



Sobolev estimates for non-uniformly parabolic PDEs

Steve Schochet¹

Received: 4 October 2021 / Accepted: 8 January 2022 / Published online: 24 January 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract

Sobolev energy estimates are proven for solutions of initial-value-problems for non-uniformly parabolic second-order PDEs having symmetric coefficients depending on the independent and dependent variables. Local-in-time existence of solutions to initial-value problems for such systems and convergence results for singular limits involving such systems are consequences of those uniform bounds.

Keywords Non-uniformly parabolic · Sobolev energy estimates · singular limit

Mathematics Subject Classification 35K65 · 35K45 · 35B45 · 35B25

1 Introduction

There are many ways that a parabolic second-order PDE can fail to be uniformly parabolic. Many results for such PDEs concern equations that fail to be uniformly parabolic on specified lower-dimensional sets, such as when the dependent variable equals zero (e.g. the porous media equation [1]) or tends to infinity [2], or the time variable equals zero [3]. Other results concern partially parabolic systems, which are uniformly parabolic with respect to some of the spatial variables and contain no second derivatives with respect to the remaining spatial variables [4], or hyperbolic-parabolic systems [5], which are uniformly parabolic with respect to certain equations and contain no second derivatives in the remaining equations. The results presented here place no restrictions on the locations where uniform parabolicity fails and do not assume uniform parabolicity with respect to certain spatial or dependent variables. Linear non-uniformly parabolic equations and systems without such restrictions are a special case of the degenerate elliptic equations and systems treated in [6–8], but the results there cannot be generalized to the nonlinear equations and systems considered here because, as detailed in Appendix B, the estimates require more derivatives of the coefficients than are estimated for the solution.

This article is part of the section “Theory of PDEs” edited by Eduardo Teixeira.

✉ Steve Schochet
schochet@tauex.tau.ac.il

¹ School of Mathematical Sciences, Tel Aviv University, 69978 Tel Aviv, Israel

Methods specific to parabolic equations and systems [9] can often be utilized even when uniform parabolicity does not hold at specified locations [10], but do not seem to be applicable under the conditions considered here. The energy method, which is applicable to a wide variety of evolutionary PDEs, will be used instead. The difficulty in deriving Sobolev energy estimates for parabolic equations is due to the troublesome contributions to the energy estimates that arise from differentiating the coefficients of the second-order terms. For uniformly parabolic, partially-parabolic, or hyperbolic-parabolic equations [11, Theorem 4] the troublesome terms can be estimated by cancellation against the helpful contribution obtained on account of full or partial uniform parabolicity. The main point of this paper is that those troublesome terms can be controlled even when the parabolic terms are nonlinear and make no helpful contribution whatsoever.

Initial-value problems will be considered here for PDEs having one of the forms

$$A^0(t, x, u)u_t + \sum_{j=1}^d A^j(t, x, u)u_{x_j} = \sum_{j,k=1}^d D^{j,k}(t, x, u)\partial_{x_j}\partial_{x_k}u + F(t, x, u) \quad (1)$$

or

$$A^0(\varepsilon u)u_t + \sum_{j=1}^d A^j(t, x, u)u_{x_j} + \frac{1}{\varepsilon} \sum_{j=1}^d C^j u_{x_j} = \sum_{j,k=1}^d D^{j,k}(t, x, u)\partial_{x_j}\partial_{x_k}u + F(t, x, u), \quad (2)$$

where the first-order part is symmetric hyperbolic and the second-order terms are symmetric and satisfy the non-strict Legendre condition, as detailed in Assumption 2.1 below. In contrast to the corresponding hyperbolic systems in which the second-order spatial derivatives are absent, when the second derivative terms are present but make no helpful contribution it is not possible in general to obtain a closed energy estimate for just spatial derivatives of solutions to (1) or (2). The problem is that the time derivative appears in the equations for spatial derivatives, and when that time derivative is eliminated by solving the original PDE for u_t and substituting the result for u_t wherever it appears on the right side of energy estimates for the spatial derivatives then the resulting expressions involve one more spatial derivative than is being estimated. The main results Theorems 2.7 and 3.1 of this paper say that under the hypotheses mentioned above a local in time closed energy estimate can be obtained when both space and time derivatives are estimated together, where for (2) the norms involving time derivatives are weighted by appropriate powers of ε and the resulting estimate is uniform in ε .

The difficulty in proving estimates for the solution and its derivatives when the parabolic terms provide no helpful contribution is that those parabolic terms involve one more derivative than is dealt with in classical estimates for hyperbolic systems. The key estimates (13)–(15) provide bounds for the parabolic terms for the case when more derivatives have been applied to a single factor of the solution than are being estimated. Although those estimates, like the key estimate [6, (3.4)–(3.5)] in the theory of degenerate elliptic systems, involve repeated integrations by parts, the two sets of estimates are otherwise different in every possible manner.

However, a problem remains even for the terms in which the maximum number of derivatives applied to any one factor of the solution is at most the number of derivatives being estimated, because the total number of derivatives applied to all factors of the solution remains larger than for the hyperbolic case. The standard proof of the commutator estimate [11, Lemma A.1] for the product of two functions used in hyperbolic estimates yields

the slightly more general result

$$\sum_{\substack{\ell \text{ } |\alpha_\ell| \leq s+1, \max_\ell |\alpha_\ell| \leq s \\ \text{provided } s \geq s_0 + 1,}} \int \prod_\ell [D^{\alpha_\ell} w^{(\ell)}]^2 dx \leq c \prod_\ell \|w^{(\ell)}\|_{H^s}^2 \tag{3}$$

where D^{α_ℓ} is a spatial derivative of order $|\alpha_\ell|$ and

$$s_0 := \lfloor \frac{d}{2} \rfloor + 1 \tag{4}$$

is the Sobolev embedding index, i.e., the smallest integer such that $H^{s_0} \subset L^\infty$. The requirement that s be at least $s_0 + 1$ is therefore a standard requirement in the local existence theory for nonlinear symmetric hyperbolic systems. In order to deal with the parabolic terms and in light of the fact that we need to estimate time as well as spatial derivatives, we require a generalization of (3) to the case when the total number of derivatives is $s + 2$ rather than $s + 1$, and moreover some of those derivatives may be with respect to time, although the integral and norms still involve integration only over the spatial variables. The required calculus estimate, which more generally allows the total number of derivatives to be $s + r$ provided that $s \geq s_0 + r$, is presented in Lemma A.1 in Appendix A. The estimate proven there is also used in [12]. Since $s + 2$ derivatives appear in the parabolic terms when s derivatives are taken of the PDE, the condition on s in Lemma A.1 requires that s be at least $s_0 + 2$, which will therefore be assumed in both Theorems 2.7 and 3.1.

Since the presentation of the assumptions and the norms used in the theorems is somewhat lengthy, the full statements of the theorems on uniform bounds for (1) and for (2) are given in the same section as the corresponding proof. The uniform bound for (1) is proven in Sect. 2; for emphasis, the key estimates for the nonuniformly parabolic term are presented there in Lemma 2.3. The modifications needed to obtain uniform bounds for (2) using ε -weighted norms of time derivatives are shown in Sect. 3, both for the case when the first time derivative is uniformly bounded at time zero (known as the case of “well-prepared” initial data), and for the general case. Obtaining uniform estimates for (2) is more complicated than for purely hyperbolic systems having large constant-coefficient terms on account of the need to estimate time derivatives in order to obtain bounds for spatial derivatives. Moreover, in contrast to the purely hyperbolic case, the estimates here for well-prepared initial data are more complicated than those for the general case. The key estimate (14) for the parabolic terms is the source of the chief difficulty, since it makes the nonuniform second time derivative u_{tt} appear in the estimate for the first time derivative u_t that we want to show to be uniformly bounded. In order to prove that u_t nevertheless remains uniformly bounded for positive times it is necessary to use an intricate system of ε -dependent weights.

Once the energy estimates are obtained, local-in-time existence of solutions to the initial-value problem for those equations, for a time independent of ε in the case of (2), then follows by a mollification argument as in [13, Section 5.2, Section 7.1]. Furthermore, the local-in-time convergence of solutions of (2) to corresponding solutions of appropriate limit or profile equations as $\varepsilon \rightarrow 0+$ then follows as in [14] since the convergence parts of the theorems there only require uniform estimates and times of existence plus an easy L^2 estimate for the difference of sufficiently smooth solutions of the limit system. The statements of those theorems and some indications of the proofs are presented in Sect. 4. The full details of the existence and convergence proofs are omitted since they are minor variations of those in the references mentioned above.

In addition to the existence and convergence results just described, several additional results follow from the methods here. First, uniform estimates can be obtained for two variants of (1)–(2):

1. If the diagonal terms $(D^{j,k})_{\ell\ell}$ of the second-order coefficients depend only on t, x , and u_ℓ , the off-diagonal terms depend only on t and x , and all second-order coefficients $D^{j,k}$ have one more bounded derivative than assumed below, then the second-order terms can be allowed to have the divergence form $\sum_{j,k} \partial_{x_j} [D^{j,k} \partial_{x_k} u]$. In particular, divergence form is allowed for scalar PDEs. The additional assumptions just mentioned are needed in order to eliminate via integration by parts the term in the Sobolev energy estimates in which ∂_{x_j} and all derivatives applied to the PDE are placed on the variable u appearing inside $D^{j,k}$.
2. The solution u can be required to satisfy a linear constant-coefficient constraint $Lu = 0$, enforced by adding the adjoint Lagrange multiplier term $L^* \phi$ to the equation. This situation is familiar from the incompressible Navier-Stokes equations.

For brevity, further details of the extension to such variant equations are omitted. Applying both extensions yields uniform estimates for the incompressible Navier-Stokes equations with eddy viscosity considered in [15], even when the variable-coefficient eddy viscosity coefficient is merely non-negative, since the incompressibility constraint can be used to write the viscosity terms in symmetric form.

Finally, the results here also make it possible to generalize geometric optics results of [16] to the case when the second-order terms are not uniformly parabolic, by using scaled norms $\|u\|_{L^2} + \sum_{1 \leq \ell+m \leq s} \varepsilon^{\ell+m-1} \|\partial_t^m u\|_{H^\ell}$. The verification that the bound obtained for the scaled norms is independent of ε is similar to the calculations in [16, Section IV] or [17, Proof of Theorem 3.6].

2 Uniform bounds

Since $\partial_{x_j} \partial_{x_k} = \partial_{x_k} \partial_{x_j}$ we can ensure without loss of generality that

$$D^{j,k} = D^{k,j} \tag{5}$$

by replacing $D^{j,k}$ with $\frac{1}{2}(D^{j,k} + D^{k,j})$ if necessary. The following assumptions will be made on $D^{j,k}$ and the other coefficient matrices appearing in (1) and (2).

