular treatment of the lower-order nonlinearity.

Key words. Nonlinear Galerkin method, inertial manifolds, inertial algorithms, dissipative problems, finite element method, Burgers equation, Kuramoto-Sivashinsky equation

AMS subject classifications. 65M60, 58F39, 35Q53, 76M10

1. Introduction.

The solutions of certain semi-linear dissipative evolution equations converge in time asymptotically towards a compact set in function space, named *global attractor*. This global attractor may be rather complicated and far beyond reach of computational methods. In order to simplify the situation, the concept of an *inertial manifold* has been introduced which is supposed to be a smooth, low-dimensional manifold containing the attractor and attracting exponentially all solution trajectories. In the past years a new class of numerical schemes termed *nonlinear Galerkin methods* (NGMs) or *inertial algorithms* has been constructed with the aim to approximate and utilize this inertial manifold. The purpose of this paper is to investigate the potential of this approach for solving dissipative evolution problems using spatial discretization by finite element methods.

Foias, Sell and Temam [9] define an *inertial manifold* as a finite-dimensional algebraic relation between "high" and "low" frequency modes. For a number of differential equations occuring in modeling certain reaction diffusion processes the existence of such inertial manifolds is known. For results in this direction see Foias, Sell and Temam [9], Fabes, Luskin and Sell [5], Sell and You [26], and the literature cited therein. However, for the most challenging case, the Navier-Stokes equations in fluid dynamics, the existence of an inertial manifold is still an open problem, even for two-dimensional spatially periodic flow (for an attempt of proof see Kwak [20]).

Despite of these theoretical problems, a number of mostly heuristic approaches have been developed in the past years for using these ideas in practical computations. In accordance with the basic idea of an inertial manifold, namely the splitting of the state space into low-frequency and high-frequency components, the concept of approximate inertial manifolds was originally used in the context of spectral Galerkin

† Institut für Angewandte Mathematik and Graduiertenkolleg "Modellierung und Wissenschaftliches Rechnen", Universität Heidelberg, INF 293, D-69120 Heidelberg, Germany

approximations. But later also extensions to other discretization methods like general Galerkin as well as finite difference and finite element schemes appeared in the literature; see the papers of Marion and Temam [24], Devulder and Marion [4], Chen and Temam [2], Shen and Temam [27], and the literature cited therein.

A strong impulse for research in this field resulted from theoretical papers as the one of Dedulvier, Marion and Titi [3], in which for a spectral NGM applied to the Navier-Stokes equations with Dirichlet boundary conditions a higher convergence rate was established than what is usually possible for the corresponding *standard Galerkin method* (SGM). This result was taken as justification for the superior potential of the NGM and it even seemed to support the claim that these new methods possess the potential to lead towards mathematically based turbulence modeling (see Foias, Manley and Temam [7] and [8]). This claim, however, was questioned in a paper of Heywood und Rannacher [16] who showed that the increase in order of the NGM reported in [3] is not so much a matter of better nonlinear modeling but rather that of reducing Gibb's phenomenon along the rigid boundary caused by the particular basis function representation. The same effect even occurs for a simple linear equation where the modification in comparison to the SGM is localized to an arbitrarily small strip along the boundary. Subsequently, Jones, Margolin and Titi [19] also came to the same conclusion that a higher convergence rate of the NGM does not seem to be a truely nonlinear effect.

The question whether the NGM approach can at least lead to an effective reduction of computational work, despite to the lacking theoretical support, is not conclusively answered by the existing literature. For example, Gottlieb and Temam [12], and Jauberteau, Rosier and Temam [17] reported of semi-implicit spectral solvers which led to a work reduction of about 50% for the spatially periodic two-dimensional Navier-Stokes equations. This saving is merely due to the reduced number of nonlinear terms in the NGM to be evaluated on the high-frequency level within the time stepping process. This effect cannot be utilized in the presence of rigid boundaries as then the cost for inverting the Stokes part becomes dominant. On the other hand, Garcia-Archilla and Frutos [10] found almost no cost reduction at all in their implicit spectral solver for the two-dimensional Kuramoto-Sivashinsky equation.

For non-spectral Galerkin discretizations only very few results on the behavior of the NGM are known. The problem here is that even the definition of what a NGM looks like is not so clear, as a natural splitting of the discrete solution space into "low" and "high" frequency modes is not obvious. Marion and Temam [23] analysed various NGMs and proved qualitative convergence based on asymptotic stability estimates. Marion and Xu [25] and Ammi and Marion [1] obtained higher order H^1-error estimates for a NGM-type two-level finite element discretization of general semilinear reaction diffusion and the Navier-Stokes equations. Here the strong orthogonality properties of the spectral Galerkin approximation is substituted by the weaker L_2-orthogonality in the finite element subspaces. Laminie, Pascal and Temam [21] considered two-dimensional Burgers and Navier-Stokes type equations and, by computations, came to the conclusion that in using the natural hierarchical bases (not quite the same as H^1-orthogonality in this case) the coarse mesh has to be already very fine in order to realize any positive effect by the NGM. The theoretical results in [25] were later on extended and improved by Walsh [28] who considered finite element discretizations of the two-dimensional Navier-Stokes equations in bounded domains following the analysis of Heywood and Rannacher [15]. He used a mixed L_2/H^1-orthogonal space splitting which appeared natural in the time stepping process and

proved also optimal L_2-error estimates. However, his numerical tests showed that in this situation the NGM became unstable and the correction could even damage the results unless the coarse mesh was already fine enough to represent the essential features of the flow. Only then, higher accuracy might be achieved by the nonlinear correction, but to unacceptably high extra cost.

In this paper we try to contribute to this yet unresolved question in the context of finite element discretization of a one-dimensional model problem containing the Burgers and Kuramoto-Sivashinsky equation with Dirichlet boundary conditions. These problems have a similar typ of nonlinearity, uu_x, as the Navier Stokes equations. For earlier related work see Foias, Nicolaenko, Sell and Temam [6], Jolly, Kevrekidis and Titi [18], and Margolin and Jones [22]. In particular, we investigate the dominant mechanisms active in the NGM which determine the error behavior at realistic mesh sizes. We mainly use L_2-orthogonal splittings as proposed by Goubet [13] and consider some two-grid finite element schemes, including the original NGM but also a scheme in which all nonlinear terms are dropped on the fine-grid level. Even for this "minimal" scheme we obtain optimal-order energy as well as L_2-error estimates which shows that mere asymptotic error analysis is not sufficient for predicting the performance of these new schemes. This result also applies to more complex models like, for example, the Navier-Stokes equations as long as the nonlinearity is of lower order than the linear main part of the operator. We also address the question of whether formal simplifications on the nonlinear terms in the variational problem really lead to a reduction in the numerical cost.

Our theoretical analysis is supplemented by a series of computational tests for the Burgers and Kuramoto-Sivashinsky equations. The performance of the methods considered strongly depends on the choice of the inner product in the space decomposition and the basis functions used. Further, for fine enough coarse mesh all these methods including the "minimal" scheme have almost the same accuracy (within the relevant range) as the standard Galerkin method. This indicates that the dominant mechanism governing the error behavior is more the quality in the approximation of the linear main part of the operator, i.e., its critical eigenvalues, rather than the particular treatment of the lower-order nonlinearity. Only on rather coarse meshes (e.g., 8 unknowns for a fourth-order problem) we find cases in which one of these schemes yield more accurate solutions than the others. However, these quantitative effects are rather unpredictable, as they depend on the particular choice of initial conditions, and cannot be explained by asymptotic error analysis.

The general conclusion of our study is that, for finite element Galerkin discretization, the new nonlinear Galerkin-type methods have only little potential for providing superior accuracy or efficiency in solving evolution problems of the type considered. This may be due to the fact that in all these cases, the linear main parts of the problems do not have a spectrum that naturally splits into low- and high-frequency parts, so that a space decomposition into low- and high-frequency components appears rather arbitrary. Such a *spectral gap property* is inherent for example to certain chemical reaction systems modeling combustion processes, where the use of low-dimensional "inertial manifolds" can actually lead to a significant reduction in the computational work. In the present case, the improvements occasionally observed for the NGM and its variants have not so much to do with modeling of the nonlinearity but rather with the particular approximation properties of the ansatz functions used.

2. Nonlinear Galerkin methods. We consider the following initial boundary problem posed on the unit intervall $\Omega = (0,1)$ with given smooth force f, smooth and compatible initial value u_0, and constants $\nu > 0$ and $\alpha \in \mathbf{R}$:

$$
\begin{aligned}
& u_t + (-1)^n \nu \partial_x^{2n} u + \alpha \partial_x^n u + u u_x \;=\; f \quad \text{in } \Omega, \\
& u_{|t=0} = u_0, \quad \partial_x^i u_{|\partial\Omega} = 0,\ i \in \{0, \ldots, n-1\}.
\end{aligned} \tag{1}
$$

For simplicity the boundary data are assumed as zero. This family of problems includes, e.g., the ordinary Burgers equation, for $n = 1$ and $\alpha = 0$, and the Kuramoto-Sivashinsky equation, for $n = 2$ and $\alpha > 0$.

We introduce the bilinear forms

$$
a(u,v) = a_p(u,v) + a_n(u,v), \quad a_p(u,v) = \nu(\partial_x^n u, \partial_x^n v), \quad a_n(u,v) = \alpha(\partial_x^n u, v),
$$

and the trilinear form

$$
b(u,v,w) = (uv_x, w),
$$

where $(\cdot,\cdot)$ is the inner product of the Lebesgue space $L_2(\Omega)$. The form $a_p(u,v)$ is positive definit, and $a_n(u,v)$ is either positive or negative definit depending on the sign of α. Further, for homogenous Dirichlet boundary conditions, there holds $b(u,u,u) = 0$.

Using this notation, the variational formulation of problem (1) reads as follows: Find a function $U(t) \in V$, such that $U(0) = u_0$, and

$$
(U_t, \phi) + a(U, \phi) + b(U, U, \phi) = (f, \phi) \quad \forall \phi \in V. \tag{2}
$$

Here, $V = H_0^n(\Omega)$ is the usual n-th-order Sobolev space over Ω including the appropriate zero boundary conditions. Below, $H^r = H^r(\Omega)$ will denote the Sobolev spaces of order r, $r \in Z$, and $\|\cdot\|_r$ its usual norm, with $\|\cdot\| = \|\cdot\|_0$ being the L_2-norm. Further, we will use the notation $|\cdot|_\infty$ for the maximum-norm over Ω. By standard arguments, problem (2) has always a unique solution which is smooth for regular data, i.e., has uniformly bounded norms $\|U(t)\|_{2n}$ and $\|U_t(t)\|_{2n}$ over the time interval $[0,T]$.

For discretizing (2), we let $V_h \subset V$ be a finite element subspace consisting of piecewise polynomial functions on a quasi-uniform decomposition of the unit interval into subintervals of maximal length h. Here we restrict us to the case of piecewise polynomials of degree $2n-1$.

Standard Galerkin Method (SGM). The approximation $\bar{u}(t) \in V_h$ to the solution $U(t)$ of (2) is determined by the discrete variational equation termed in the following *Standard Galerkin Method* (SGM):

$$
(\bar{u}_t, \phi) + a(\bar{u}, \phi) + b(\bar{u}, \bar{u}, \phi) = (f, \phi), \quad \forall \phi \in V_h. \tag{3}
$$

Next, we extract a coarse grid space $V_H \subset V_h$ out of the fine grid space V_h. Then, by selecting an appropriate inner product in V_h we induce an orthogonal decomposition of the ansatz space

$$
V_h = V_H + W_h^H,
$$

and a corresponding decomposition of the solution $\bar{u}$ of problem (3)

$$
\bar{u} = \bar{p} + \bar{q}, \qquad \bar{p} \in V_H, \quad \bar{q} \in W_h^H.
$$

Accordingly the SGM (3) can be written in the following equivalent form for the function $\bar{u} = \bar{p} + \bar{q}$:

$$
\begin{aligned}
(\bar{p}_t + \bar{q}_t, \Phi) + a(\bar{p} + \bar{q}, \Phi) + b(\bar{p} + \bar{q}, \bar{p} + \bar{q}, \Phi) &= (f, \Phi), \quad \forall \Phi \in V_H, \\
(\bar{p}_t + \bar{q}_t, \psi) + a(\bar{p} + \bar{q}, \psi) + b(\bar{p} + \bar{q}, \bar{p} + \bar{q}, \psi) &= (f, \psi), \quad \forall \psi \in W_h^H.
\end{aligned} \tag{4}
$$

The new reduced schemes are now derived from (4) as a formal simplification of this system by ignoring certain terms, in order to weaken the coupling between the components in V_H and W_h^H. From the different possibilities for choosing an orthogonal decomposition of V_h, we will consider mainly the case in which the elements of W_h^H are orthogonal in the L_2-sense to the elements of V_H,

$$
(\Phi, \psi) = 0, \qquad \forall \Phi \in V_H, \quad \forall \psi \in W_h^H.
$$

In this case the coupling terms containing the time derivative vanish.

Micro-Scale Linearization Method (MSLM). Ignoring in (4) all nonlinear terms that are cubic and quadratic in elements of W_h^H leads to

$$
\begin{aligned}
(p_t + q_t, \Phi) + a(p + q, \Phi) + b(p + q, p, \Phi) + b(p, q, \Phi) &= (f, \Phi), \quad \forall \Phi \in V_H, \\
(p_t + q_t, \psi) + a(p + q, \psi) + b(p, p, \psi) &= (f, \psi), \quad \forall \psi \in W_h^H,
\end{aligned} \tag{5}
$$

for a modified solution $u = p + q$, $p \in V_H$, $q \in W_h^H$. This approach is subsequently called *micro-scale linearization method*.

Nonlinear Galerkin Method (NGM). Based on concepts of dynamical system theory, one may postulate the existence of an algebraic relation determining the "high-frequency" component q in terms of the "low-frequency" component p as $q = \chi(p)$. In the literature (see, e.g., [23]), such a relation is approximately sought in a form which originates from the second equation in (4) by simply omitting time derivative and nonlinear terms. This leads to the following differential-algebraic system, called *nonlinear Galerkin method*:

$$
\begin{aligned}
(p_t + q_t, \Phi) + a(p + q, \Phi) + b(p + q, p, \Phi) + b(p, q, \Phi) &= (f, \Phi), \quad \forall \Phi \in V_H, \\
a(p + q, \psi) + b(p, p, \psi) &= (f, \psi), \quad \forall \psi \in W_h^H.
\end{aligned} \tag{6}
$$

Compared to the MSLM, in the NGM the time derivative terms $(p_t + q_t, \psi)$ are dropped on the fine-grid level. Normally the term (q_t, Φ), is also ignored, for an easier decoupling within the solution process. We note that this term automatically vanishes in the case of an L_2-orthogonal space decomposition. However, in the general case, an additional "linear" error contribution occurs.

Reduced Nonlinear Coupling Methods (RNCM). Since we want to focus on the treatment of the nonlinearity, we introduce still another scheme called *reduced nonlinear coupling method*. We formally drop in (4) all terms of the nonlinearity which involve linear, quadratic or cubic elements of W_h^H:

$$
\begin{aligned}
(p_t, \Phi) + a(p + q, \Phi) + b(p, p, \Phi) &= (f, \Phi), \quad \forall \Phi \in V_H, \\
(q_t, \psi) + a(p + q, \psi) &= (f, \psi), \quad \forall \psi \in W_h^H.
\end{aligned} \tag{7}
$$

Notice that in this case the equation for the evolution of q is purely linear. Here, the simplification of the model has been carried to the extreme as from the eight

nonlinear terms in (4) only one is kept. Since this is just the one which is also used by a SGM, the computation with this method should be rather cheap. The RNCM can be interpreted as a kind of *artifical viscosity model* adapted to the solution and is of particular interest in the case of an L_2-orthogonal space decomposition. Due to our error analysis presented below, the RNCM shows an optimal-order asymptotic error behavior in H^n as well as in L_2.

We emphasize that the nonlinear terms which are kept in the schemes defined above preserve the property of the SGM, that testing the system by the solution $\Phi = p$ and $\psi = q$ makes the nonlinear terms vanish. This crucial property guarantees a certain degree of stability for all these schemes which allows to prove the existence of solutions. Further, by standard arguments these solutions $u = p + q$ always admit the a priori bound

$$\max_{0\le t\le T} \|u(t)\|_{2n} \le C(T, \nu, \alpha, f, u_0)\,. \tag{8}$$

In the following, overlined quantities like $\bar{p}, \bar{q}, \bar{u}$ are solutions of the SGM, quantities like p, q, u are solutions of the simplified scheme considered (MSLM, NGM or RNCM), and U is the exact solution of the given problem (in our calculations the reference solution).

REMARK 2.1. *When using a spectral Galerkin scheme with ansatz functions which are orthogonal with respect to the L_2 as well as the energy product, there would be no coarse-fine-grid coupling at all in the RNCM. The same would be true with the decomposition ansatz used in [28], where after discretization in time the basis was constructed to be orthogonal with respect to the (time-step dependent) inner product $(u, v) + ka(u, v)$.*

For the error analysis of the schemes defined above we need the usual approximation properties of finite elements,

$$\|v - P_h v\|_i \le c h^{j-i} \|v\|_j, \quad v \in V \cap H^j, \tag{9}$$

for $-2n \le i \le 2n$, and $max\{i, 0\} \le j \le 2n$, where P_h is the L_2-projection into V_h, defined by

$$(v - P_h v, \phi) = 0, \qquad \forall \phi \in V_h\,.$$

Consequently, for an L_2-orthogonal space decomposition, there holds,

$$\|Q_h^H v\|_i \le c H^{n-i} \|\partial_x^n v\|_n, \quad v \in V, \tag{10}$$

for $-2n \le i \le n$, where Q_h^H is the L_2-projection into W_h^H. This is the key for extracting additional orders of convergence in terms of powers of H for the schemes considered. In the case of energy (respectively H^n-) orthogonality, these estimates hold true only for $i \ge 0$. They are available for the common finite element spaces on quasi-uniform meshes.

Here, and below c denotes a generic positive constant which may depend on the solution U of equation (1), on the length of the time interval $[0, T]$, on the viscosity parameter ν, and on the characteristics of the discretization, but it is always independent of the mesh-size parameters h and H.

First, we recall the well-known convergence results for the SGM on the fine grid, namely

$$\|U - \bar{u}\| + h^n \|U - \bar{u}\|_n \leq c h^{2n} . \tag{11}$$

The main theoretical results of this paper are summarized in the following two theorems.

THEOREM 2.1. *For the MSLM, NGM and RNCM with an L_2-orthogonal space decomposition, there holds the H^n-error estimate*

$$\|(u - \bar{u})(t)\|_n \leq c H^{3n-1}, \quad 0 \leq t \leq T, \tag{12}$$

with u being the solution of the reduced scheme considered and $\bar{u}$ that of the SGM.

The H^n-norm (*energy norm*) error estimate (12) is of optimal order compared to what is possible in estimating terms of the type

$$|(pq, \Phi_x)| \leq c H^{3n-1} \|\Phi\|_n, \quad \Phi \in V_H . \tag{13}$$

Combining this result with the error estimate for the SGM, we find that, in the case of an L_2-orthogonal space decomposition, the MSLM, NGM and RNCM admit the estimate

$$\|(u - U)(t)\|_n \leq c_{SGM} h^n + c_{NEW} H^{3n-1}. \tag{14}$$

Hence, for $H = O(h^{\frac{n}{3n-1}-\epsilon})$, these reduced schemes are asymptotically of the same order of approximation with respect to the energy norm, even with the same leading error constant, as the SGM, when all solutions are calculated on the same fine grid.

For the MSLM and RNCM we can supplement the H^n-error estimate of Theorem 2.1 by an improved L_2-error estimate.

THEOREM 2.2. *For the MSLM and RNCM with an L_2-orthogonal space decomposition, there holds the L_2-error estimate*

$$\|(u - \bar{u})(t)\| \leq c\, min\{t^{-1/2} H^{4n-1}, H^{7n/2-1}\}, \quad 0 \leq t \leq T \tag{15}$$

with u being the solution of the reduced scheme considered and $\bar{u}$ that of the SGM.

We note that also the estimate (15) is of optimal order. The blow-up factor $t^{-1/2}$ is due to the fact that for the schemes considered, in general, the initial error $(u-\bar{u})(0)$ is lacking compatibility, i.e., the time derivative $(u - \bar{u})_t$ does not converge to zero as $t \to 0$, in contrast to that of the error $U - \bar{u}$ in the SGM. Combining (15) with the corresponding result for the SGM gives us the estimate

$$\|(u - U)(t)\| \leq c_{SGM} h^{2n} + c_{NEW} H^{4n-1} , \quad 0 < t_0 \leq t \leq T . \tag{16}$$

Hence, in the case of an L_2-orthogonal space decomposition, the MSLM and RNCM are asymptotically of the same accuracy as the SGM also with respect to the L_2-norm.

REMARK 2.2. *For the NGM, we obtain an L_2-norm error estimate of the order $O(H^{3n})$. This is again the best one can achieve in estimating terms of the form*

$$|(q_t, \psi)| \leq c H^{3n} \|\psi\|_n, \quad \psi \in W_h^H .$$

Some further remarks on this point will be given in the proof of Theorem 2.2.

REMARK 2.3. *The results of Theorems 2.1 and 2.2 remain true for the MSLM if in the space decomposition the L_2-orthogonality is replaced by the energy-orthogonality. In this case, additional errors occur for the NGM,*

$$|(p_t, \psi)| \le cH^n \|\psi\|_n, \quad \psi \in W_h^H ,$$

as well as for the RNCM,

$$|(pp_x, \psi)| \le cH^n \|\psi\|_n, \quad \psi \in W_h^H ,$$

which limit the order of the error in (14) to $O(H^n)$, and therefore effect the leading error constant. Further, the dropping of terms like

$$|(q_t, \Phi)| + |((pq)_x, \Phi)| \le cH^{2n} \|\Phi\|_n, \quad \Phi \in V_H ,$$

leads to the same error constant in (14), but does not allow an improvement as in (16), which is confirmed by our numerical results.

3. Proof of Theorem 2.1. We will provide the detailed proof of the H^n-error estimate in Theorem 2.1 for the RNCM only. This argument carries over to the MSLM, as this scheme is by construction certainly closer to the SGM than the RNCM. The corresponding result for the NGM follows by a simple modification of the argument to be presented. In order to make the resulting powers in H more plausible, we will omit the usage of the abstract formulation with bi- and trilinear forms. Let

$$P := P_H, \qquad Q := Q_h^H = I - P_H,$$

denote the L_2-projections onto V_H and its complement W_h^H, respectively, and let the initial values be taken as $p(\cdot, 0) = P\bar{u}_0$ and $q(\cdot, 0) = Q\bar{u}_0$, where again $\bar{u}$ is the solution of the SGM with initial value $\bar{u}_0 = P_h U_0$. The difference between the solutions of the SGM and the MSLM, RNCM, or NGM, is abbreviated by $e = u - \bar{u}$. Due to our assumptions, there holds

$$\|Pe\|_i + \|Qe\|_i \le c\|e\|_i,$$

for $i = 0, \ldots, n$, which will extensively be used in the following without explicit mention. Further, for estimating the nonlinear terms, we will use the estimate

$$|(uv, w)| \le |u|_\infty \|v\| \|w\| \le \|u\|_1 \|v\| \|w\|,$$

which is valid in one space dimension.