Assumption 2.1 1. For every $k < \infty$ there is a constant $b(k)$ such that

$$\begin{cases} \|A^0(t, x, v)\|_{C^s([0, \infty) \times \mathbb{R}^d \times \{|v| \leq k\})} \leq b(k) & \text{for (1)} \\ \|A^0(v)\|_{C^s(\{|v| \leq k\})} \leq b(k) & \text{for (2)} \end{cases} \tag{6:s}$$

and

$$\begin{aligned} & \| \{A^j(t, x, v)\}_{j=1}^d, \{D^{j,k}(t, x, v)\}_{j,k=1}^d \|_{C^s([0, \infty) \times \mathbb{R}^d \times \{|v| \leq k\})} \\ & + \|F(t, x, v), \int_0^1 \frac{\partial F}{\partial u}(t, x, rv) dr \|_{C^s([0, \infty) \times \mathbb{R}^d \times \{|v| \leq k\})} \leq b(k), \end{aligned} \tag{7:s}$$

where s is a positive integer to be specified.

2. The matrices A^j for $0 \leq j \leq d$, the matrices C^j for $1 \leq j \leq d$, and the matrices $D^{j,k}$ for $1 \leq j, k \leq d$ are symmetric, and (5) holds.

3. The matrix A^0 is positive definite, i.e., for every $k < \infty$ there is a $\delta(k) > 0$ such that

$$\begin{cases} w^T A^0(t, x, v)w \geq \delta(k)|w|^2 & \text{for } |v| \leq k & \text{for (1)} \\ w^T A^0(v)w \geq \delta(k)|w|^2 & \text{for } |v| \leq k & \text{for (2).} \end{cases} \tag{8}$$

4. The matrices $D^{j,k}$ satisfy the non-strict Legendre condition

$$\sum_{j,k=1}^d (w^{(j)})^T D^{j,k}(t, x, u)w^{(k)} \geq 0$$

for all (t, x, u) and all sets of vectors $\{w^{(j)}\}_{j=1}^d$. (9)

Remark 2.2 The non-strict Legendre condition (9) in Assumption 2.1 ensures that

$$- \int_{\mathbb{R}^d} \sum_{j,k=1}^d (\partial_{x_j} v(t, x))^T D^{j,k}(t, x, u(t, x)) \partial_{x_k} v(t, x) \leq 0, \tag{10}$$

which will be vital for the proof of Theorem 2.7 below. In contrast, if the $D^{j,k}$ are not constant then the non-strict Legendre-Hadamard condition

$$\sum_{j,k=1}^d \xi_j \xi_k w^T D^{j,k} w \geq 0 \quad \text{for all real numbers } \{\xi_j\}_{j=1}^d \text{ and vectors } w \tag{11}$$

only implies the the sharp Gårding inequality (e.g., [13, Section 0.7])

$$- \int_{\mathbb{R}^d} \sum_{j,k=1}^d (\partial_{x_j} v(t, x))^T D^{j,k}(t, x, u(t, x)) \partial_{x_k} v(t, x) \leq c \|v\|_{H^{1/2}}^2, \tag{12}$$

which is not sufficient for the proof of Lemma 2.3 because there is no helpful contribution to compensate for the extra half derivative on the right side of (12).

The following lemma contains the key estimates involving the nonuniformly parabolic terms that will be used for the troublesome case when more derivatives are applied to the solution than are being estimated.

Lemma 2.3 *Let the spatial domain be the torus \mathbb{T}^d or the whole space \mathbb{R}^d . Suppose that Assumption 2.1 holds, with(6:s)–(7:s) holding for $s := s_0 + 2$. In addition, let u and v be functions in $\cap_{m=0}^2 C^m([0, T_u]; H^{s-m})$ and $\cap_{m=0}^1 C^m([0, T_u]; H^{2-m})$, respectively. Then*

1.

$$\int \sum_{j,k} v^T D^{j,k}(t, x, u(t, x)) \partial_{x_j} \partial_{x_k} v \leq b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+2}}) \|v\|_{L^2}^2, \tag{13}$$

where here and later P denotes a polynomial, which may be different in different occurrences.

2.

$$\begin{aligned}
 & 2 \int \sum_{j,k} (\partial_t v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_j} \partial_{x_k} v \\
 & \leq -\frac{d}{dt} \left[\int \sum_{j,k} (\partial_{x_j} v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_k} v \right] \\
 & \quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+1}}) (1 + \|u_t\|_{H^{s_0+1}}) \|v_t\|_{L^2} \|v\|_{H^1} \\
 & \quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0}}) (1 + \|u_t\|_{H^{s_0}}^2 + \|u_{tt}\|_{H^{s_0}}) \|v\|_{H^1}^2,
 \end{aligned} \tag{14}$$

where here and later $\partial_t D^{j,k}(t, x, u(t, x))$ means

$$\left[\partial_t (D^{j,k}(t, x, w)) + (u_t \cdot \nabla_w)(D^{j,k}(t, x, w)) \right] \Big|_{w=u(t,x)}$$

and $\partial_{x_\ell} D^{j,k}$ has an analogous meaning.

3.

$$\begin{aligned}
 & \int \sum_{j,k} (\partial_{x_\ell} v)^T (\partial_{x_m} D^{j,k}(t, x, u(t, x))) \partial_{x_j} \partial_{x_k} v \\
 & \leq b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+2}}) \|v\|_{H^1}^2.
 \end{aligned} \tag{15}$$

Proof To estimate the left side of (13), the two derivatives ∂_{x_j} and ∂_{x_k} must both be transferred from v to $D^{j,k}$. That can be accomplished by using integration by parts, the product rule for derivatives, the consequence (10) of the assumed Legendre condition (9), the assumed symmetry of the coefficient matrices $D^{j,k}$, the integration-by-parts identity

$$abc' + a'bc = (abc)' - ab'c, \tag{16}$$

and the facts that the integral over \mathbb{R}^d or \mathbb{T}^d of a spatial derivative vanishes and that H^{s_0+2} is an algebra. This yields,

$$\begin{aligned}
 & 2 \int \sum_{j,k} v^T D^{j,k} \partial_{x_j} \partial_{x_k} v = -2 \int \sum_{j,k} \partial_{x_j} \left[v^T D^{j,k} \right] \partial_{x_k} v \\
 & = -2 \left\{ \int \sum_{j,k} [\partial_{x_j} v]^T D^{j,k} \partial_{x_k} v + \int \sum_{j,k} v^T [\partial_{x_j} D^{j,k}] \partial_{x_k} v \right\} \\
 & \leq -2 \int \sum_{j,k} v^T [\partial_{x_j} D^{j,k}] \partial_{x_k} v \\
 & = - \int \sum_{j,k} v^T [\partial_{x_j} D^{j,k}] \partial_{x_k} v - \int \sum_{j,k} \left\{ \partial_{x_k} v^T \right\} [\partial_{x_j} D^{j,k}] v \\
 & = - \sum_{j,k} \int \partial_{x_k} \left\{ v^T [\partial_{x_j} D^{j,k}] v \right\} + \int v^T \left[\sum_{j,k} \partial_{x_k} \partial_{x_j} D^{j,k} \right] v \\
 & = \int v^T \left[\sum_{j,k} \partial_{x_k} \partial_{x_j} D^{j,k} \right] v \leq \left\| \sum_{j,k} \partial_{x_k} \partial_{x_j} D^{j,k} \right\|_{L^\infty} \|v\|_{L^2}^2 \\
 & \leq b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+2}}) \|v\|_{L^2}^2,
 \end{aligned} \tag{17}$$

where here and later P denotes a polynomial, which may be different in different occurrences.

To estimate the left side of (14), integrate by parts, use the fact that H^s is an algebra for $s \geq s_0$ to estimate the term in which two derivatives are applied to the $D^{j,k}$, use the symmetry of the matrices $D^{j,k}$ and the symmetry (5) of $D^{j,k}$ under interchange of the indices j, k and switch the labels j and k of the indices to write the term in which a time derivative is applied to $\partial_{x_j} v$ as half the sum of terms in which a time derivative is applied once to $\partial_{x_j} v$ and once to $\partial_{x_k} v$, and use the identity (16) to obtain a term in which a time derivative is applied to an entire integral minus a term in which both time derivatives are applied to the coefficients $D^{j,k}$. Upon estimating the last term in similar fashion to the previous term in which two derivatives were applied to $D^{j,k}$, this yields

$$\begin{aligned}
 & 2 \int \sum_{j,k} (\partial_t v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_j} \partial_{x_k} v \\
 &= -2 \int \sum_{j,k} (\partial_t \partial_{x_j} v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_k} v \\
 &\quad - 2 \int \sum_{j,k} (\partial_t v)^T (\partial_{x_j} \partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_k} v \\
 &\leq -2 \int \sum_{j,k} (\partial_t \partial_{x_j} v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_k} v \\
 &\quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+1}})(1 + \|u_t\|_{H^{s_0+1}}) \|v_t\|_{L^2} \|v\|_{H^1} \\
 &= - \int \sum_{j,k} (\partial_t \partial_{x_j} v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_k} v \\
 &\quad - \int \sum_{j,k} (\partial_{x_j} v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_t \partial_{x_k} v \\
 &\quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+1}})(1 + \|u_t\|_{H^{s_0+1}}) \|v_t\|_{L^2} \|v\|_{H^1} \\
 &= -\frac{d}{dt} \int \sum_{j,k} (\partial_{x_j} v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_k} v \\
 &\quad + \int \sum_{j,k} (\partial_{x_j} v)^T (\partial_t^2 D^{j,k}(t, x, u(t, x))) \partial_{x_k} v \\
 &\quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+1}})(1 + \|u_t\|_{H^{s_0+1}}) \|v_t\|_{L^2} \|v\|_{H^1} \\
 &\leq -\frac{d}{dt} \int \sum_{j,k} (\partial_{x_j} v)^T (\partial_t D^{j,k}(t, x, u(t, x))) \partial_{x_k} v \\
 &\quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0}})(1 + \|u_t\|_{H^{s_0}}^2 + \|u_{tt}\|_{H^{s_0}}) \|v\|_{H^1}^2 \\
 &\quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+1}})(1 + \|u_t\|_{H^{s_0+1}}) \|v_t\|_{L^2} \|v\|_{H^1}.
 \end{aligned}$$

The process of estimating the left side of (15) is similar, but since spatial rather than time derivatives are involved the term $-\int \partial_{x_t} [(\partial_{x_j} v)^T (\partial_{x_m} D^{j,k}(t, x, u(t, x))) \partial_{x_k} v]$ that vanishes is obtained instead of the time derivative of an integral, and spatial derivatives rather than time derivatives are applied to v and u . Hence the estimate on the right side of (15) is obtained. \square

Before stating the theorem on uniform bounds for (1) it is necessary to define the expression that will satisfy a differential inequality and verify that it is the square of a norm.