In the first step, we subtract the equations of the SGM (4) from those of the RNCM (7) resulting in

$$\begin{aligned} (e_t, \phi) + \nu(\partial_x^n e, \partial_x^n \phi) \quad &= \quad -\alpha(\partial_x^n e, \phi) + (\bar{p}\bar{p}_x - pp_x, P\phi) \\ &\quad + ((\bar{p}\bar{q})_x, \phi) + (\bar{q}\bar{q}_x, \phi) + (\bar{p}\bar{p}_x, Q\phi), \quad \forall \phi \in V_h. \end{aligned} \tag{17}$$

Note that for the MSLM the additional terms $((pq)_x, P\phi)$ and $(pp_x, Q\phi)$ occur, and in the case of the NGM another term

$$(q_t, Q\phi) \le cH^{6n} + \epsilon \|\partial_x^n \phi\|^2,$$

which does not effect the following lemma which provides a first, though still sub-optimal error estimate.

LEMMA 3.1. *Under the assumptions of Theorem 2.1 there holds*

$$\|e\|^2 + \nu \int_0^t \|\partial_x^n e\|^2 ds \le cH^{6n-2}, \quad 0 \le t \le T. \tag{18}$$

Proof. Taking $\phi = e$ in (17), we get

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|e\|^2 + \nu\|\partial_x^n e\|^2 &= -\alpha(\partial_x^n e, e) + (\bar{p}\bar{p}_x - pp_x, Pe) \\ &\quad +((\bar{p}\bar{q})_x, e) + (\bar{q}\bar{q}_x, e) + (\bar{p}\bar{p}_x, Qe). \end{aligned} \tag{19}$$

We will estimate the different terms of the right hand side separately. For gaining powers of H, we will make frequent use of the inequalities

$$\|\bar{q}\| \le cH^{2n}, \qquad \|Qe\|_1 \le cH^{n-1}\|Qe\|_n \le cH^{n-1}\|\partial_x^n e\|,$$

as well as of the Hölder and Young inequalities. Further, the a priori bound (8) for the approximate solutions u and $\bar{u}$ will be used without extra mention.

We break the first term in (19) into

$$|(\partial_x^n e, e)| \le \epsilon\|\partial_x^n e\|^2 + c_\epsilon\|e\|^2.$$

For the second term, we have

$$|(pp_x - \bar{p}\bar{p}_x, Pe)| = |((p-\bar{p})p_x + \bar{p}(p-\bar{p})_x, Pe)| \le \epsilon\|\partial_x^n e\|^2 + c_\epsilon\|e\|^2.$$

The third term is a crucial one, as it determines the achievable order in the error estimate. It can be estimated, after integration by parts, by

$$|((\bar{p}\bar{q})_x, e)| \le \|\bar{q}\|_{1-n}\|\bar{p}e_x\|_{n-1} \le \epsilon\|\partial_x^n e\|^2 + c_\epsilon H^{6n-2}.$$

We note that, in the case of an energy-orthogonal decomposition, the order of this term is limited to $O(H^{4n})$. For the fourth term, we have

$$|(\bar{q}\bar{q}_x, e)| \le |e|_\infty\|\bar{q}\|\|\bar{q}_x\| \le \epsilon\|\partial_x^n e\|^2 + c_\epsilon H^{8n-2}.$$

For estimating the last term, we replace $\bar{p}$ by $\bar{u} - \bar{q}$ and obtain

$$(\bar{p}\bar{p}_x, Qe) = (\bar{u}\bar{u}_x, Qe) - ((\bar{u}\bar{q})_x, Qe) + (\bar{q}\bar{q}_x, Qe). \tag{20}$$

The last two terms on the right can be estimated analogously to the third term above by $\epsilon\|\partial_x^n e\|^2 + c_\epsilon H^{6n-2}$. The first term is rewritten as

$$(\bar{u}\bar{u}_x, Qe) = (\bar{u}(\bar{u}-U)_x, Qe) + ((\bar{u}-U)U_x, Qe) + (UU_x, Qe),$$

where the first two terms can be estimated, using the known error estimate $\|\bar{u}_x - U_x\| \le ch^{2n-1}$ for the SGM, by

$$|(\bar{u}(\bar{u}-U)_x, Qe) + ((\bar{u}-U)U_x, Qe)| \le \epsilon\|\partial_x^n e\|^2 + c_\epsilon H^{2n}h^{4n-2}.$$

Finally, for the last term (UU_x, Qe) we obtain

$$|(UU_x, Qe)| \leq \epsilon \|\partial_x^n e\|^2 + c_\epsilon H^{6n-2}.$$

This implies for the last term in (17) that

$$|(\bar{p}\bar{p}_x, Qe)| \leq \epsilon \|\partial_x^n e\|^2 + c_\epsilon H^{6n-2}.$$

We note that in the case of H^n-orthogonality, this term would only be of order $O(H^{2n})$. Combining the above estimates and chosing a suitable ϵ, we arrive at

$$\frac{d}{dt}\|e\|^2 + \nu \|\partial_x^n e\|^2 \leq c\|e\|^2 + cH^{6n-2},$$

and use of Gronwall's inequality concludes the proof. □

In the next step, we will derive the desired estimate for $\|\partial_x^n e\|$. Taking $\phi = e_t$ in (17) we get

$$\begin{aligned} \frac{\nu}{2}\frac{d}{dt}\|\partial_x^n e\|^2 + \|e_t\|^2 &= -\alpha(\partial_x^n e, e_t) + (\bar{p}\bar{p}_x - pp_x, Pe_t) \\ &\quad + (\bar{q}\bar{q}_x, e_t) + ((\bar{p}\bar{q})_x, e_t) + (\bar{p}\bar{p}_x, Qe_t). \end{aligned} \tag{21}$$

The different terms of the right hand side will be estimated similarly as in the previous proof. The first two terms are treated completely analogously the only difference being the change from e and Pe to e_t and Pe_t, respectivly,

$$|\alpha(\partial_x^n e, e_t) + (pp_x - \bar{p}\bar{p}_x, Pe_t)| \leq \epsilon \|e_t\|^2 + c_\epsilon \|\partial_x^n e\|^2. \tag{22}$$

For the three remaining terms in (21) we have to change the argument a bit. For estimating $(\bar{q}\bar{q}_x, e_t)$, we write

$$(\bar{q}\bar{q}_x, e_t) = \frac{d}{dt}(\bar{q}\bar{q}_x, e) - ((\bar{q}\bar{q}_x)_t, e) \leq \frac{d}{dt}(\bar{q}\bar{q}_x, e) + \epsilon \|\partial_x^n e\|^2 + c_\epsilon H^{8n-2}.$$

This leads us to

$$(\bar{q}\bar{q}_x, e_t) \leq \frac{d}{dt}(\bar{q}\bar{q}_x, e) + \epsilon \|\partial_x^n e\|^2 + c_\epsilon H^{8n-2}.$$

For the fourth term in (21) we proceed in the same way obtaining

$$((\bar{p}\bar{q})_x, e_t) = -\frac{d}{dt}(\bar{p}\bar{q}, e_x) + ((\bar{p}\bar{q})_t, e_x) \leq -\frac{d}{dt}(\bar{p}\bar{q}, e_x) + \epsilon \|\partial_x^n e\|^2 + c_\epsilon H^{6n-2}.$$

We note, that at this point again energy-orthogonality would limit the achievable order for the middle term to $O(H^{4n})$. For the last term in (21) we can argue analogously as in the previous proof. Combining the above estimates while choosing ϵ appropriately, we arrive at

$$\frac{\nu}{2}\frac{d}{dt}\|\partial_x^n e\|^2 + \|e_t\|^2 \leq -\frac{d}{dt}\{(\bar{p}\bar{q}, e_x) - (\bar{q}\bar{q}_x, e)\} + c\|\partial_x^n e\|^2 + cH^{6n-2}.$$

Integrating this over $[0, t]$ and observing that $\partial_x^i e|_{t=0} = 0$, $i \in \{0, \ldots, n-1\}$, we finally conclude by Gronwall's lemma and the results of the previous proof that

$$\nu \|\partial_x^n e\|^2 + \int_0^t \|e_s\|^2 ds \leq cH^{6n-2}. \tag{23}$$

This completes the proof of the theorem for the RNCM, and in view of the above remarks also for the MSLM and NGM.

4. Proof of Theorem 2.2. The proof parallels very much that of Theorem 2.1. Therefore, only the key steps will be described and the technical details left to the reader. For deriving the L_2-error estimates, we use a simple linearization argument. Let $\hat{u} = \hat{p} + \hat{q}$, with $\hat{p} \in V_H$, $\hat{q} \in W_h^H$, be the solution of the linear system

$$(24)\qquad \begin{aligned} (\hat{p}_t, \Phi) + \nu(\partial_x^n(\hat{p}+\hat{q}), \partial_x^n \Phi) &= (f - \bar{p}\bar{p}_x - \alpha\partial_x^n \bar{u}, \Phi), \quad \forall \Phi \in V_H, \\ (\hat{q}_t, \psi) + \nu(\partial_x^n(\hat{p}+\hat{q}), \partial_x^n \psi) &= (f - \alpha\partial_x^n \bar{u}, \psi), \quad \forall \psi \in W_h^H, \end{aligned}$$

with initial values $\hat{p}_{|t=0} = P\bar{u}_{|t=0}$ and $\hat{q}_{|t=0} = Q\bar{u}_{|t=0}$.

We define $\hat{e} = \bar{u} - \hat{u}$ as the difference of the solution of the SGM to that of the auxiliar problem (24). Then, subtracting (24) from (4) gives us

$$(25)\qquad (\hat{e}_t, \phi) + \nu(\partial_x^n \hat{e}, \partial_x^n \phi) = -(\bar{q}\bar{q}_x, \phi) - (\bar{p}\bar{p}_x, Q\phi) - ((\bar{p}\bar{q})_x, \phi), \quad \forall \phi \in V_h,$$

with initial value $\hat{e}_{|t=0} = 0$. Using this notation, the following lemma can be proven by arguments already used in the proof of Theorem 2.1.

LEMMA 4.1. *Under the assumptions of Theorem 2.2 there holds,*

$$(26)\qquad \|\hat{e}\|^2 + \nu \int_0^t \|\partial_x^n \hat{e}\|^2 ds \le cH^{6n-2}, \quad 0 \le t \le T,$$

$$(27)\qquad \nu\|\partial_x^n \hat{e}\|^2 + \int_0^t \|\hat{e}_s\|^2 ds \le cH^{6n-2}, \quad 0 \le t \le T.$$

In the following proof we will use a (linear) parabolic duality argument. For fixed $t \in (0, T]$, let $z = z(x, s)$ be the solution of the auxiliary problem

$$(28)\qquad z_s + (-1)^{n+1}\nu\partial_x^{2n} z = g, \quad z_{|s=t} = 0, \quad \partial_x^i z|_{\partial\Omega} = 0, \quad i = 0, ..., n-1.$$

For this problem the usual a priori estimates hold true which are summarized in the following lemma.

LEMMA 4.2. *Under the assumptions of Theorem 2.2 there holds*

$$(29)\qquad \|z(0)\|_n^2 + \int_0^t \{\|z\|_{2n}^2 + \|z_s\|^2\}\, ds \le c \int_0^t \|g\|^2\, ds, \quad 0 \le t \le T.$$

In the next step we prove an optimal-order estimate for the error $\hat{e} = \bar{u} - \hat{u}$.

LEMMA 4.3. *Under the assumptions of Theorem 2.2 there holds*

$$(30)\qquad \int_0^t \|\hat{e}\|^2 ds \le cH^{8n-2}, \quad 0 \le t \le T.$$

Proof. For fixed $t \in (0, T]$, let z be the solution of the auxiliary problem (28) with right hand side $g = \hat{e}$. Multiplying (28) by $\hat{e}$ gives us

$$\|\hat{e}\|^2 = (z_s, \hat{e}) - \nu(\partial_x^n z, \partial_x^n \hat{e}) = \frac{d}{ds}(z, \hat{e}) - (\hat{e}_s, z) - \nu(\partial_x^n \hat{e}, \partial_x^n z).$$

Observing that $QPz = 0$, we get for $\phi = Pz$ in (25)

$$(\hat{e}_t, Pz) + \nu(\partial_x^n \hat{e}, \partial_x^n Pz) + ((\bar{p}\bar{q})_x, Pz) + (\bar{q}\bar{q}_x, Pz) = 0.$$

Adding this equation to the previous one, we get (observe $Q = I - P$)

$$\|\hat{e}\|^2 = \frac{d}{ds}(z, \hat{e}) - (\hat{e}_s, Qz) - \nu(\partial_x^n \hat{e}, \partial_x^n Qz) + ((\bar{p}\bar{q})_x, Pz) + (\bar{q}\bar{q}_x, Pz).$$

For the terms on the right hand side, we obtain by standard arguments the estimates

$$\begin{aligned}
|(\hat{e}_s, Qz)| &\leq \epsilon \|z\|_{2n}^2 + c_\epsilon H^{4n} \|\hat{e}_s\|^2, \\
|(\partial_x^n \hat{e}, \partial_x^n Qz)| &\leq \epsilon \|z\|_{2n}^2 + c_\epsilon H^{2n} \|\partial_x^n \hat{e}\|^2, \\
|(\bar{q}, \bar{p}(Pz)_x)| &\leq c\|\bar{q}\|_{1-2n}\|z\|_{2n} \leq \epsilon \|z\|_{2n}^2 + c_\epsilon H^{8n-2}, \\
|(\bar{q}\bar{q}_x, Pz)| &\leq \|\bar{q}\|\|\bar{q}_x\|\|(Pz)_x\| \leq \epsilon \|z\|_{2n}^2 + c_\epsilon H^{8n-2}.
\end{aligned}$$

We note that for energy-orthogonality, the order achievable in the third estimate would be only $O(H^{4n})$. Combining the above estimates gives us

$$\|\hat{e}\|^2 \leq \frac{d}{ds}(z, \hat{e}) + \epsilon \|z\|_{2n}^2 + c_\epsilon H^{2n}(\|\partial_x^n \hat{e}\|^2 + \|\partial_x^n e\|^2) + c_\epsilon H^{8n-2} + c_\epsilon H^{4n} \|\hat{e}_s\|^2.$$

Integrating this over $[0, t]$ and using of the results of the previous lemma, we get

$$\int_0^t \|\hat{e}\|^2 \, ds \leq \epsilon \int_0^t \|z\|_{2n}^2 \, ds + c_\epsilon H^{8n-2}.$$

This eventually leads us to the desired result by using the a priori estimate (29) for the dual solution z. □

Next, we estimate the error $\tilde{e} = u - \hat{u}$.

LEMMA 4.4. *Under the assumptions of Theorem 2.2 there holds*

$$\|\tilde{e}\|^2 + \int_0^t \|\partial_x^n \tilde{e}\|^2 \, ds \leq c H^{8n-2}, \quad 0 \leq t \leq T. \tag{31}$$

Proof. We subtract the equations for $\hat{u}$ from that for u and set $\phi = \tilde{e}$, to obtain

$$\frac{1}{2}\frac{d}{dt}\|\tilde{e}\|^2 + \nu\|\partial_x^n \tilde{e}\|^2 = (\bar{p}\bar{p}_x - pp_x, P\tilde{e}) + \alpha(\partial_x^n e, \tilde{e}).$$

For the terms on the right hand side we have

$$|(pp_x - \bar{p}\bar{p}_x, P\tilde{e}) + (\partial_x^n e, \tilde{e})| \leq \epsilon \|\partial_x^n \tilde{e}\|^2 + c_\epsilon \|\hat{e}\|^2 + c_\epsilon \|\tilde{e}\|^2,$$

and, consequently,

$$\frac{d}{dt}\|\tilde{e}\|^2 + \nu\|\partial_x^n \tilde{e}\|^2 \leq c\|\tilde{e}\|^2 + c\|\hat{e}\|^2.$$

Integrating this with respect to time and using (30) yields the desired result. □

We note that, if also the time derivative term is dropped as in the NGM, an additional error of the order $O(H^{6n})$ occurs.

From the previous lemmata it follows that

$$\int_0^t \|e\|^2 ds \le 2\int_0^t (\|\hat{e}\|^2 + \|\tilde{e}\|^2) ds \le cH^{8n-2}. \tag{32}$$

Next, we want to convert this integral error estimate into a uniform one in time, for which we need appropriate bounds for the time derivative e_t. Since we cannot expect that $\lim_{t\to 0}\|e_t\| = 0$, a time-weight t has to be introduced, so that $\lim_{t\to 0} t\|e_t\| = 0$.

We begin by differentiating equation (17) with respect to time,

$$\begin{aligned}(e_{tt},\phi) + \nu(\partial_x^n e_t, \partial_x^n \phi) + \alpha(\partial_x^n e_t, \phi) - ((\bar{q}\bar{q}_x)_t, \phi) \\ -((\bar{p}\bar{p}_x - pp_x)_t, P\phi) - ((\bar{p}\bar{q})_{xt}, P\phi) \\ -((\bar{p}\bar{p}_x)_t, Q\phi) - ((\bar{p}\bar{q})_{xt}, Q\phi) \quad = \quad 0, \quad \forall \phi \in V_h.\end{aligned} \tag{33}$$

LEMMA 4.5. *Under the assumptions of Theorem 2.2 there holds*

$$t\|e_t\|^2 + \nu\int_0^t s\|\partial_x^n e_s\|^2\, ds \le cH^{6n-2}, \quad 0 \le t \le T. \tag{34}$$

Proof. Taking $\phi = e_t$ in (33), we get

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|e_t\|^2 + \nu\|\partial_x^n e_t\|^2 \quad = \quad & -\alpha(\partial_x^n e_t, e_t) + ((\bar{q}\bar{q}_x)_t, e_t) + ((\bar{p}\bar{p}_x - pp_x)_t, Pe_t) \\ & +((\bar{p}\bar{q})_{xt}, Pe_t) - ((\bar{p}\bar{p}_x)_t, Qe_t) + ((\bar{p}\bar{q})_{xt}, Qe_t).\end{aligned}$$

The argument for estimating the several terms on the right hand side is analogous to that already used in the proof for the estimate of $\|e\|$. We suppress the details. Making use of the previously proven error estimates, we arrive at

$$\frac{d}{dt}\|e_t\|^2 + \nu\|\partial_x^n e_t\|^2 \le c\|e_t\|^2 + cH^{6n-2}.$$

Multiplying this through by t, we obtain

$$\frac{d}{dt}\{t\|e_t\|^2\} + t\nu\|\partial_x^n e_t\|^2 \le ct\|e_t\|^2 + \|e_t\|^2 + ctH^{6n-2},$$

from which, by integration with respect to time, we conclude the desired result. □

Next, we differentiate the equation derived for the auxiliary problem (25) with respect to time, to obtain

$$(\hat{e}_{tt},\phi) + \nu(\partial_x^n \hat{e}_t, \partial_x^n \phi) + ((\bar{q}\bar{q}_x)_t, \phi) + ((\bar{p}\bar{p}_x)_t, Q\phi) + ((\bar{p}\bar{q})_{xt}, \phi) = 0, \tag{35}$$

for all $\phi \in V_h$. Then, analogously as in the previous proof, we obtain the following result.

LEMMA 4.6. *Under the assumptions of Theorem 2.2 there holds*

$$t\|\hat{e}_t\|^2 + \nu\int_0^t s\|\partial_x^n \hat{e}_t\|^2\, ds \le cH^{6n-2}, \quad 0 \le t \le T. \tag{36}$$

In the next step, we will derive an optimal-order integral estimate for $t\|\hat{e}_t\|$.

LEMMA 4.7. *Under the assumptions of Theorem 2.2 there holds*

$$\int_0^t s^2 \|\hat{e}_t\|^2 \, ds \leq cH^{8n-2}, \quad 0 \leq t \leq T. \tag{37}$$

Proof. We use again the auxiliary problem (28), this time with the right hand side $g = s\hat{e}_s$. Multiplying (28) by $\hat{e}_s$ and integrating the result with respect to time gives us

$$s\|\hat{e}_s\|^2 = (z_s, \hat{e}_s) - \nu(\partial_x^n z, \partial_x^n \hat{e}_s) = \frac{d}{ds}(z, \hat{e}_s) - (\hat{e}_{ss}, z) - \nu(\partial_x^n \hat{e}_s, \partial_x^n z).$$

Taking $\phi = Pz$ in (35), and observing that $QPz = 0$, we get

$$(\hat{e}_{ss}, Pz) + \nu(\partial_x^n \hat{e}_s, \partial_x^n Pz) + ((\bar{q}\bar{q}_x)_s, Pz) + ((\bar{p}\bar{q})_{xs}, Pz) \;=\; 0\,.$$

Adding this to the preceding identity results in

$$s\|\hat{e}_s\|^2 \;=\; \frac{d}{ds}(z, \hat{e}_s) - (\hat{e}_{ss}, Qz) - (\partial_x^n \hat{e}_s, \partial_x^n Qz) + ((\bar{p}\bar{q})_{xs}, Pz) + ((\bar{q}\bar{q}_x)_s, Pz).$$

Estimating the right hand side as before and observing that

$$(\hat{e}_s, Qz_s) = (Q\hat{e}_s, z_s),$$

we can proceed as previously in the proof of Lemma 4.1, to obtain

$$s\|\hat{e}_s\|^2 \leq \frac{d}{ds}(\hat{e}_s, Pz) + c_\epsilon H^{2n}\{\|\partial_x^n \hat{e}_s\|^2 + \|\partial_x^n e_s\|^2\} + c_\epsilon H^{8n-2} + \epsilon\{\|z\|_{2n}^2 + \|z_s\|^2\}\,.$$

We multiply this estimate through by s. The first term on the right hand side is rewritten as

$$s\frac{d}{ds}(\hat{e}_s, Pz) = \frac{d}{ds}\{s(\hat{e}_s, Pz)\} - \frac{d}{ds}(\hat{e}, Pz) + (\hat{e}, Pz_s),$$

which after integration becomes

$$\int_0^t s\frac{d}{ds}(\hat{e}_s, Pz)\, ds = \int_0^t (\hat{e}, Pz_s)\, ds \leq c_\epsilon \int_0^t \|\hat{e}\|^2 \, ds + \epsilon \int_0^t \|z_s\|^2 \, ds\,,$$

where we have used that $\hat{e}|_{s=0} = 0$, $Pz|_t = 0$ and $\lim_{s\to 0}\{s\|\hat{e}_s\|\} = 0$. Then, in view of the above estimates for $\hat{e}_s$ and e_s, and the a priori bounds for z, we conclude that

$$\int_0^t s^2\|\hat{e}_s\|^2 \, ds \leq c_\epsilon H^{8n-2} + \epsilon \int_0^t s^2\|\hat{e}_s\|^2 \, ds\,.$$

From this we obtain the desired result by choosing a suitable ϵ. □

Combining the above results we can now complete the proof. We write

$$t\|\hat{e}\|^2 \;=\; \int_0^t \frac{d}{ds}\{s\|\hat{e}\|^2\}\, ds \;\leq\; 2\int_0^t \|\hat{e}\|^2 \, ds + \int_0^t s^2\|\hat{e}_s\|^2 \, ds \;\leq\; cH^{8n-2}\,.$$

This together with the already proven estimate $\|\tilde{e}\| \leq cH^{4n-1}$, leads us to the desired result for $e = \tilde{e} - \hat{e}$. We note that, for the NGM we have $\|\tilde{e}\| \leq cH^{6n}$ and therefore only get $t^{\frac{1}{2}}\|e\| \leq cH^{3n}$.

5. Numerical results. In the following we present the results of some test computations for the one-dimensional Burgers and Kuramoto-Sivashinsky equations. Besides demonstrating the sharpness of our theoretical results, this is mainly intended to clearify the mechanisms active in the various schemes considered. Although, the Burgers equation has only a very simple solution dynamics, it has been chosen to study the asymptotic convergence behavior of the schemes and to check their quantitative approximation properties. The solution of the Kuramoto-Sivashinsky equation has a more complex behavior, apparently time periodic for the parameters chosen. This allows us to study the schemes' ability to capture the essential features of this dynamics depending on the discretization parameters.