Definition 2.4 For any nonnegative integer r and function $u \in L^\infty$ define a weighted H^r norm by

$$\|v\|_{r,u} := \sqrt{\sum_{0 \leq |\alpha| \leq r} \int_{\mathbb{R}^d} (D^\alpha v)^T A^0(t, x, u) (D^\alpha v) dx}, \tag{18}$$

and for any $(u, u_t) \in L^\infty$ and any set $W := \{w_m\}_{m=0}^s$ of positive constants, define

$$\| \|v\| \|_{s,u,W} := \sqrt{Q_{s,u,W}(v, v_t, \dots, \partial_t^s v)}, \tag{19}$$

where

$$\begin{aligned} Q_{s,u,W}(v^{(0)}, \dots, v^{(s)}) := & \sum_{m=0}^{s-1} \|v^{(m)}\|_{s-1-m,u}^2 + \sum_{m=0}^s w_m \sum_{|\alpha|=s-m} \|D^\alpha v^{(m)}\|_{0,u}^2 \\ & + \sum_{m=0}^{s-1} (m+1)w_{m+1} \sum_{j,k=1}^d \sum_{|\gamma|=s-1-m} \int (\partial_{x_j} D^\gamma v^{(m)})^T [\partial_t D^{j,k}] \partial_{x_k} D^\gamma v^{(m)}. \end{aligned} \tag{20}$$

Although the inclusion of the term involving $\partial_t D^{j,k}$ in (20) may seem strange, that term will arise naturally in the proof of Theorem 2.7 on account of the use of estimate (14).

Lemma 2.5 *Let s be a positive integer, and assume that the positivity condition (8) holds and the boundedness conditions (6:s)–(7:s) hold for that value of s . Assume in addition that*

$$\|u\|_{L^\infty} \leq M_1 \quad \text{and} \quad \|u_t\|_{L^\infty} \leq M_2, \tag{21}$$

and that $W := \{w_m\}_{m=0}^s$ are positive constants satisfying

$$\begin{aligned} \max_{0 \leq m \leq s-1} \frac{(m+1)w_{m+1}}{w_m} & \leq \frac{\mu}{d^2(1+NM_2)b(M_1)} \\ \text{for some } \mu \text{ satisfying } 0 & < \mu < \delta(M_1), \end{aligned} \tag{22}$$

where N denotes the number of components of the vector u , d is the number of spatial variables, and $\delta(\cdot)$ and $b(\cdot)$ are defined in Assumption 2.1. Then

1. $\|\cdot\|_{s,u}$ is a norm on H^s equivalent to the standard H^s norm; specifically,

$$\delta(M_1)\|v\|_{H^s} \leq \|v\|_{s,u} \leq b(M_1)\|v\|_{H^s}. \tag{23}$$

2. $\sqrt{Q_{s,u,W}(v^{(0)}, \dots, v^{(s)})}$ is a norm on $\prod_{s=0}^m H^{s-m}$ equivalent to sum of the standard H^{s-m} norms of the components. Quantitatively,

$$\begin{aligned} Q_{s,u,W}(v^{(0)}, \dots, v^{(s)}) \geq & \delta(M_1) \sum_{m=0}^{s-1} \|v^{(m)}\|_{H^{s-1-m}}^2 + \delta(M_1)w_s \|v^{(s)}\|_{L^2}^2 \\ & + (\delta(M_1) - \mu) \sum_{m=0}^{s-1} w_m \sum_{|\alpha|=s-m} \|D^\alpha v^{(m)}\|_{L^2}^2 \end{aligned} \tag{24}$$

and

$$\begin{aligned} Q_{s,u,W}(v^{(0)}, \dots, v^{(s)}) \leq & b(M_1) \sum_{m=0}^{s-1} \|v^{(m)}\|_{H^{s-1-m}}^2 + b(M_1)w_s \|v^{(s)}\|_{L^2}^2 \\ & + (b(M_1) + \mu) \sum_{m=0}^{s-1} w_m \sum_{|\alpha|=s-m} \|D^\alpha v^{(m)}\|_{L^2}^2. \end{aligned} \tag{25}$$

Proof The bounds (23) follow directly from (21), (8), and (6:s). Straightforward estimation of the term involving $D^{j,k}$ in (20) using (21) and (7:s) shows that

$$\begin{aligned} & \sum_{j,k=1}^d \sum_{|\gamma|=s-1-m} \left| \int (\partial_{x_j} D^\gamma v^{(m)})^T [\partial_t D^{j,k}] \partial_{x_k} D^\gamma v^{(m)} \right| \\ & \leq d^2 \|\partial_t D^{j,k}\|_{L^\infty} \sum_{|\alpha|=s-m} \|D^\alpha v^{(m)}\|_{L^2}^2 \\ & \leq d^2(1 + NM_2)b(M_1) \sum_{|\alpha|=s-m} \|D^\alpha v^{(m)}\|_{L^2}^2, \end{aligned} \tag{26}$$

and (24)–(25) follow directly from (26), formula (20) for $Q_{s,u,W}$, and the bounds (22), (8), and (6:s). □

In order to simplify the statement of the theorem, it will be convenient to state and prove a lemma estimating $Q_{u,s,W}(u, u_t, \dots, \partial_t^s u)$ in terms of $\|u\|_{H^{2s}}$ when u is a solution of (1). This result is only used to bound the initial value of Q .

Lemma 2.6 *Let the assumptions of Lemma 2.5 hold, and assume that $s \geq s_0$, where s_0 is defined in (4). Then there exists a continuous increasing function C_{init} such that any solution u of (1) satisfies*

$$\| \|u\|_{u,s,W}^2 \leq C_{\text{init}}(\|u\|_{H^{2s}}). \tag{27}$$

Proof Since H^r is an algebra for $r \geq s_0$, solving the PDE (1) for u_t , and then repeatedly differentiating with respect to t , substituting into the result the formulas for lower-order time derivatives obtained at earlier stages, and estimating an appropriate norm of the result yields a bound $\sum_{m=0}^s \|\partial_t^m u\|_{H^{s-m}} \leq C(\|u\|_{H^{2s}})$, because each time derivative adds at most two spatial derivatives. Combining this with (25) yields (27). □

The theorem on bounds for (1) includes two parts: an estimate for $Q_{s,u,W}$ and an estimate for $\|u\|_{s,u}^2$ when the equation satisfies additional conditions that make it possible to eliminate u_t from that estimate. In the latter case an estimate for u_t can then be obtained by solving the PDE for u_t and estimating the expression so obtained. Since that expression contains second derivatives of u , the Sobolev index of the bound for u_t for that case is two less than the index of the bound for u . Although one might expect that should hold for the first case as well, the expression $Q_{s,u,W}$ actually includes only one less derivative of u_t than of u . The reason is that the estimate for derivatives of order s of u involves derivatives of order $s - 1$ of u_t , so it is necessary to estimate u_t in a Sobolev space having index only one less than the index for u . As a consequence, in order to obtain an H^s estimate for u its initial data must belong to H^{2s} , in accordance with the estimate (27).

Theorem 2.7 *Let the spatial domain be the torus \mathbb{T}^d or the whole space \mathbb{R}^d . Let s be an integer satisfying*

$$s \geq s_0 + 2, \tag{28}$$

where s_0 is defined in (4). Suppose that Assumption 2.1 holds, with (6:s)–(7:s) holding for the given value of s , and that $\sum_{m=0}^s \|\partial_t^m F(t, x, 0)\|_{C^0([0,\infty); H^{s-m})}$ is finite.

1. Let m_1, m_2 , and m_3 be arbitrary positive constants, and let M_1, M_2 , and M_3 satisfy

$$M_1 > m_1, \quad M_2 > m_2, \quad \text{and} \quad M_3 > C_{\text{init}}(m_3). \tag{29}$$

Then there exist positive weights $W := \{w_m\}_{m=0}^s$ satisfying (22), a continuously differentiable nondecreasing function G , and a positive time T such that every solution u of (1) that satisfies

$$\|u(0, \cdot)\|_{L^\infty} \leq m_1, \quad \|u_t(0, \cdot)\|_{L^\infty} \leq m_2, \quad \text{and} \quad \|u(0, x)\|_{H^{2s}} \leq m_3, \quad (30)$$

and belongs to $\cap_{m=0}^s C^m([0, T_u]; H^{s-m})$ also satisfies the estimates

$$\|u\|_{L^\infty} \leq M_1 \quad \text{and} \quad \|u_t\|_{L^\infty} \leq M_2$$

from (21) and the differential and integrated energy estimates

$$\frac{d}{dt} \| \|u(t, \cdot)\|_{s,u,W}^2 \leq G(\| \|u(t, \cdot)\|_{s,u,W}^2) \quad (31)$$

and

$$\| \|u\|_{s,u,W}^2 \leq U(t, C_{\text{init}}(m_3)) \leq M_3 \quad (32)$$

for $0 \leq t \leq \min(T, T_u)$, where $U(t, U_0)$ is the unique solution of the ODE $U' = G(U)$ satisfying $U(0) = U_0$. The constant T , the weights W , and the function G are all independent of the particular solution u satisfying the above conditions.

2. Assume that in addition either the PDE (1) is scalar or the matrix A^0 is independent of x and u . Let m_1 and m_3 be arbitrary constants, and let M_1 and M_3 satisfy $M_1 > m_1$, and $M_3 > b(M_1)^2 m_3^2$. Then there exist a constant M_2 , a continuous nondecreasing function \tilde{G} , and a positive time \tilde{T} such that any sufficiently smooth solution u of (1) that satisfies

$$\|u(0, \cdot)\|_{L^\infty} \leq m_1 \quad \text{and} \quad \|u(0, x)\|_{H^s} \leq m_3 \quad (33)$$

and belongs to $\cap_{m=0}^s C^m([0, T_u]; H^{s-m})$ also satisfies the estimates

$$\|u\|_{L^\infty} \leq M_1 \quad \text{and} \quad \|u_t\|_{L^\infty} \leq M_2$$

from (21) and the estimates

$$\frac{d}{dt} \| \|u(t, \cdot)\|_{s,u}^2 \leq \tilde{G}(\| \|u(t, \cdot)\|_{s,u}^2) \quad (34)$$

and

$$\| \|u\|_{s,u}^2 \leq \tilde{U}(t, b(m_1)^2 m_3^2) \leq M_3 \quad (35)$$

for $0 \leq t \leq \min(\tilde{T}, T_u)$, where $\tilde{U}(t, U_0)$ is the unique solution of the ODE $\tilde{U}' = \tilde{G}(\tilde{U})$ satisfying $\tilde{U}(0) = U_0$. The constants M_2 and \tilde{T} and the function \tilde{G} are independent of the particular solution u satisfying the above conditions.