5.1. The Burgers test. We consider the Burgers equation

$$U_t - \nu U_{xx} + UU_x = f,$$

on the unit intervall $\Omega = (0,1)$, with boundary conditions $U_{|\partial\Omega} = 0$, and with viscosity parameter $\nu = \pi^{-2}$. The exact (smooth) solution is prescribed as

$$U(x,t) = [1 - \exp(-t)]\sin(\pi x).$$

We integrate from $t = 0$ until $t = 1$ which induces the initial condition $U_0 = 0$ and the right hand side $f(x,t) = \sin(\pi x) + (1 - \exp(-t))^2 \cos(\pi x)\sin(\pi x)$.

The spatial discretization is restricted to the simplest case of linear finite elements where the space decomposition $V_h = V_H + W_h^H$ is realized for various mesh size ratios $H/h = 2, 4, 8$. For time integration we use a SDIRK4-method (see [14]) with automatic step-size control. The error of the time integration is in all cases less than the number of digits given in the tables.

In the following we display the L_2- and H^1-errors of the fine-grid solution $u = p+q$, the coarse grid component p, and the corresponding convergence rates. All figures are related to $T = 1$, while N and n are the number of intervals of the coarse and fine grids, respectively.

n	$\|\|U - u\|\|$		$\|\|U_x - u_x\|\|$	
4	1.458e-2	-	3.187e-1	-
8	3.533e-3	4.12	1.592e-1	2.00
16	8.763e-4	4.03	7.960e-2	2.00
32	2.186e-4	4.01	3.980e-2	2.00
64	5.464e-5	4.00	1.990e-2	2.00
128	1.366e-5	4.00	9.949e-3	2.00
256	3.414e-6	4.00	4.975e-3	2.00

TABLE 1
Burgers Test, SGM

For the SGM, we clearly see the expected convergence of second order in L_2 and of first order in H^1. In this case there is no a priori splitting into p and q, and n is equal to N.

For the NGM, we also observe the expected higher convergence rates in the L_2-norm, $c_{SGM}h^2 + c_{NGM}H^3$, especially for higher grid-size ratios H/h. It is interesting that for a grid-size ratio of $H/h = 2$, the solution $u = p + q$ is nearly of the same accuracy as that of the fine-grid SGM (calculated with mesh size h). However, the solution $u = p$ is only slightly better than the corresponding coarse-grid solution of

n	N	$\|\|U-u\|\|$		$\|\|U_x-u_x\|\|$		$\|\|U-p\|\|$		$\|\|U_x-p_x\|\|$	
4	2	1.797e-2	-	3.258e-1	-	5.472e-2	-	6.557e-1	-
8	4	3.537e-3	5.08	1.598e-1	2.04	1.128e-2	4.85	3.216e-1	2.04
16	8	8.709e-4	4.06	7.961e-2	2.01	2.681e-3	4.21	1.596e-1	2.01
32	16	2.182e-4	3.99	3.980e-2	2.00	6.618e-4	4.05	7.965e-2	2.00
64	32	5.461e-5	4.00	1.990e-2	2.00	1.649e-4	4.01	3.980e-2	2.00
128	64	1.366e-5	4.00	9.949e-3	2.00	4.120e-5	4.00	1.990e-2	2.00
256	128	3.414e-6	4.00	4.975e-3	2.00	1.030e-5	4.00	9.949e-3	2.00
8	2	2.251e-2	-	2.677e-1	-	5.426e-2	-	6.594e-1	-
16	4	2.278e-3	9.88	9.383e-2	2.85	1.107e-2	4.90	3.227e-1	2.04
32	8	3.111e-4	7.32	4.152e-2	2.26	2.621e-3	4.05	1.597e-1	2.02
64	16	6.068e-5	5.13	2.011e-2	2.07	6.462e-4	4.01	7.966e-2	2.00
128	32	1.404e-5	4.32	9.976e-3	2.02	1.610e-4	4.00	3.981e-2	2.00
256	64	3.439e-6	4.08	4.978e-3	2.00	4.022e-5	4.00	1.990e-2	2.00
16	2	2.492e-2	-	2.623e-1	-	5.419e-2	-	6.601e-1	-
32	4	2.452e-3	10.2	6.920e-2	3.79	1.106e-2	4.90	3.230e-1	2.04
64	8	2.713e-4	9.04	2.407e-2	2.87	2.617e-3	4.23	1.598e-1	2.02
128	16	3.481e-5	7.80	1.050e-2	2.29	6.452e-4	4.06	7.967e-2	2.01
256	32	5.233e-6	6.65	5.044e-3	2.08	1.608e-4	4.01	3.981e-2	2.00

TABLE 2
Burgers Test, NGM (L_2-orthogonal)

the SGM (calculated with mesh-size H). Hence, at least in this test case, the "fine-mesh" component q should not be neglected in the approximate solution $u = p + q$, in contrast to what is sometimes suggested in the literature.

n	N	$\|\|U-u\|\|$		$\|\|U_x-u_x\|\|$		$\|\|U-p\|\|$		$\|\|U_x-p_x\|\|$	
4	2	1.781e-2	-	3.212e-1	-	6.502e-2	-	6.538e-1	-
8	4	3.611e-3	4.93	1.597e-1	2.01	1.130e-2	5.75	3.215e-1	2.03
16	8	8.762e-4	4.12	7.961e-2	2.01	2.682e-3	4.21	1.596e-1	2.01
32	16	2.186e-4	4.01	3.980e-2	2.00	6.619e-4	4.05	7.965e-2	2.00
64	32	5.463e-5	4.00	1.990e-2	2.00	1.650e-4	4.01	3.980e-2	2.00
128	64	1.366e-5	4.00	9.949e-3	2.00	4.120e-5	4.00	1.990e-2	2.00
256	128	3.414e-6	4.00	4.975e-3	2.00	1.030e-5	4.00	9.949e-3	2.00
8	2	1.941e-2	-	2.494e.1	-	5.434e-2	-	6.586e-1	
16	4	2.252e-3	8.62	9.328e-2	2.67	1.108e-2	4.90	3.226e-1	2.04
32	8	3.141e-4	7.17	4.151e-2	2.25	2.621e-3	4.23	1.597e-1	2.02
64	16	6.098e-5	5.15	2.011e-2	2.06	6.463e-4	4.06	7.966e-2	2.00
128	32	1.406e-5	4.34	9.976e-3	2.02	1.610e-4	4.01	3.981e-2	2.00
256	64	3.440e-6	4.09	4.978e-3	2.00	4.022e-5	4.00	1.990e-2	2.00
16	2	2.131e-2	-	2.385e-1	-	5.424e-2	-	6.596e-1	-
32	4	2.385e-3	8.94	6.821e-2	3.50	1.106e-2	4.90	3.229e-1	2.04
64	8	2.708e-4	8.81	2.404e-2	2.84	2.617e-3	4.23	1.598e-1	2.02
128	16	3.491e-5	7.76	1.050e-3	2.29	6.452e-4	4.06	7.967e-2	2.01
256	32	5.247e-6	6.65	5.044e-3	2.08	1.608e-4	4.01	3.981e-2	2.00

TABLE 3
Burgers Test, MSLM (L_2-orthogonal)

The MSLM does not show a significantly different behavior than the NGM. Accordingly, the ignoring of the linear term (q_t, ψ) appears justified.

In the case of the RNCM it strikes that despite large errors for small N (coarse grid resolution), already for $N = 16$ the error is nearly of the same size as that of the MSLM or the NGM. This result confirms our theoretical analysis according to which the RNCM is asymptotically of the same accuracy as the other schemes.

In the previous calculations exclusively L_2-orthogonal space decomposition has

n	N	$\|\|U-u\|\|$		$\|\|U_x-u_x\|\|$		$\|\|U-p\|\|$		$\|\|U_x-p_x\|\|$	
4	2	7.008e-2	-	5.713e-1	-	5.460e-2	-	6.565e-1	-
8	4	4.126e-3	17.0	1.615e-1	3.54	1.142e-2	4.78	3.222e-1	2.04
16	8	8.849e-4	4.66	7.963e-2	2.03	2.685e-3	4.25	1.596e-1	2.02
32	16	2.189e-4	4.04	3.980e-2	2.00	6.621e-4	4.06	7.965e-2	2.00
64	32	5.465e-5	4.01	1.990e-2	2.00	1.650e-4	4.01	3.980e-2	2.00
128	64	1.366e-5	4.00	9.949e-3	2.00	4.121e-5	4.00	1.990e-2	2.00
256	128	3.415e-6	4.00	4.975e-3	2.00	1.030e-5	4.00	9.949e-3	2.00
8	2	7.622e-2	-	5.159e-1	-	5.781e-2	-	6.660e-1	-
16	4	3.270e-3	23.3	8.405e-2	6.14	1.119e-2	5.17	3.233e-1	2.06
32	8	2.430e-4	13.5	3.986e-2	2.10	2.622e-3	4.27	1.598e-1	2.02
64	16	5.515e-5	4.41	1.990e-2	2.00	6.463e-4	4.06	7.966e-2	2.00
128	32	1.368e-5	4.03	9.949e-3	2.00	1.610e-4	4.01	3.981e-2	2.00
256	64	3.415e-6	4.01	4.975e-3	2.00	4.022e-5	4.00	1.990e-2	2.00
16	2	7.809e-2	-	5.002e-1	-	5.376e-2	-	6.685e-1	-
32	4	2.251e-3	34.7	4.836e-2	10.3	1.117e-2	4.81	3.236e-1	2.07
64	8	1.159e-4	19.4	2.003e-2	2.41	2.618e-3	4.27	1.598e-1	2.03
128	16	1.502e-5	7.72	9.952e-3	2.01	6.453e-4	4.06	7.967e-2	2.01
256	32	3.446e-6	4.36	4.975e-3	2.00	1.608e-4	4.01	3.981e-2	2.00

TABLE 4
Burgers Test, RNCM (L_2-orthogonal)

been used. To demonstrate the importance of the choice of the type of orthogonality, we repeat the same test with H^1-orthogonal space decomposition. In the present situation, this results in the usual hierarchical basis functions. We restrict us to the NGM and the MSLM. The RNCM does not make sense in this case due to the lack of coupling. This time we use a variant of the NGM, named NGM2, where in comparison to the standard scheme only the time derivative term (q_t, Φ) in (4) is ignored, while the nonlinearity remains unchanged.

For the NGM2, we observe an asymptotic behavior similar to that of a fine-grid SGM in the H^1-Norm but no asymptotic improvement when comparing the errors in the L_2-Norm. Dropping of terms like $((pq)_x, \Phi)$ would lead to the same effect. A further reduction, namely ignoring (p_t, ψ), will destroy even the positive effect in the H^1-Norm. Consequently, the neglection of the time derivative terms which couple the systems is not justified in the case of an H^1-orthogonal space decomposition, as predicted by our theoretical analysis. Furthermore, the difference between the solutions $u = p + q$ and $u = p$ is much smaller, i.e., the correction q is in this case not as relevant as in the previous cases. In contrast, the MSLM still behaves as good as for the case of L_2-orthogonal base functions, so that again the reduced treatment of the nonlinearity is justified. In addition, we observe that the NGM produces instabilities resulting in oscillatory solutions. Similar results had been found by Walsh in [28] for his time step-dependent orthogonality.