Proof Let the assumptions of the first part hold. The standard method of mollifying u allows us to assume that as many derivatives of it as needed exist and belong to L^2 , which justifies the calculations below that involve placing more than s derivatives on u or use the fact that the integral of a spatial derivative of any expression vanishes even when that expression contains a derivative of order $s + 1$ of u .

The difference $F(t, x, u) - F(t, x, 0)$ can be written as $\int_0^1 \partial_r F(t, x, ru) dr$, and the derivative inside the integral equals $u \cdot \frac{\partial F}{\partial u}(t, x, ru)$, which yields the identity

$$F(t, x, u) = F(t, x, 0) + H(t, x, u)u, \quad \text{where} \quad H(t, x, u) := \int_0^1 \left(\frac{\partial F}{\partial u}(t, x, ru) \right)^T dr. \quad (36)$$

Substitute (36) into (1) and take m time derivatives and a spatial derivative D^α of the result, with $0 \leq m + |\alpha| \leq s$. Rewrite the expression so obtained as the sum of the terms in which all derivatives are applied to the derivatives of u already appearing in (1) plus commutator terms, multiply the result on the left by $2(D^\alpha \partial_t^m u)^T$, integrate over the spatial domain, and use the symmetry of the A^j to obtain

$$2 \int v^T A(Dv) = \int v^T A(Dv) + \int (Dv)^T Av = \int D(v^T Av) - v^T (DA)v, \tag{37}$$

where $v = D^\alpha \partial_t^m u$, A is any of the matrices A^j for $0 \leq j \leq d$, and the operator D is ∂_t if $j = 0$ and is ∂_{x_j} if $j > 0$. After noting that $\int \partial_{x_j}(v^T A^j v) = 0$ by the periodicity of u and its derivatives or their decay at infinity, this yields

$$\begin{aligned} & \frac{d}{dt} \int (D^\alpha \partial_t^m u)^T A^0 (D^\alpha \partial_t^m u) \\ &= \int (D^\alpha \partial_t^m u)^T \left[\partial_t A^0 + \sum_{j=1}^d \partial_{x_j} A_j \right] (D^\alpha \partial_t^m u) + 2 \int (D^\alpha \partial_t^m u)^T (D^\alpha \partial_t^m F(t, x, 0)) \\ &+ 2 \int (D^\alpha \partial_t^m u)^T (D^\alpha \partial_t^m (Hu)) + 2 \int \sum_{j,k} (D^\alpha \partial_t^m u)^T D^{j,k} \partial_{x_j} \partial_{x_k} (D^\alpha \partial_t^m u) \\ &- 2 \int (D^\alpha \partial_t^m u)^T \left\{ [D^\alpha \partial_t^m, A^0]u_t + \sum_j [D^\alpha \partial_t^m, A^j]u_{x_j} - \sum_{j,k} [D^\alpha \partial_t^m, D^{j,k}] \partial_{x_j} \partial_{x_k} u \right\}. \end{aligned} \tag{38}$$

By the assumption on $F(t, x, 0)$, the Cauchy-Schwartz inequality and the elementary estimate $2ab \leq a^2 + b^2$ can be used to obtain

$$2 \int (D^\alpha \partial_t^m u)^T D^\alpha \partial_t^m F(t, x, 0) \leq 2 \|\partial_t^m u\|_{H^{|\alpha|}} \|\partial_t^m F(t, x, 0)\|_{H^{|\alpha|}} \leq c + \|\partial_t^m u\|_{H^{|\alpha|}}^2. \tag{39}$$

All the remaining terms on the right side of (38) have the form

$$c \int (D^\alpha \partial_t^m u_i) g(t, x, u) \prod_{\ell} D^{\alpha_\ell} \partial_t^{m_\ell} u_{i_\ell} \tag{40}$$

where u_i is a component of u and g is some component of some derivative of a coefficient $A^0, A^j, D^{j,k}$, or H . Because each term of the PDE (1) contains at most two derivatives, and at most s derivatives were applied to that equation,

$$\sum_{\ell} |\alpha_\ell| + m_\ell \leq s + 2. \tag{41}$$

The terms in (40) that satisfy in addition

$$|\alpha_\ell| + m_\ell \leq s \quad \text{for all } \ell \tag{42}$$

will be estimated using Lemma A.1. Since no derivatives are applied to the coefficients in the original PDE, the factors g in (40) involve at most derivatives of order s of the coefficients. Hence by the assumed bound (6:s)–(7:s) for the coefficients and the Cauchy-Schwartz

inequality,

[sum of all terms on right side of (38) having form (40) with (42) holding]

$$\begin{aligned} &\leq \sum_{\text{terms (40) satisfying (42)}} c \|g\|_{L^\infty} \int |D^\alpha \partial_t^m u_i| \left| \prod_\ell [D^{\alpha_\ell} \partial_t^{m_\ell} u_{i_\ell}] \right| \tag{43} \\ &\leq \sum_{\{i_\ell\}, \{m_\ell\}, \{\alpha_\ell\} \text{ satisfying (41), (42)}} c b(\|u\|_{L^\infty}) \|D^\alpha \partial_t^m u\|_{L^2} \left\{ \int \prod_\ell [D^{\alpha_\ell} \partial_t^{m_\ell} u_{i_\ell}]^2 \right\}^{1/2} \end{aligned}$$

Since the conditions (41)–(42) hold and s satisfies (28), applying Lemma A.1 with $r := 2$ shows that for all the terms appearing in (43)

$$\left[\int \prod_\ell [D^{\alpha_\ell} \partial_t^{m_\ell} u_{i_\ell}]^2 \right]^{1/2} \leq \prod_\ell \|\partial_t^{m_\ell} u\|_{H^{s-m_\ell}}. \tag{44}$$

Substituting (44) back into (43) yields the estimate

$$\begin{aligned} &\text{[sum of all terms on right side of (38) having form (40) with (42) holding]} \\ &\leq c b(\|u\|_{L^\infty}) \|\partial_t^m u\|_{H^{|\alpha|}} \sum_{\{m_\ell\} \text{ satisfying } m_\ell \leq s, \sum m_\ell \leq s+1} \prod_\ell \|\partial_t^{m_\ell} u\|_{H^{s-m_\ell}}, \tag{45} \end{aligned}$$

where the restriction $\sum m_\ell \leq s + 1$ in the final sum comes from the fact that only one time derivative appears in the PDE and at most s were applied to that equation.

The remaining terms on the right side of (38) have at least $s + 1$ derivatives applied to some occurrence of u , and all such terms involve the viscosity matrix $D^{j,k}$ since the presence of the commutators in the terms involving A^0 and A^j prevents all $s + 1$ derivatives in those terms from being applied to one factor. To treat the term in which $D^{j,k}$ appears undifferentiated, apply (13) with $v := D^\alpha \partial_t^m u$, which yields

$$2 \int \sum_{j,k} (D^\alpha \partial_t^m u)^T D^{j,k} \partial_{x_j} \partial_{x_k} (D^\alpha \partial_t^m u) \leq b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+2}}) \|\partial_t^m u\|_{H^{s-m}}^2. \tag{46}$$

The other term involving $D^{j,k}$ is the commutator term, which has the form

$$2 \sum_{j,k} \sum_{\substack{0 \leq l \leq m, 0 \leq \beta_i \leq \alpha_i \\ \ell + |\beta| \geq 1}} \binom{m}{\ell} \binom{\alpha}{\beta} \int (D^\alpha \partial_t^m u)^T (\partial_t^\ell D^\beta D^{j,k}) \partial_t^{m-\ell} D^{\alpha-\beta} \partial_{x_j} \partial_{x_k} u, \tag{47}$$

where $\binom{\alpha}{\beta}$ is the product binomial coefficient $\prod_{j=1}^d \binom{\alpha_j}{\beta_j}$. Since $\ell + |\beta| \leq m + |\alpha| \leq s$, the terms for which $m - \ell + |\alpha| - |\beta| + 2 \leq s$ have already been estimated in (45), so only the cases for which

$$m + |\alpha| - [\ell + |\beta|] \geq s - 1 \tag{48}$$

need be considered. Since $m + |\alpha| \leq s$ and $\ell + |\beta| \geq 1$, (48) can only hold when $m + |\alpha| = s$ and $\ell + |\beta| = 1$, i.e., when a total of s derivatives have been applied to the PDE and only one of those derivatives is applied to the coefficient $D^{j,k}$. When the derivative applied to $D^{j,k}$ is

a spatial derivative, use (15) with $v := D^{\alpha-\beta} \partial_t^m u$, which yields

$$\begin{aligned}
 & 2 \sum_{j,k} \sum_{\substack{0 \leq \beta_i \leq \alpha_i \\ |\beta|=1}} \binom{\alpha}{\beta} \int (D^\alpha \partial_t^m u)^T (D^\beta D^{j,k}) \partial_{x_j} \partial_{x_k} D^{\alpha-\beta} \partial_t^m u \\
 & \leq b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+2}}) \|\partial_t^m u\|_{H^{|\alpha|-m}}^2.
 \end{aligned}
 \tag{49}$$

When the derivative applied to $D^{j,k}$ is a time derivative, use (14) with $v := D^\alpha \partial_t^{m-1} u$, which yields

$$\begin{aligned}
 & 2m \sum_{j,k} \int (D^\alpha \partial_t^m u)^T (\partial_t D^{j,k}) \partial_{x_j} \partial_{x_k} D^\alpha \partial_t^{m-1} u \\
 & \leq -m \sum_{j,k} \frac{d}{dt} \int (\partial_{x_j} D^\alpha \partial_t^{m-1} u)^T (\partial_t D^{j,k}) \partial_{x_k} D^\alpha \partial_t^{m-1} u \\
 & \quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0+1}}) (1 + \|u_t\|_{H^{s_0+1}}) \|\partial_t^m u\|_{H^{|\alpha|}} \|\partial_t^{m-1} u\|_{H^{|\alpha|+1}} \\
 & \quad + b(\|u\|_{L^\infty}) P(\|u\|_{H^{s_0}}) (1 + \|u_t\|_{H^{s_0}}^2 + \|u_{tt}\|_{H^{s_0}}) \|\partial_t^{m-1} u\|_{H^{|\alpha|+1}}^2.
 \end{aligned}
 \tag{50}$$

As noted above, only the case when $|\alpha| + m = s$ needs to be estimated by (50).