Recently, Garcia and Titi [11] have proposed a method, in the context of spectral approximations, which they called *AIM-Postprocessing*. It is claimed that under certain assumptions, one may solve time dependent problems with higher accuracy by first integrating with SGM on a coarse mesh, and then making use of a projection into a larger space by employing an approximate inertial manifold, i.e., by doing a final step with a NGM. For testing this approach, we integrate with the SGM until a time apart 10^{-5} from $T = 1$. Then, one step of an implicit Euler scheme is added in combination with the NGM. In this case, the component q does not have to be calculated over the whole time interval. We present here only the results for an

n	N	$\|\|U-u\|\|$		$\|\|U_x-u_x\|\|$		$\|\|U-p\|\|$		$\|\|U_x-p_x\|\|$	
4	2	2.287e-2	-	3.465e-1	-	6.720e-2	-	6.255e-1	-
8	4	7.272e-3	3.15	1.643e-1	2.11	1.529e-2	4.40	3.178e-1	1.97
16	8	1.926e-3	3.78	8.028e-2	2.05	3.746e-3	4.08	1.591e-1	2.00
32	16	4.883e-4	3.94	3.988e-2	2.01	9.319e-4	4.02	7.959e-2	2.00
64	32	1.225e-4	3.94	1.991e-2	2.00	2.327e-4	4.00	3.980e-2	2.00
128	64	3.065e-5	4.00	9.950e-2	2.00	5.816e-5	4.00	1.990e-2	2.00
256	128	7.664e-6	4.00	4.975e-3	2.00	1.454e-5	4.00	9.949e-3	2.00
8	2	3.501e-2	-	2.227e-1	-	6.875e-2	-	6.237e-1	-
16	4	1.089e-2	3.22	9.186e-2	2.42	1.574e-2	4.37	3.189e-1	1.96
32	8	2.865e-3	3.80	4.150e-2	2.21	3.877e-3	4.06	1.592e-1	2.00
64	16	7.251e-4	3.95	2.012e-2	2.06	9.661e-4	4.01	7.960e-2	2.00
128	32	1.818e-4	3.99	9.977e-3	2.02	2.413e-4	4.00	3.980e-2	2.00
256	64	4.549e-5	4.00	4.978e-3	2.00	6.032e-5	4.00	1.990e-2	2.00
16	2	3.858e-2	-	1.789e-1	-	6.916e-2	-	6.233e-1	-
32	4	1.189e-2	3.25	6.173e-2	2.90	1.587e-2	4.36	3.183e-1	1.96
64	8	3.122e-3	3.81	2.329e-2	2.65	3.914e-3	4.05	1.592e-1	2.00
128	16	7.899e-4	3.95	1.040e-2	2.24	9.757e-4	4.01	7.960e-2	2.00
256	32	1.980e-4	3.99	5.032e-3	2.07	2.438e-4	4.00	3.980e-2	2.00

TABLE 5
Burgers Test, NGM2 (H_0^1-orthogonal)

n	N	$\|\|U-u\|\|$		$\|\|U_x-u_x\|\|$		$\|\|U-p\|\|$		$\|\|U_x-p_x\|\|$	
4	2	1.706e-2	-	3.271e-1	-	8.444e-2	-	6.130e-1	-
8	4	3.588e-3	4.75	1.594e-1	2.05	2.185e-2	3.86	3.153e-1	1.94
16	8	8.781e-4	4.09	7.960e-2	2.00	5.501e-3	3.97	1.588e-1	1.99
32	16	2.187e-4	4.02	3.980e-2	2.00	1.378e-3	4.00	7.955e-2	2.00
64	32	5.464e-5	4.00	1.990e-2	2.00	3.446e-4	4.00	3.979e-2	2.00
128	64	1.366e-5	4.00	9.949e-3	2.00	8.615e-5	4.00	1.990e-2	2.00
256	128	3.414e-6	4.00	4.975e-3	2.00	2.154e-5	4.00	9.949e-3	2.00
8	2	1.590e-2	-	2.194e-1	-	9.123e-2	-	6.114e-1	
16	4	2.008e-3	7.92	9.164e-2	2.39	2.400e-2	3.80	3.151e-1	1.94
32	8	3.101e-4	6.48	4.145e-2	2.21	6.071e-3	3.95	1.588e-1	1.98
64	16	6.108e-5	5.08	2.011e-2	2.06	1.522e-3	3.99	7.954e-2	2.00
128	32	1.408e-5	4.34	9.976e-3	2.02	3.808e-4	4.00	3.979e-2	2.00
256	64	3.441e-6	4.09	4.978e-3	2.00	9.523e-5	4.00	1.990e-2	2.00
16	2	1.759e-2	-	1.893e-1	-	9.306e-2	-	6.112e-1	-
32	4	2.103e-3	8.36	6.533e-2	2.90	2.456e-2	3.79	3.151e-1	1.94
64	8	2.625e-4	8.01	2.391e-2	2.73	6.217e-3	3.95	1.588e-1	1.98
128	16	3.475e-5	7.55	1.049e-2	2.28	1.559e-3	3.99	7.954e-2	2.00
256	32	5.251e-6	6.62	5.044e-3	2.08	3.901e-4	4.00	3.979e-2	2.00

TABLE 6
Burgers Test, MSLM (H_0^1-orthogonal)

L_2-orthogonal space decomposition. For H^1-orthogonal bases, this correction process does not improve the error, due to the dropping of the time derivative on the q-level. In Table 7, N is the number of intervals used by the SGM, and n that used for the whole space when doing the last step with the NGM.

We see that the accuracy of the L_2-error does not change much, whereas it improves asymptotically for the H^1-error as if one had used the NGM from the very beginning, an effect similar to the one observed when using H^1-orthogonal space decomposition in the NGM2. The explanation for this behavior lies in the fact that the postprocessing does not effect the coarse grid part (small integration step) which plays the dominant part for the L_2-error, $\|e\|^2 = \|Pe\|^2 + \|Qe\|^2$. The H^1-error,

n	N	$\|\|U-u\|\|$	$\frac{\|\|U-p\|\|}{\|\|U-u\|\|}$	$\|\|U_x-u_x\|\|$	$\frac{\|\|U_x-u_x\|\|}{\|\|U_x-p_x\|\|}$
16	8	2.461e-3	1.44	7.974e-2	2.00
32	8	2.390e-3	1.48	4.217e-2	3.78
64	8	2.389e-3	1.48	2.537e-2	6.28
32	16	6.145e-4	1.43	3.981e-2	2.00
64	16	5.950e-4	1.47	2.020e-2	3.94
128	16	5.940e-4	1.48	1.069e-2	7.45
64	32	1.536e-4	1.43	1.990e-2	2.00
128	32	1.486e-4	1.48	9.987e-3	4.00
256	32	1.483e-4	1.48	5.070e-3	7,85

TABLE 7
Burgers Test, AIM Postprocessing (L_2-orthogonal)

$\|e_x\|^2 = \|(Pe)_x\|^2 + 2((Pe)_x, (Qe)_x) + \|(Qe)_x\|^2$, behaves differently because of the presence of the coupling terms and, what seems more important, $(Pe)_x$ scales by H, whereas $(Qe)_x$ scales by h. We conclude that AIM-Postprocessing can significantly improve the error only in special norms, e.g., the energy norm of the problem, and is therefore not to be considered as a universal approach.

5.2. The Kuramoto-Sivashinsky test. The Burgers equation has only stationary shock-like solutions. Any other dynamics has to be enforced by the right hand side and does not reflect the coupling mechanisms active in a general nonlinear equation. A more interesting nonstationary behavior can be achieved by the solutions of the Kuramoto-Sivashinsky equation,

$$U_t + U_{xxxx} + \alpha(U_{xx} + UU_x) = 0\,,$$

posed on the unit interval $\Omega = (0,1)$ with Dirichlet boundary conditions $U|_{\partial\Omega} = U^0$, $U_x|_{\partial\Omega} = U^1$. Here the choice of the parameter α is critical, as it determines the degree of relative instability present in the linear part of the equation which initiates the desired nonstationary behavior of the solution due to the nonlinearity. To this end, we first consider the corresponding eigenvalue problem

$$w_{xxxx} + \lambda w_{xx} = 0, \qquad w|_{\partial\Omega} = w_x|_{\partial\Omega} = 0\,,$$

whose eigenvalues are just the critical values for α, for which the linear main part becomes singular. On the unit intervall these eigenfunctions are $w(x) = \sin^2(m\pi x)$, $m \in N$. For our test we choose $\alpha = 250$, in which case at least two negative critical eigenvalues exist, and we are still far enough away from a singular point. The initial conditions are $U(x,0) = 1$, for $x \in (0.4, 0.6)$, and $U(x,0) = 0$, else. First using the SGM, we calculate a time periodic starting solution from which the initial data for the various tests are taken. The time discretization is by the Crank-Nicolson scheme with 2048 time steps of uniform length $k = 9.765625 \cdot 10^{-7}$ covering about 2.5 periods of the solution. The spatial discretization uses the cubic Hermite element with a uniform mesh spacing of $h = 2^{-r}$ for $r = 3, ..., 7$, where for $r = 7$ the space and time errors are well balanced. On the finest mesh reference solutions are calculated by using the SDIRK4 scheme from [14] with fixed time step-size $k = 9.765625 \cdot 10^{-7}$.

For the first test (Test 1) we simply start from the starting solution taken at some time t_0 and use homogeneous boundary data $U^0 = U^1 = 0$. In the second test (Test 2) we again start from the starting solution at time t_0, but now continue the

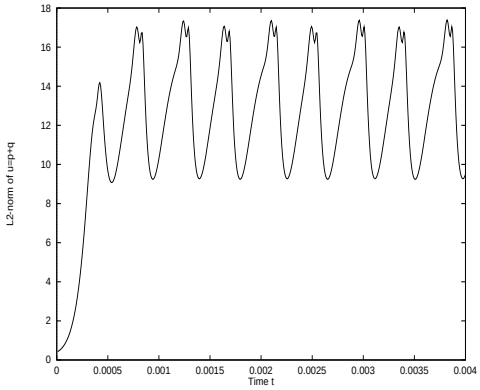

FIG. 1. *K-S Test: Evolution of the starting solution U (L_2-Norm)*

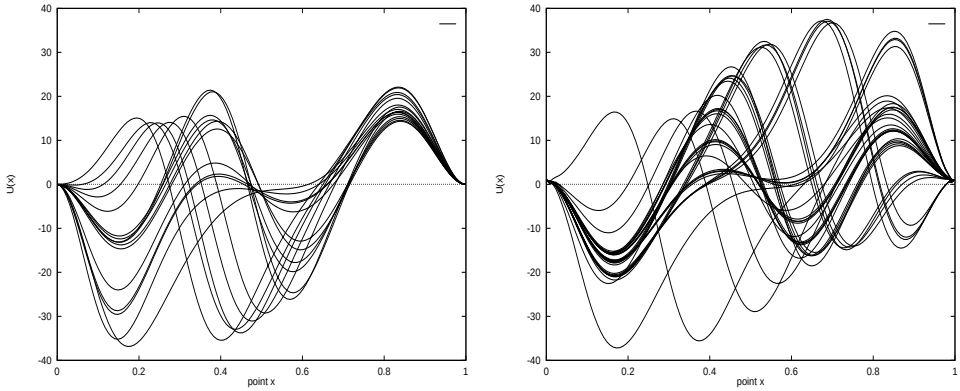

FIG. 2. *K-S Test: Evolution of U (Snapshots: Test 1 left, Test 2 right)*

computation with inhomogeneous boundary data $U^0 = 1$, $U^1 = 0$. This leads to a qualitatively different periodic solution than in Test 1.

In the case of only low resolution, i.e., $N = 4$ and $n = 8, 16$, all the schemes show bad approximation quality. Neither the period nor the amplitude of the solutions are satisfactorily represented. The RNCM becomes even stationary as does the SGM for $n = N = 4$. Hence, in this case on such a coarse mesh, MSLM and NGM are superior over the RNCM. Increasing the grid ratio H/h reduces the accuracy for both, MSLM and NGM. In the second test case (Test 2), the picture is different. Here, for increasing mesh ratios the NGM performs better than the MSLM. However, already for $N = 8$, all schemes are of almost equally good accuracy, i.e. close to that of the SGM calculated on the fine mesh with h. This is surprising since one might suspect that the over-simplified RNCM should behave much worse than the other schemes at low resolution. This indicates that it is essentially the quality in the approximation of the linear main part of the problem more precisely its critical negative eigenvalues, by the fine-grid correction, which determines the overall accuracy. The representation of the (lower-order) nonlinearity seems, in this case, to be only of minor importance for capturing the dynamics of the solution.

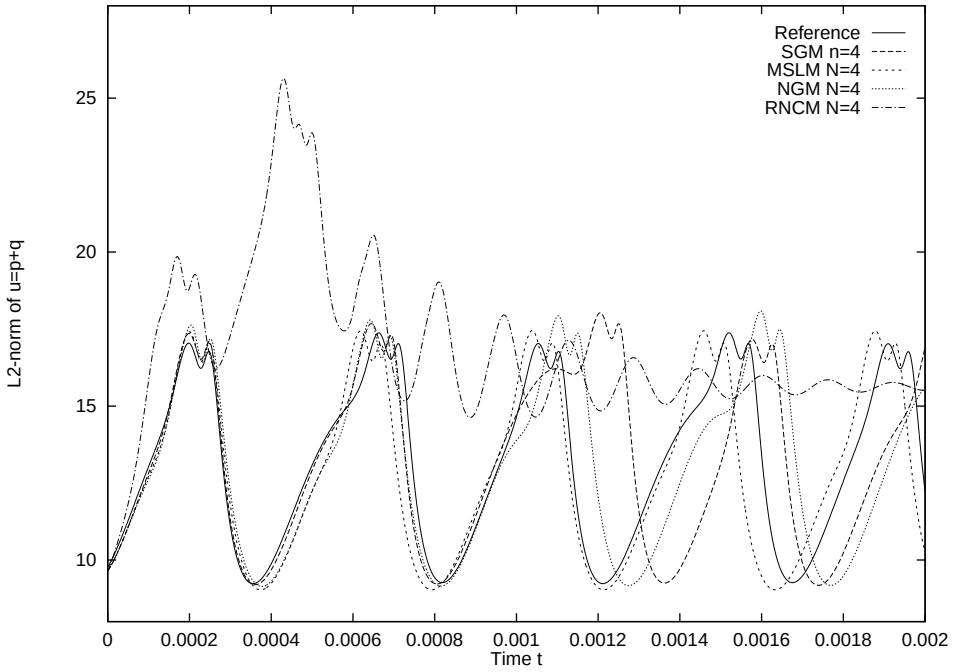

FIG. 3. *K-S Test 1: L_2-Norm plots (n=8, N=4)*

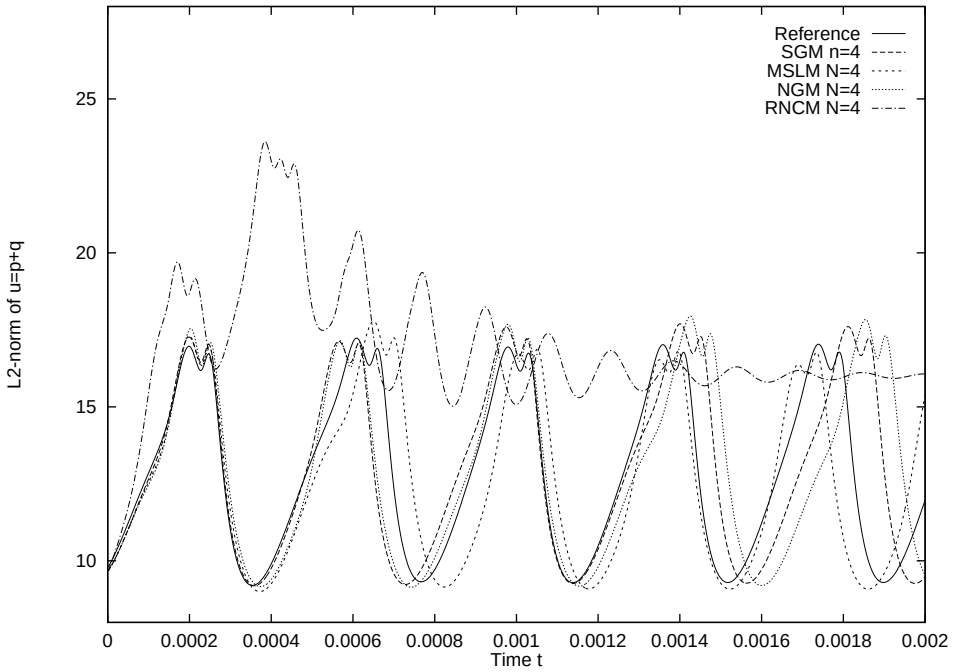

FIG. 4. *K-S Test 2: L_2-Norm plots (n=8, N=4)*

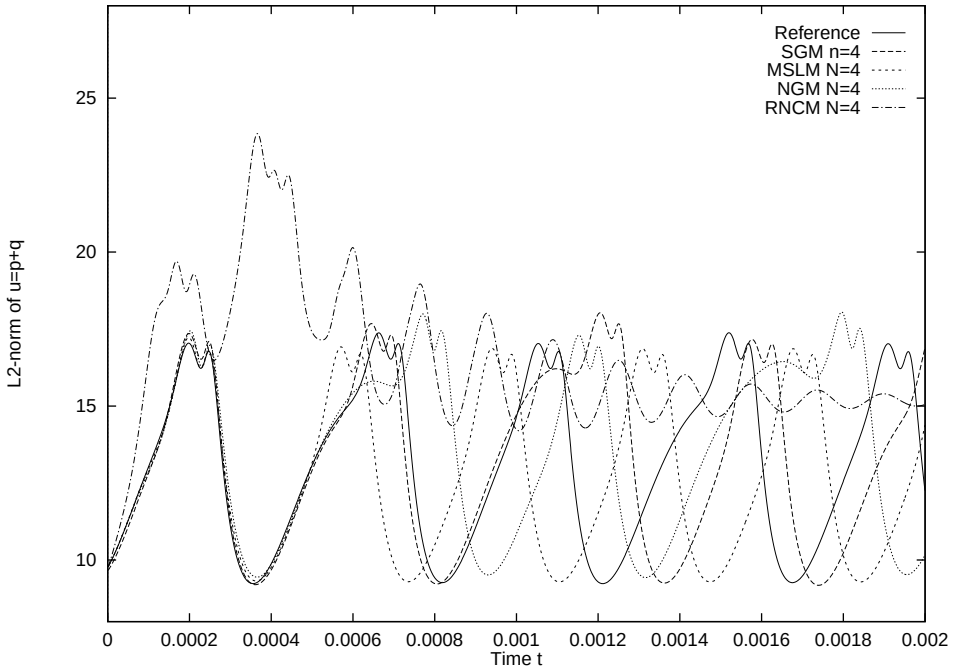

FIG. 5. *K-S Test 1: L_2-norm plots (n=16, N=4)*

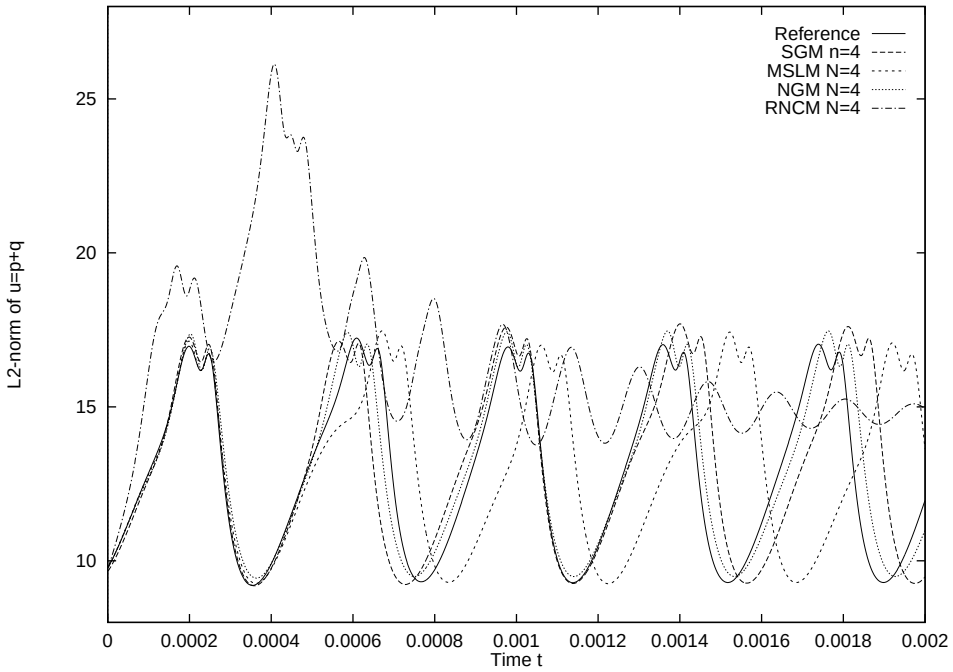

FIG. 6. *K-S Test 2: L_2-norm plots (n=16, N=4)*

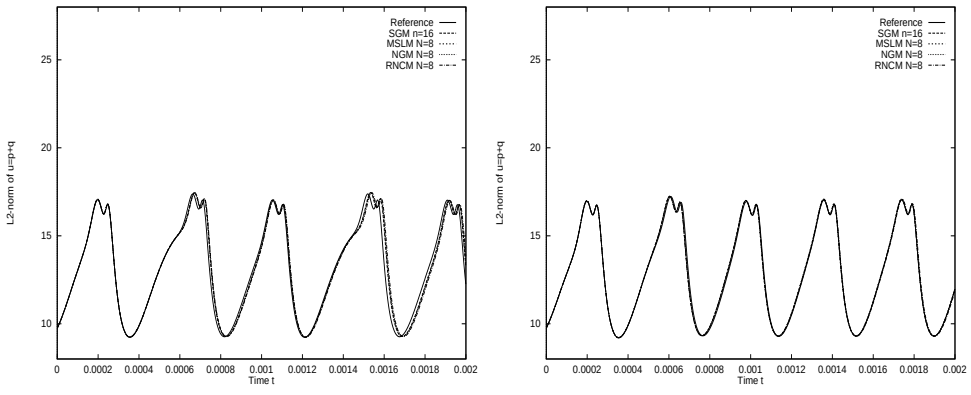

FIG. 7. *K-S Test 1 (left) and Test 2 (right): L_2-norm plots (n=16, N=8)*

6. Estimate of computational work. For illustration, we give a work estimate for the various schemes considered applied to the one-dimensional Burgers equation. As usual the finite element solution is expressed in the form

$$u(x,t) = \sum_i \alpha_i(t)\phi_i(x),$$

where the ϕ_i are nodal basis functions of V_h with local support. Then, one constructs a basis in the complement space W_h^H which is almost L_2-orthogonal to the standard nodal basis of the coarse grid space V_H. For the sake of computational efficiency, the basis functions should have small support while vanishing at the boundary (Dirichlet problem). The last condition cannot be strictly fulfilled for the first and the last interior node when full L_2-orthogonality is required as, due to the explicit setting of the values at the boundary points to zero, the corresponding base functions are not L_2-orthogonal to those corresponding to the boundary nodes of the coarse grid. For example, in the case of linear finite elements on an equidistant mesh and only one refinement step, i.e., the fine grid (with $2N$ nodes) results from the coarse grid (with N nodes) by one bisection step, this leads to basis functions of the form

$$\psi_i = \phi_{i-2} - 6\phi_{i-1} + 10\phi_i - 6\phi_{i+1} + \phi_{i+2}, \quad i = 2j,\ j \in \{2,\ldots,N-1\},$$

at the inner grid points, and $\psi_1 = 9\phi_1 - 6\phi_2 + \phi_3$, at the first inner point. With

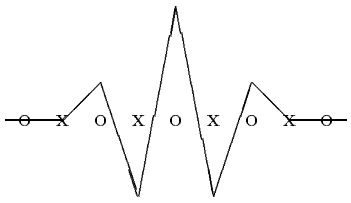

FIG. 8. *L_2-orthogonal finite element basis function*

our construction this lack of orthogonality at the boundary is not critical since, for

a Dirichlet problem, the boundary values are given and in a Galerkin method the system is only tested by basis functions corresponding to interior nodes. However, the resulting system matrices are no longer tri-diagonal as in the SGM.