Now pick μ satisfying $0 < \mu < \delta(M_1)$. Then define $w_0 := 1$ and successively choose the w_m for $1 \leq w_m \leq s$ satisfying (22) for the given values of M_1, M_2 , and μ , so the conclusions of Lemma 2.5 will hold for as long as (21) holds. Multiply (38) by w_m if $|\alpha| + m = s$ and by 1 otherwise, add the resulting equations, use the estimates (39), (45), (46), (49), and (50) to estimate the right side of the result, and move the time derivative terms arising from (50) to the left side. By the definition (19)–(20), the left side of the combined estimate is the time derivative of $\|u\|_{s,u,W}^2$. Hence, after replacing $b(\|u\|_{L^\infty})$ by its upper bound $b(M_1)$ and using the elementary estimate $x \leq 1 + x^2$ to eliminate odd powers of norms, the combined estimate has the form

$$\frac{d}{dt} \|u\|_{s,u,W}^2 \leq b(M_1) P(\sum_{m=0}^s \|\partial_t^m u\|_{H^{s-m}}^2).
 \tag{51}$$

The right side of (24) with $v^{(m)}$ replaced by $\partial_t^m u$ is bounded from above and below by constants times $\sum_{m=0}^s \|\partial_t^m u\|_{H^{s-m}}^2$. Hence by Lemma 2.5 and the definition (19) of $\|u\|_{s,u,W}$, the right side of (51) can be bounded by a polynomial of $\|u\|_{s,u,\varepsilon,W}^2$ as long as (21) holds. Once (21) has been shown to hold for some positive time this will yield (31).

By the comparison principle for ODEs, also known as the fence theorem [18, Theorem 4.7.1], as long as (21) holds then (31) together with (27) and (30) implies the first inequality in (32). In view of the last condition in (29), there exists a time T_3 such that $U(t, C_{\text{init}}(m_3)) \leq M_3$ for $0 \leq t \leq T_3$, which is the second inequality in (32). Moreover, the bounds

$$\|u\|_{L^\infty} \leq m_1 + t \sup_{[0,t]} \|u_t\|_{L^\infty} \quad \text{and} \quad \|u_t\|_{L^\infty} \leq m_2 + ct \sup_{[0,t]} \|u\|_{s,u,W}
 \tag{52}$$

yield times $T_1 := \frac{M_1 - m_1}{M_2}$ and $T_2 := \frac{M_2 - m_2}{c\sqrt{M_3}}$ such that

$$\|u_t\|_{L^\infty} \leq M_2 \quad \text{on the time interval } 0 \leq t \leq \min(T_2, T_u) \text{ for as long as (32) holds,}
 \tag{53}$$

$$\|u\|_{L^\infty} \leq M_1 \quad \text{on the time interval } 0 \leq t \leq \min(T_1, T_u) \text{ for as long as (53) holds,}
 \tag{54}$$

where T_u is the time of existence of u defined in the statement of the theorem. Since the estimates (52) and (32) imply bounds smaller than (M_1, M_2, M_3) for

$$(\|u\|_{L^\infty}, \|u_t\|_{L^\infty}, \|u\|_{s,u,\varepsilon,W}^2)
 \tag{55}$$

for times less than the minimum of $T := \min(T_1, T_2, T_3)$ and T_u as long as the bounds (M_1, M_2, M_3) hold, the continuity of the expressions (55) implies that the bounds (M_1, M_2, M_3) , and hence also (31)–(32), remain valid on $[0, \min(T, T_u)]$.

Now assume that the additional hypothesis of the second part of the theorem holds. In this case we take only spatial derivatives of the PDE (1), so (38) is replaced by

$$\begin{aligned} & \frac{d}{dt} \int (D^\alpha u)^T A^0 (D^\alpha u) \\ &= \int (D^\alpha u)^T \left[\partial_t A^0 + \sum_{j=1}^d \partial_{x_j} A_j \right] (D^\alpha u) + 2 \int (D^\alpha u)^T (D^\alpha F(t, x, 0)) \\ & \quad + 2 \int (D^\alpha u)^T (D^\alpha (Hu)) + 2 \int (D^\alpha u)^T D^{j,k} \partial_{x_j} \partial_{x_k} (D^\alpha u) \\ & \quad - 2 \int (D^\alpha u) \cdot \left\{ [D^\alpha, A^0] u_t + \sum_j [D^\alpha, A^j] u_{x_j} - \sum_{j,k} [D^\alpha, D^{j,k}] \partial_{x_j} \partial_{x_k} u \right\}. \end{aligned} \tag{56}$$

The terms on the right side of (56) that do not contain the time derivative u_t can be estimated in the same way as was done for (38), which yields a bound $C(\|u\|_{H^s})$ since no time derivatives are present. When A^0 is independent of x and v then $[D^\alpha, A^0]$ vanishes, so there is no term on the right of (56) that involves a time derivative of u . When the equation is scalar, then modify (56) by using the PDE (1) and the identity (36) to replace $[D^\alpha, A^0] u_t$ in (56) with the equal expression

$$[D^\alpha, A^0] \left\{ (A^0)^{-1} \left(\sum_{j,k} D^{j,k} \partial_{x_j} \partial_{x_k} u + F(t, x, 0) + H(t, x, u)u - \sum_j A^j u_{x_j} \right) \right\}. \tag{57}$$

Since (57) is only used to eliminate u_t when the equation is scalar, the term $[D^\alpha A^0] (A^0)^{-1} \partial_{x_j} D^{j,k}$ appearing in (57) is a 1×1 matrix, and so is automatically symmetric. Hence the term in the modified (56) in which that expression appears can be estimated in the same way that the term in (38) containing $[D^\alpha, D^{j,k}] \partial_{x_j} \partial_{x_k} u$ was estimated. The terms in (56) arising from the remaining terms in (57) can also be estimated in the same manner as for the terms in (38). Since no time derivatives appear, summing the result over $0 \leq |\alpha| \leq s$ and proceeding in similar fashion to the general case yields (34) and (35). In particular, (52) is replaced by

$$\|u_t\|_{L^\infty} \leq c \|u_t\|_{H^{s_0}} \leq C(\|u\|_{H^{s_0+2}}) \leq C \left(\frac{\|u\|_{s,u}}{\delta(M_1)} \right) \leq M_2 := C \left(\frac{\sqrt{M_3}}{\delta(M_1)} \right),$$

where u_t has been estimated in terms of u in similar fashion to (27), and

$$\|u\|_{L^\infty} \leq m_1 + t \sup_{[0,t]} \|u_t\|_{L^\infty} \leq m_1 + ct M_2.$$

□

3 Uniform H^s bounds and weighted time-derivative bounds for singular limit equations

For the PDE (2) the norm (18) is naturally replaced by

$$\|v\|_{r,u,\varepsilon} := \sqrt{\sum_{0 \leq |\alpha| \leq r} \int_{\mathbb{R}^d} (D^\alpha v)^T A^0(\varepsilon u) (D^\alpha u) dx} \tag{58}$$

In order to obtain uniform spatial bounds for solutions of (2), the expression Q whose time derivative is estimated must be modified to include powers of ε in the terms involving time derivatives. Just as for hyperbolic singular limits, for general initial data one factor of ε is applied to each time derivative, so $\partial_t^m u$ is multiplied by ε^m . For hyperbolic singular limits the well-prepared case is treated by simply replacing that factor ε^m by $\varepsilon^{\max(m-1,0)}$, but this does not quite work for the PDE (2), because the energy equation (38) for the case when $m = 1$ and $|\alpha| = s - 1$ would not be multiplied by a power of ε yet the term (50) of its estimate contains a term with a factor u_{tt} that needs to be multiplied by ε to be uniformly bounded. Hence the term $\sum_{|\alpha|=s-1} \|D^\alpha u_t\|_{0,u,\varepsilon}^2$ in Q must be multiplied by ε , which will necessitate a more careful treatment of some of the estimates. Thus, for general initial data (19)–(20) are replaced by

$$\|v\|_{s,u,\varepsilon,W} := \sqrt{Q_{s,u,\varepsilon,W}(v, v_t, \dots, \partial_t^s v)}, \tag{59}$$

where

$$\begin{aligned} Q_{s,u,\varepsilon,W}(v^{(0)}, \dots, v^{(s)}) &:= \sum_{m=0}^{s-1} \varepsilon^{2m} \|v^{(m)}\|_{s-1-m,u,\varepsilon}^2 \\ &+ \sum_{m=0}^s w_m \varepsilon^{2m} \sum_{|\alpha|=s-m} \|D^\alpha v^{(m)}\|_{0,u,\varepsilon}^2 \\ &+ \sum_{m=0}^{s-1} (m+1) w_{m+1} \varepsilon^{2(m+1)} \sum_{j,k=1}^d \sum_{|\gamma|=s-1-m} \int (\partial_{x_j} D^\gamma v^{(m)})^T [\partial_t D^{j,k}] \partial_{x_k} D^\gamma v^{(m)}, \end{aligned} \tag{60}$$

while for well prepared initial data

$$\begin{aligned} \|v\|_{s,u,\varepsilon,W,\text{alt}} &:= \sqrt{Q_{s,u,\varepsilon,W,\text{alt}}(v, v_t, \dots, \partial_t^s v)}, \tag{61} \\ Q_{s,u,\varepsilon,W,\text{alt}}(v^{(0)}, \dots, v^{(s)}) &:= \|v^{(0)}\|_{s-1,u,\varepsilon}^2 + \sum_{m=1}^{s-1} \varepsilon^{2(m-1)} \|v^{(m)}\|_{s-1-m,u,\varepsilon}^2 \\ &+ w_0 \sum_{|\alpha|=s} \|D^\alpha v^{(0)}\|_{0,u,\varepsilon}^2 + \varepsilon w_1 \sum_{|\alpha|=s-1} \|D^\alpha v^{(1)}\|_{0,u,\varepsilon}^2 \\ &+ \sum_{m=2}^s w_m \varepsilon^{2(m-1)} \sum_{|\alpha|=s-m} \|D^\alpha v^{(m)}\|_{0,u,\varepsilon}^2 \\ &+ w_1 \varepsilon \sum_{j,k=1}^d \sum_{|\gamma|=s-1} \int (\partial_{x_j} D^\gamma v)^T [\partial_t D^{j,k}] \partial_{x_k} D^\gamma v \end{aligned}$$