REMARK 6.1. *In the case of periodic boundary conditions we have a fully L_2-orthogonal basis and the new basis functions in the space W_h^H build a pre-wavelet basis as shown in [13].*

Next, we consider the evaluation of the nonlinearity in the case $n = 2N$. The critical points are the cost for i) the transfer between the representation of the solution $u = p + q$ with respect to the natural nodal basis of V_h and the orthogonal bases of V_H and W_h^H, and vice versa, and ii) the evaluation of nonlinear terms of the form $b(u, \Phi, q)$ and $b(p, p, \psi)$ for elements $p, \Phi \in V_H$ and $q, \psi \in W_h^H$. In the SGM the work required for building up the nonlinearity

$$b(u, u, \phi_k) = \sum_{i,j} u_i u_j b(\phi_i, \phi_j, \phi_k)$$

depends on the size of the support of the base functions ϕ_i of V_h. In the case of piecewise linear finite elements, only directly neighbored basis functions have a common support. Neglecting the neighborhood of the boundary, for each k at most 7 of the terms $b(\phi_i, \phi_j, \phi_k)$ (observing $b(\phi_i, \phi_i, \phi_i) = 0$, actually only 6 terms) have to be evaluated. Hence, for building up the nonlinearity, at most $7n$ operations (Op) are needed.

In using one of the reduced Galerkin schemes, we have to select a proper basis presentation for the discrete solution $u = p + q$. At the initial state, we have to compute initial values for $p \in V_H$ and $q \in W_h^H$ from u_0, and later on, we have to transfer the computed coarse- and fine-grid components p and q, respectively, back to V_h for evaluating the composed solution $u = p + q \in V_h$. This transfer is cheap in the case of an H^1-orthogonal "hierarchical" basis, but it requires the solution of a globally coupled system in the case of L_2-orthogonality. Accordingly, the reduction of terms like $b(p, p, \phi)$ to the standard nodal basis representation is unfavorable as the corresponding support becomes global. This spoils the locality of finite elements and lets the complexity increase from the usual $O(n)$ Op to $O(n^2)$ Op making the scheme similar to a spectral method. In the case of hierarchical bases (H^1-orthogonality) the situation is more favorable, but still the number of operations is almost the same as for a SGM, as the support of the coarse grid basis functions is local but enlarged. In view of these observations, in using L_2-orthogonality, it would not be appropriate to work with the standard nodal basis but rather directly with the (mutually orthogonal) bases of V_H and W_h^H described above. This, however, excludes the use of hierarchical bases preconditioning which would otherwise be the natural approach for the efficient iterative solution of the discrete problems.

Next, we estimate the work required in using these orthogonal bases. For the NGM and the MSLM, the building up of the nonlinear terms $b(p, p, \Phi) - b(p, \Phi, q)$ and $b(p, p, \psi)$ costs about $7N + 10N$ Op, respectively, $10N$ Op, and consequently, all together $27N$ Op. This means, compared to the SGM (with $7n = 14N$ Op) an increase of almost 100%. In this estimate the additional costs of $4n$ operations arising by the transformation from the L_2-orthogonal to the standard nodal basis representation has not even been counted. Consequently, there is no gain in the cost of setting up the formulation with the NGM or the MSLM, at least in the case of an L_2-orthogonal decomposition. In the RNCM the nonlinearity has to be evaluated only on the coarse

mesh which, in the case $H = 2h$, leads to a saving of 50%. However, as we have seen, this scheme is useful only for high-accuracy computations when the solution is already sufficiently represented on the coarse mesh. When using the hierarchical basis representation (H^1-orthogonality), the situation is different. In this case, in view of our computational results, only the MSLM may be useful, as the other methods are lacking accuracy. But again the evaluation of the nonlinear terms $b(p,p,\Phi)-b(q,\Phi,p)$ and $b(p,p,\psi)$ require all together $15N$ Op which does not mean a reduction in cost. This work estimate for the one-dimensional Burgers equation is even optimistic for the general multi-dimensional case, as here, due to the larger support of basis functions, the cost for computing the nonlinear terms drastically increase.

The second aspect to be considered in the work estimate of the reduced Galerkin schemes is the solution of the resulting algebraic systems which are of the form

$$\begin{bmatrix} PM & SM \\ SM^t & QM \end{bmatrix} \begin{bmatrix} p_t \\ q_t \end{bmatrix} + \begin{bmatrix} PA+PN_1 & SA+PN_2 \\ SA^t+QN_2 & QA+QN_1 \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} Pf \\ Qf \end{bmatrix}.$$

In the present one-dimensional situation, the matrices PM, PA and PN_1 are all tri-diagonal. The structure of the other submatrices depends on the basis representation and on the ratio between coarse and fine grid size. The number of diagonals for the matrices QM, QA and QN_1 is at most seven (for $n = 2N$ only five) when using L_2-orthogonal bases, and at most three (for $n = 2N$ only one) when using hierarchical bases. The other matrices representing the coarse-fine-grid coupling have about $4(n-N)$ non-zero elements. In view of this situation, the only chance for gaining a reduction in the total number of operations would be a faster convergence of the iterative process for solving the nonlinear system. However, in our computations we did not observe such a reduction, the number of iterations increased (by a factor of up to ten) and in some cases not even convergence could be achieved especially with the NGM.

7. Conclusions. From the above discussion we draw the following conclusions: In the context of finite element discretization the choice of the inner product for the orthogonal decomposition is essential. L_2-orthogonality is required for most of the NGM-type schemes in order to achieve optimal order accuracy. In the approximate solution $u = p + q$ also the fine-grid component q is relevant for the accuracy. In the case of an L_2-orthogonal decomposition, all the different ways of treating the nonlinearity lead to the same asymptotic error behavior, even with the same error constants. This indicates that mere asymptotic error analysis does not say much about the practical performance of these schemes.

For capturing the dynamics of the solution, the most economical schemes are either the standard Galerkin method (SGM) on the coarse grid, if it works, or the simplest scheme (RNCM) using L_2-orthogonal finite element bases and dropping all nonlinear terms on the fine-grid level. The additional effort of the full SGM pays in cases where the RNCM does not improve the coarse grid solution due to nonlinear effects. In this situation the NGM as well as the MSLM show a rather unpredictable behaviour depending on the particular data of the problem. Altogether the new schemes are useful only if really high accuracy is required for the approximate solution, which might be of interest for certain one-dimensional problems. It seems that the observable effect of the NGMs has not so much to do with modeling of the nonlinearity but rather with the particular approximation properties of the basis functions used. For most NGM-type schemes the formal reduction in the number of nonlinear terms does not automatically lead to a practical advantage. The resulting schemes are too costly and often lack stability.

REFERENCES

[1] A. Ammi and M. Marion, *Nonlinear Galerkin methods and mixed finite elements: Two-grid algorithms for the Navier-Stokes equations,* Numer. Math, 68 (1994), pp. 189-213.

[2] M. Chen and R. Temam, *Nonlinear Galerkin method in the finite difference case and wavelet-like incremental unknowns,* Numer. Math., 64 (1993), pp. 271-294.

[3] C. Dedulvier, M. Marion and E. Titi, *On the rate of convergence of the nonlinear Galerkin methods,* Math. Comp., 60 (1993), pp. 495-514.

[4] C. Devulder and M. Marion, *A class of numerical algorithms for large time integration: the nonlinear Galerkin methods,* SIAM J. Numer. Anal., 29 (1992), pp. 462-483.

[5] E. Fabes, M. Luskin and G. R. Sell, *Construction of inertial manifolds by elliptic regularization,* J. Differ. Equations, 89 (1991), pp. 355-387.

[6] C. Foias, B. Nicolaenko, G. R. Sell and R.Temam, *Inertial manifolds for the Kuramoto-Sivashinsky-Equation and an estimate of their lowest dimension,* J. Math. Pures et Appl., 67 (1988), pp. 197-226.

[7] C. Foias, O. Manley and R.Temam, *Modeling of the interaction of small and large eddies in two dimensional turbulent flows,* M^2AN, 22 (1988), pp. 93-114.

[8] C. Foias, O. Manley and R.Temam, *Approximate inertial manifolds and effective viscosity in turbulent flows,* Phys. Fluids, A 3, No. 5, Part 1 (1991), pp. 898-911.

[9] C. Foias, R. Sell and R.Temam, *Inertial manifolds for nonlinear evolution equations,* J. Diff. Eq., 73 (1988), pp. 309-353.

[10] B. Garcia-Archilla and J. de Frutos, *Time integration of the nonlinear Galerkin method,* IMA J. Num. Anal., 15 (1995), 221-244.

[11] B. Garcia-Archilla and E. Titi, *Postprocessing the Galerkin method: A novel approach to approximate inertial manifolds,* submitted to SIAM J. Numer. Anal. (1995).

[12] D. Gottlieb and R. Temam, *Implementation of the nonlinear Galerkin method with pseudospectral (collocation) discretizations,* Appl. Numer. Math., 12 (1993), pp. 119-134.

[13] O. Goubet, *Nonlinear Galerkin methods using hierarchical almost-orthogonal finite elements bases,* Nonlinear Analysis, 20 (1993), pp. 223-247.

[14] E. Hairer and G.Wanner, *Solving Ordinary Differential Equations II,* Springer, Berlin, 1991.

[15] J. G. Heywood and R. Rannacher, *Finite element approximation of the nonstationary Navier-Stokes-Problem. Part I. Regularity of solutions and second-order error estimates for spatial discretizations,* SIAM J. Numer. Anal., 19 (1982), pp. 275-311.

[16] J. G. Heywood and R. Rannacher, *On the question of turbulence modeling by approximate inertial manifolds and the nonlinear Galerkin scheme,* SIAM J. Numer. Anal., 30 (1993), pp. 1603-1621.

[17] F. Jauberteau, C. Rosier and R. Temam, *A nonlinear Galerkin method for the Navier-Stokes-Equations,* Comp. Meth. in Appl. Mech. Eng., 80 (1990), pp. 245-260.

[18] M. S. Jolly, I. G. Kevrekidis and E. S. Titi, *Approximate inertial manifolds for the Kuramoto-Sivashinsky equation: Analysis and computation,* Physica, D 44 (1990), pp. 38-60.

[19] D. A. Jones, L. G. Margolin and E. S. Titi, *On the effectiveness of the approximate inertial manifold - a computational study,* to appear in Theo. and Comp. Fluid Dynamics, 1995.

[20] M. Kwak, *Finite dimensional inertial manifolds for the 2D Navier-Stokes equations,* Technical Report 91-30, AHPCRC, University of Minnesota, 1991.

[21] J. Laminie, F. Pascal and R. Temam, *Implementation of finite element nonlinear Galerkin methods using hierarchical bases,* Comput. Mech., 11 (1993), pp. 384-407.

[22] L. G. Margolin and D. A. Jones, *An approximate inertial manifold for computing Burgers' equation,* Physica, D60 (1992), pp. 175-184.

[23] M. Marion and R. Temam, *Nonlinear Galerkin methods,* SIAM J. Numer. Anal., 26 (1989), pp. 1139-1157.

[24] M. Marion and R. Temam, *Nonlinear Galerkin methods: The finite element case,* Numer. Math., 57 (1990), pp. 205-226.

[25] M. Marion and J. Xu, *Error estimates on a new Nonlinear Galerkin method based on two-grid finite elements,* Preprint, 1994.

[26] G. R. Sell and Y. You, *Inertial manifolds: The non-self-adjoint case,* J. Differ. Equations, 96 (1992), pp. 203-255.

[27] J. Shen and R. Temam, *Nonlinear Galerkin method using Chebyshev and Legendre polynomials. I: The one-dimensional case,* SIAM J. Numer. Anal., 32 (1995), pp. 215-234.

[28] O. Walsh, *On approximate inertial manifolds for the Navier-Stokes-Equations using finite elements,* Ph.D thesis, UBC, Vancouver, 1994.