$$+ \sum_{m=1}^{s-1} (m+1)w_{m+1}\varepsilon^{2m} \sum_{j,k=1}^d \sum_{\substack{|\gamma|=s-1-m}} \int (\partial_{x_j} D^\gamma v^{(m)})^T [\partial_t D^{j,k}] \partial_{x_k} D^\gamma v^{(m)} \quad (62)$$

will be used. Calculations similar to those in the proof of Lemma 2.5 show that if

$$\varepsilon \|u\|_{L^\infty} \leq M_1 \quad (63)$$

then (23) remains valid when $\|\cdot\|_{s,u}$ is replaced by $\|\cdot\|_{s,u,\varepsilon}$, and if in addition, for positive $(M_2, \varepsilon_0, \mu)$ satisfying $\mu < \delta(M_1)$,

$$\varepsilon \|u_t\|_{L^\infty} \leq M_2 \quad (64)$$

and

$$\max_{0 \leq m \leq s-1} \frac{(m+1)w_{m+1}}{w_m} \leq \frac{\mu}{d^2 \varepsilon_0 (\varepsilon_0 + NM_2) b(M_1)}, \quad (65)$$

then (24)–(25) remain valid for $0 < \varepsilon \leq \varepsilon_0$ when $Q_{s,u,W}$ is replaced by $Q_{s,u,\varepsilon,W}$ and $v^{(m)}$ for $0 \leq m \leq s$ is replaced by $\varepsilon^m v^{(m)}$ on the right sides of those inequalities. Similarly, if instead

$$\|u_t\|_{L^\infty} \leq M_2 \quad (66)$$

and

$$\begin{aligned} \max_{0 \leq m \leq s-1} \frac{\varepsilon_0^{p(m)} (m+1)w_{m+1}}{w_m} &\leq \frac{\mu}{d^2 (1 + NM_2) b(M_1)} \\ \text{with } p(m) &:= \begin{cases} 1 & m \leq 1 \\ 2 & m \geq 2 \end{cases}, \end{aligned} \quad (67)$$

then (24)–(25) remain valid for $0 < \varepsilon \leq \varepsilon_0$ when $Q_{s,u,W}$ is replaced by $Q_{s,u,\varepsilon,W, \text{alt}}$ and powers of ε matching the corresponding powers on the first two lines of (62) are inserted into the right sides of those inequalities.

In similar fashion to (27), there exist continuous increasing functions $C_{\text{init},\varepsilon}$, C_* , and $C_{\text{init},\varepsilon,\text{alt}}$ such that when Assumption 2.1, (63), (64), and (65) hold and $s \geq s_0$ defined in (4) then solutions of (2) satisfy

$$\|v\|_{s,u,\varepsilon,W}^2 \leq C_{\text{init},\varepsilon} (\|u\|_{H^{2s}}), \quad (68)$$

while if Assumption 2.1, (63), (66), and (67) hold then solutions of (2) satisfy

$$\|u_t\|_{H^{2(s-1)}} \leq C_* \left(\left\| \frac{1}{\varepsilon} \sum_j C^j \partial_{x_j} u \right\|_{H^{2(s-1)}}, \|u\|_{H^{2s}} \right) \quad (69)$$

and

$$\|v\|_{s,u,\varepsilon,W,\text{alt}}^2 \leq C_{\text{init},\varepsilon,\text{alt}} (\|u\|_{H^{2s}}, \|u_t\|_{H^{2(s-1)}}). \quad (70)$$

Like (27), (68) is obtained by solving the PDE for u_t and then taking time derivatives of the result and using the formulas obtained previously to express time derivatives in terms of u and its spatial derivatives. Similarly, (70) is obtained by refraining from substituting for u_t in terms of u but substituting for u_{tt} and higher time derivatives in terms of u and u_t , while (69) is obtained by solving (2) for u_t and estimating the result while keeping the large term intact.

The theorem for (2) includes three parts: The first part is an estimate for $Q_{s,u,\varepsilon,W}$ when u_t is not assumed to be uniformly bounded at time zero, and the second is an estimate for $Q_{s,u,\varepsilon,W,alt}$ assuming well-prepared initial data. The final part of the theorem is an estimate for $\|u\|_{s,u,\varepsilon}^2$ when the equation satisfies additional conditions that make it possible to eliminate u_t from that estimate. In that case, a uniform estimate for εu_t can then be obtained by solving the PDE for u_t and estimating the expression so obtained. If the initial data is well prepared, then a uniform estimate for u_t without a power of ε can be obtained by differentiating the PDE with respect to t and using the result to estimate $\|u_t\|_{s-2,u,\varepsilon}^2$ in similar fashion to the estimate for $\|u\|_{s,u,\varepsilon}^2$.

Theorem 3.1 *Let the spatial domain be the torus \mathbb{T}^d or the whole space \mathbb{R}^d . Let s be an integer satisfying $s \geq s_0 + 2$, where the Sobolev embedding index s_0 is defined in (4). Suppose also that Assumption 2.1 holds, with (6:s)–(7:s) holding for the given value of s , and that $\sum_{m=0}^s \|\partial_t^m F(t, x, 0)\|_{C^0([0,\infty); H^{s-m})}$ is finite.*

1. *Let ε_0, m_1 , and m_3 be arbitrary positive constants, and let M_1 and M_3 satisfy $M_1 > m_1$ and $M_3 > C_{init,\varepsilon}(m_3)$. Then there exist a positive M_2 , positive weights $W := \{w_m\}_{m=0}^s$ satisfying (65), a continuously differentiable nondecreasing function G , and a positive time T such that every solution u of (2) with $0 < \varepsilon \leq \varepsilon_0$ that satisfies*

$$\varepsilon \|u(0, \cdot)\|_{L^\infty} \leq m_1 \quad \text{and} \quad \|u(0, x)\|_{H^{2s}} \leq m_3 \tag{71}$$

and belongs to $\cap_{m=0}^s C^m([0, T_u]; H^{s-m})$ also satisfies

$$\varepsilon \|u\|_{L^\infty} \leq M_1 \quad \text{and} \quad \varepsilon \|u_t\|_{L^\infty} \leq M_2$$

from (63) to (64),

$$\frac{d}{dt} \| \|u(t, \cdot)\| \|_{s,u,\varepsilon,W}^2 \leq G(\| \|u(t, \cdot)\| \|_{s,u,\varepsilon,W}^2), \tag{72}$$

and

$$\| \|u\| \|_{s,u,\varepsilon,W}^2 \leq U(t, C_{init,\varepsilon}(m_3)) \leq M_3 \tag{73}$$

for $0 \leq t \leq \min(T, T_u)$, where $U(t, U_0)$ is the unique solution of the ODE $U' = G(U)$ satisfying $U(0) = U_0$. The constants M_2 and T , the weights W , and the function G are all independent of $\varepsilon \in (0, \varepsilon_0]$ and of the particular solution u satisfying the above conditions.

2. *Let ε_0, m_1, m_3 , and m_4 be arbitrary positive constants, and let M_1 and M_3 satisfy $M_1 > m_1$ and $M_3 > C_{init,\varepsilon,alt}(m_3, C_*(m_4, m_3))$. Then there exist a positive M_2 , positive weights $W := \{w_m\}_{m=0}^s$ satisfying (67), a continuously differentiable nondecreasing function G , and a positive time T such that every solution u of (2) with $0 < \varepsilon \leq \varepsilon_0$ that satisfies (71) and the well-preparedness condition*

$$\| \sum_j C^j \partial_{x_j} u(0, x) \|_{H^{2(s-1)}} \leq m_4 \varepsilon \tag{74}$$

and belongs to $\cap_{m=0}^s C^m([0, T_u]; H^{s-m})$ also satisfies

$$\varepsilon \|u\|_{L^\infty} \leq M_1 \quad \|u_t\|_{L^\infty} \leq M_2$$

from (63), (66),

$$\frac{d}{dt} \| \|u(t, \cdot)\| \|_{s,u,\varepsilon,W}^2 \leq G(\| \|u(t, \cdot)\| \|_{s,u,\varepsilon,W}^2), \tag{75}$$

and

$$\| \|u\| \|_{s,u,\varepsilon,W}^2 \leq U(t, C_{init,\varepsilon,alt}(m_3, C_*(m_4, m_3))) \leq M_3 \tag{76}$$

for $0 \leq t \leq \min(T, T_u)$, where $U(t, U_0)$ is the unique solution of the ODE $U' = G(U)$ satisfying $U(0) = U_0$. The constants M_2 and T , the weights W , and the function G are all independent of $\varepsilon \in (0, \varepsilon_0]$ and of the particular solution u satisfying the above conditions.

- Assume that in addition either the PDE (2) is scalar or the matrix A^0 is constant. Let m_1 and m_3 be arbitrary constants, and let M_1 and M_3 satisfy $M_1 > m_1$, and $M_3 > b(M_1)^2 m_3^2$. Then there exist a constant M_2 , a continuous nondecreasing function \tilde{G} , and a positive time \tilde{T} such that for any sufficiently smooth solution u of (2) that satisfies (71) and belongs to $\cap_{m=0}^s C^m([0, T_u]; H^{s-m})$, also satisfies

$$\varepsilon \|u\|_{L^\infty} \leq M_1 \quad \text{and} \quad \varepsilon \|u_t\|_{L^\infty} \leq M_2$$

from (63) to (64),

$$\frac{d}{dt} \|u(t, \cdot)\|_{s,u,\varepsilon}^2 \leq \tilde{G}(\|u(t, \cdot)\|_{s,u,\varepsilon}^2), \tag{77}$$

and

$$\|u\|_{s,u,\varepsilon}^2 \leq \tilde{U}(t, b(m_1)^2 m_3^2) \leq M_3 \tag{78}$$

for $0 \leq t \leq \min(\tilde{T}, T_u)$, where $\tilde{U}(t, U_0)$ is the unique solution of the ODE $\tilde{U}' = \tilde{G}(\tilde{U})$ satisfying $\tilde{U}(0) = U_0$. The constants M_2 and \tilde{T} and the function \tilde{G} are independent of the particular solution u satisfying the above conditions.

Proof Assume that the conditions of the first part hold. The derivation of the estimate (72) for solutions of (2) proceeds in similar fashion to the derivation of estimate (31) for solutions of (1). Since the matrices C^j are symmetric and constant, (37) implies that the terms involving them drop out of the energy estimates, so (38) still holds. The terms on the right side of (38) are estimated in the same way as before. This time choose the $w_m = w_m(M_2)$ to satisfy (65) for the given M_1 and a chosen μ as a function of M_2 that will be determined later. Multiply (38) by $w_m \varepsilon^{2m}$ if $|\alpha| + m = s$ and by ε^{2m} otherwise. As before, add the resulting equations, use the estimates shown in the proof of Theorem 2.7 to estimate the right side of the result, and move the time derivatives to the left side. On account of the powers of ε , this time the resulting left side is the time derivative of $\|u\|_{s,u,\varepsilon,W}^2$. The key point is that the total number of time derivatives of u appearing in each term on the right side is less than or equal to the power of ε that term is multiplied by, because every term in (38) either had $2m$ time derivatives or had $2m + 1$ time derivatives multiplied by a derivative of A^0 , which yields a factor of ε thanks to the assumed form of that matrix, so the multiplication by ε^{2m} made the powers of ε balance the number of time derivatives (some of which may be applied to the explicit time dependence of coefficients rather than to u). Hence instead of (51) we obtain, provided (63)–(64) hold,

$$\begin{aligned} \frac{d}{dt} \|u\|_{s,u,\varepsilon,W}^2 &\leq b(M_1) P(\sum_{m=0}^s \varepsilon^{2m} \|\partial_t^m u\|_{H^{s-m}}^2) \\ &\leq b(M_1) \tilde{P}_{M_1,\mu}(\|u\|_{s,u,\varepsilon,W}^2) := G(\|u\|_{s,u,\varepsilon,W}^2), \end{aligned}$$

which is (72), from which (73) follows for some time T_3 . To define M_2 so that (64) holds, note that

$$\varepsilon \|u_t\|_{L^\infty} \leq c\varepsilon \|u_t\|_{H^{s_0}} \leq \frac{c\varepsilon}{\delta(M_1)} \|u_t\|_{s-2,u} \leq \frac{c}{\delta(M_1)} \|u\|_{s,u,\varepsilon,W} \leq \frac{cM_3}{\delta(M_1)}$$

with the final constant c independent of the W and hence of M_2 , because the weights W do not appear in the terms of $\|u\|_{s,u,\varepsilon,W}$ involving at most $s - 1$ space and time derivatives.

Hence defining $M_2 := \frac{cM_3}{\delta(M_1)}$ ensures that (64) holds as long as (63) and (73) hold. Finally, since it is only necessary to bound $\varepsilon \|u\|_{L^\infty}$, in similar fashion to the proof of Theorem 2.7,

$$\varepsilon \|u\|_{L^\infty} \leq m_1 + \varepsilon t \sup_{[0,t]} \|u_t\|_{L^\infty} \leq m_1 + ctM_2,$$

so (63) holds for some positive time T_1 . Since that estimate and (73) show that $\varepsilon \|u\|_{L^\infty}$ and $\|u\|_{u,s,\varepsilon,W}^2$ satisfy smaller bounds for times, all the claimed results hold for the minimum of $T := \min(T_1, T_3)$ and T_u .

Now assume that the conditions of the second part hold. Proceed as for the first case, but multiply (38) by the factors of ε corresponding to those in definition (62). Since the terms were all estimated in the proof of Theorem 2.7, in order to obtain (75) it suffices to verify that the powers of ε balance appropriately. In the estimates (39), (46), and (49) the only expressions on the right sides involving time derivatives have the same number of time and space derivatives as appear on the left sides, so the powers of ε balance exactly. In (50) the factor of $\|\partial_t^{m-1} u\|_{H^{|\alpha|+1}}^2$ on the right side requires at least one power less of ε than the equation was multiplied by, because the fact that the estimate is only used when $|\alpha| + m = s$ implies that the equation was multiplied by ε when $m = 1$. Indeed, that factor is used in the terms involving spatial derivatives of order $s - 1$ of u_t precisely for this reason. This factor of ε balances the factor u_{tt} that is also present on the right side of (50), so that estimate is also ε -balanced. This leaves only (45), which is more complicated and will be separated into a number of cases. Note that the ratio of the power of ε to the number of time derivatives in the weights in $Q_{s,u,\varepsilon,W,alt}$ is nondecreasing as a function of the number of time derivatives, even when the weight of u_t is taken to be $\frac{1}{2}$ as it is in the terms with $s - 1$ space derivatives. Hence splitting the time derivative ∂_t^m over several factors of u never increases the power of ε needed to balance the result. That implies that the powers of ε in the terms in (45) arising from the A^j , H , or the $D^{j,k}$ balance properly when either $m \neq 1$ or $m = 1$ and $|\alpha| = s - 1$. The case when $m = 1$ and $|\alpha| < s - 1$ must be considered separately because the powers of ε only balance in that case if the number of spatial derivatives applied to u_t in the estimate is at most $s - 1$ so no power of ε is needed for balance. That indeed holds for the terms in (45) arising from the A^j or H since at most one derivative is added by the PDE and the derivatives are split over at least two factors. The terms in (45) arising from A^0 have one more time derivative and one less spatial derivative than the terms arising from A^j , but also have an extra power of ε since A^0 is always differentiated, and that ensures that the terms balance since adding a time derivative to $D^\gamma \partial_t^k u$ increases the required power of ε by at most one. This leaves to consider only the terms in (45) with $m = 1$ and $|\alpha| \leq s - 2$ that arise from $D^{j,k}$. For those terms, the product $g \prod D^{\alpha_\ell} \partial_t^{m_\ell} u_{i_\ell}$ appearing in (40) comes from $D^\alpha \partial_t (D^{j,k} \partial_{x_j} \partial_{x_k} u)$. Suppose that the time derivative is applied to $D^{j,k}$. If ∂_t is applied to the explicit time dependence in $D^{j,k}$ then there are no factors of u_t so no power of ε is needed. If ∂_t is applied to u in $D^{j,k}$ then define $v = \partial_{x_k} u$ and apply Lemma A.1 with $s_* := s - 1$ and $r := 1$ to obtain the bound $\|u_t\|_{H^{s_*-1}} \|v\|_{H^{s_*}} \|u\|_{H^{s_*}}^{L-2}$, which equals $\|u_t\|_{H^{s-2}} \|u\|_{H^s} \|u\|_{H^{s-1}}^{L-2}$ and so does not require a power of ε to balance it since the number of spatial derivatives applied to u_t is less than $s - 1$. The case when ∂_t is applied to $\partial_{x_j} \partial_{x_k} u$ but either $|\alpha| \leq s - 4$ or else $|\alpha| = s - 3$ and at least one spatial derivative is applied to $D^{j,k}$ can be estimated similarly. The case when $|\alpha| = s - 2$ and all derivatives are applied to $\partial_{x_j} \partial_{x_k} u$ is always estimated by (17) with $v := D^\alpha \partial_t^m u$ rather than by (45), since $D^\alpha \partial_{x_j} \partial_{x_k} \partial_t u$ has more than s derivatives applied to u . The case when $|\alpha| = s - 3$ and all derivatives are applied to $\partial_{x_j} \partial_{x_k} u$ can also be estimated by (17) with $v := D^\alpha \partial_t^m u$, and as already noted that estimate is ε -balanced. Finally, the case when $|\alpha| = s - 2$ and only one spatial derivative is applied to $D^{j,k}$ can be estimated by (49), which is also ε -balanced. Since all terms have now been estimated

in balanced fashion, choosing weights $w_m(M_2)$ to satisfy (67), adding the estimates, and moving the time derivative term from (50) to the left side of the result yields (75) provided (63) and (66) hold. Those bounds can be shown similarly to the proof of the first part.

Finally, let the additional conditions of the final part hold. Since A^0 in (2) depends on εu , and in the commutator term $[D^\alpha, A^0]u_t$ at least one derivative is applied to A^0 , that term contains a factor ε . Solving (2) for u_t and multiplying the result by ε yields an expression that is bounded uniformly in ε , so after substituting for εu_t the resulting terms can be estimated as for (1), which yields the desired bounds. Also, solving the PDE (2) for u_t and using the bound (78) to estimate the result yields the required bound (64) for $\varepsilon \|u_t\|_{L^\infty}$ after defining M_2 appropriately. \square

4 Existence and convergence results

Theorem 4.1 *Under the conditions of any part of Theorem 2.7 or Theorem 3.1, there exists a unique solution of the PDE (1) or (2), respectively, with a specified initial value $u(0, x) = u_0(x)$, for at least a time T or \tilde{T} as determined in those theorems, satisfying the bounds specified there.*

Idea of the proof The solution to the initial value problem for (1) can be obtained as the limit as $\delta \rightarrow 0+$ of the solution u^δ of

$$\begin{aligned} A^0(t, x, J_\delta u^\delta)(u^\delta)_t + J_\delta \sum_{j=1}^d A^j(t, x, J_\delta u)(J_\delta u)_{x_j} \\ = J_\delta \sum_{j,k=1}^d D^{j,k}(t, x, J_\delta u) \partial_{x_j} \partial_{x_k} (J_\delta u) + J_\delta [\phi_\delta F(t, x, J_\delta u)], \\ u^\delta(0, \cdot) = J_\delta u_0, \end{aligned}$$

with an analogous formula holding for (2). Here u_0 is the initial data for (1) or (2) that belongs to a Sobolev space with sufficiently high index as detailed in the assumptions of Theorem 2.7, J_δ is a symmetric spatial mollification operator tending to the identity as $\delta \rightarrow 0$, ϕ_δ is a smooth compactly-supported function equal to one for $\delta|x| \leq 1$ whose derivatives all tend uniformly to zero as $\delta \rightarrow 0$, and the mollification parameter has been denoted δ rather than ε as in [13, Section 5.2, Section 7.1] to avoid confusion with the singular perturbation parameter of (2). Since the mollified equations are ODEs in Hilbert spaces, the existence of solutions to the initial-value problem for them follows from the ODE existence theorem, for as long as the solutions remain finite in an appropriate norm.

The energy estimates derived in Theorems 2.7 and 3.1 remain valid for the mollified equations, since the spatial cutoff applied to the undifferentiated term ensures that it is rapidly decaying. The energy estimate for the mollified equations ensures that the solutions exist and are bounded in the appropriate norm for a time and with a bound independent of the mollification parameter δ and, for (2), also independent of the small parameter ε . As in [13, Section 5.2, Section 7.1], the convergence of u^δ to the solution u of (1) or (2) satisfying the same bound for the same time then follows provided that sufficiently smooth solutions of (1) and (2) having specified initial values are unique.

As for symmetric hyperbolic systems [19, Theorem 2.1] and many other PDEs, in order to prove the uniqueness of sufficiently smooth solutions to the initial-value problem for (1)

and (2) it suffices to obtain an L^2 energy estimate

$$\frac{d}{dt} \int v^T A^0 v \, dx \leq C \int v^T A^0 v \, dx \tag{79}$$

for the difference $v := u^{(1)} - u^{(2)}$ of two solutions having the same initial data, where C depends on a norm of $u^{(1)}$ and $u^{(2)}$ whose boundedness is known, since (79) plus the fact that $\int v^T A^0 v$ vanishes at time zero implies that $v \equiv 0$. The fact that the hyperbolic terms in the PDEs yield estimates of the form (79) is standard, so it suffices to estimate the second-order terms in (1) or (2). The difference $D^{j,k}(t, x, u^{(1)})\partial_{x_j}\partial_{x_k}u^{(1)} - D^{j,k}(t, x, u^{(2)})\partial_{x_j}\partial_{x_k}u^{(2)}$ can be written as the sum of the terms $[D^{j,k}(t, x, u^{(2)} + v) - D^{j,k}(t, x, u^{(2)})]\partial_{x_j}\partial_{x_k}u^{(1)}$ and $D^{j,k}(t, x, u^{(2)})\partial_{x_j}\partial_{x_k}v$. Since the first term is $O(|v|)$ and the estimates obtained for the $u^{(j)}$ bound their C^2 norms, the first term makes a contribution of the desired form to the L^2 energy estimate. By (17) with u replaced by $u^{(2)}$, the second term above also yields a contribution of the desired form. Hence (79) indeed holds. \square

Theorem 4.2 1. *If part 2 of Theorem 3.1 holds then as $\varepsilon \rightarrow 0$ the solution of (2) with initial data converging in H^s tends to the unique solution of the limit equation*

$$A^0(0)u_t + \sum_{j=1}^d A^j(t, x, u)u_{x_j} + \sum_{j=1}^d C^j V_{x_j} = \sum_{j,k=1}^d D^{j,k}(t, x, u)\partial_{x_j}\partial_{x_k}u + F(t, x, u), \tag{80}$$

$$\sum_{j=1}^d C^j u_{x_j} = 0$$

having the limit initial data, where V is the Lagrange multiplier variable that enforces the constraint in the second equation of (80).

2. *Suppose that part 1 or part 3 of Theorem 3.1 holds, that the spatial domain is periodic, and that the initial data converges to $u^0(x)$ in H^s as $\varepsilon \rightarrow 0$. As in [14, Section 2], make the following definitions: First normalize the matrix A^0 to satisfy $A^0(0) = I$ by making the change of variables $u \mapsto (A^0(0))^{-1/2}u$ and multiplying the resulting equation by $(A^0(0))^{-1/2}$. Let $\mathcal{S}(\tau)$ be the solution operator of the fast equation $u_\tau + \sum_{j=1}^d C^j u_{x_j} = 0$, which can be written in Fourier space as*

$$[\widehat{\mathcal{S}(\tau)u}](t, \tau, k) = e^{-i\tau \sum_{j=1}^d k_j C^j} \widehat{u}(t, \tau, k).$$

Define the averaging operator

$$[\mathbb{M}(g)](t, x) := \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau g(t, \tau_1, x) \, d\tau_1$$

and the projection operator

$$\mathcal{E} := \mathcal{S}(\tau)\mathbb{M}(\mathcal{S}(-\tau)\cdot),$$

which is the orthogonal projection with respect to the inner product $[u, v] := \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau u \cdot v \, d\tau_1$ onto terms $f(t, \tau, x)$ having the form $\mathcal{S}(\tau)g(t, x)$. Then

$$u(t, x) = v^0(t, \tau, x)|_{\tau=\frac{t}{\varepsilon}} + o(1) \quad \text{in } H^{s-1}, \tag{81}$$

where v^0 is the unique solution of the limit profile equations

$$v^0 = \mathcal{E}v^0, v^0(0, \tau, x) = \mathcal{S}(\tau)u^0(x),$$

$$v_t^0 + \mathcal{E} \left\{ (v^0 \cdot A_u^0(0))v_\tau^0 + \sum_{j=1}^d A^j(t, x, v^0)v_{x_j}^0 - \sum_{j,k=1}^d D^{j,k}(t, x, v^0)\partial_{x_j}\partial_{x_k}v^0 - F(t, x, v^0) \right\} = 0. \tag{82}$$

The proof of the first part of Theorem 4.2 is essentially the same as the proof of [20, Theorem 2], and the proof of the second part is essentially the same as the proof of the convergence part of [14, Theorem 2.1], since that proof uses only the uniform bounds, some general results on averaging, and the fact that an L^2 estimate for the difference of two solutions of the limit profile equations can be shown in similar fashion to the L^2 estimate for the difference of two solutions of the original system.

Acknowledgements This research was supported by the ISF-NSFC joint research program (grant No. 2519/17).

Appendix A: Calculus estimate

The estimate derived in [17, (3.13) through paragraph after (3.22)] is

$$\int \prod_{\ell=1}^L [D^{\alpha_\ell} \partial_t^{m_\ell} v^{(\ell)}]^2 \leq C \prod_{\ell=1}^L \|\partial_t^{m_\ell} v^{(\ell)}\|_{H^{s_0+1-m_\ell}}^2 \tag{A1}$$

provided that

$$m_\ell + |\alpha_\ell| \leq s_0 + 1 \quad \text{for all } \ell \quad \text{and} \quad \sum_{\ell=1}^L m_\ell + |\alpha_\ell| \leq s_0 + 2, \tag{A2}$$

where as usual $s_0 := \lfloor \frac{d}{2} \rfloor + 1$ is the Sobolev embedding index. Although (A1) is only stated there for the case when all the $v^{(\ell)}$ are equal, the proof there first estimates the left side of (A1) by using the multi-factor Hölder’s inequality $\int \prod_{\ell=1}^L [D^{\alpha_\ell} \partial_t^{m_\ell} v^{(\ell)}]^2 \leq \prod_{\ell=1}^L [\int |D^{\alpha_\ell} \partial_t^{m_\ell} v^{(\ell)}|^{2P_\ell}]^{\frac{1}{P_\ell}}$ for appropriate P_ℓ and then estimates each factor on the right side, so there is no need to assume that the $v^{(\ell)}$ are the same. In addition, it was assumed in [17] that $m_\ell + |\alpha_\ell| \geq 1$ and $L \geq 2$ because when applied to commutators those always hold, but any undifferentiated factors can be pulled out in L^∞ norm and then bounded by the H^{s_0} norm, and the case $L = 1$ is trivial on account of the first condition in (A2), so those assumptions have also been omitted.

The estimate (A1)–(A2) implies a more general estimate:

Lemma A.1 *Suppose that*

$$s_* \geq s_0 + r \quad \text{for some integer } r \geq 1, \quad \max_\ell (m_\ell + |\alpha_\ell|) \leq s_*,$$

$$\text{and} \quad \sum_{\{\ell \mid |\alpha_\ell| + m_\ell \geq s_* - s_0\}} m_\ell + |\alpha_\ell| \leq s_* + r. \tag{A3}$$

Then

$$\int \prod_{\ell=1}^L [D^{\alpha_\ell} \partial_t^{m_\ell} v^{(\ell)}]^2 \leq C \prod_{\ell=1}^L \|\partial_t^{m_\ell} v^{(\ell)}\|_{H^{s_*-m_\ell}}^2 \tag{A4}$$

Proof Pull out in L^∞ norm all factors that do not satisfy

$$|\alpha_\ell| + m_\ell \geq s_* - s_0, \tag{A5}$$

except that if no factors satisfy that condition then leave one factor

$$[D^{\alpha_\ell} \partial_t^{m_\ell} v^{(\ell)}]^2 \tag{A6}$$

inside the integral. The Sobolev embedding theorem ensures that the L^∞ norm of each factor pulled out can be estimated by the corresponding factor on the right side of (A4), and if only one factor remains inside the integral then the integral is the square of the L^2 norm of a factor (A6), which can be trivially estimated by its corresponding factor on the right side of (A4) in view of the middle assumption in (A3), whether (A5) holds or not. This leaves only the case when at least two factors remain inside the integral, all of which satisfy (A5). For each factor (A6) remaining in the integral divide α and k into parts such that

$$\alpha_\ell = \beta_\ell + \gamma_\ell \quad \text{and} \quad m_\ell = p_\ell + q_\ell, \quad \text{where} \quad |\gamma_\ell| + q_\ell = s_* - s_0 - 1 \tag{A7}$$

and hence

$$|\beta_\ell| + p_\ell \leq s_0 + 1 \tag{A8}$$

by the middle inequality in (A3). After defining

$$w^{(\ell)} = D^{\gamma_\ell} \partial_t^{q_\ell} v^{(\ell)} \tag{A9}$$

the left side of (A4) becomes

$$\int \prod_\ell [D^{\beta_\ell} \partial_t^{p_\ell} w^{(\ell)}]^2. \tag{A10}$$

By the last inequality in (A3), the last equation in (A7), the fact that at least two factors are present, and the first inequality in (A3),

$$\begin{aligned} \sum_\ell (|\beta_\ell| + p_\ell) &= \sum_\ell (|\alpha_\ell| + m_\ell) - \sum_\ell (|\gamma_\ell| + q_\ell) \\ &\leq s_* + r - 2(s_* - s_0 - 1) = s_0 + 2 + (s_0 + r - s_*) \leq s_0 + 2. \end{aligned} \tag{A11}$$

Since (A8) and (A11) say that (A2) holds, (A1) can be applied to (A10), which together with (A9) and (A7) yields

$$\begin{aligned} \int \prod_{\ell=1}^L [D^{\alpha_\ell} \partial_t^{m_\ell} v^{(\ell)}]^2 &= \int \prod_\ell [D^{\beta_\ell} \partial_t^{p_\ell} w^{(\ell)}]^2 \\ &\leq C \prod_\ell \|\partial_t^{p_\ell} w^{(\ell)}\|_{H^{s_0+1-p_\ell}}^2 = C \prod_\ell \|\partial_t^{p_\ell} D^{\gamma_\ell} \partial_t^{q_\ell} v^{(\ell)}\|_{H^{s_0+1-p_\ell}}^2 \\ &= C \prod_\ell \|\partial_t^{m_\ell} D^{\gamma_\ell} v^{(\ell)}\|_{H^{s_0+1-p_\ell}}^2 \leq C_1 \prod_\ell \|\partial_t^{m_\ell} v^{(\ell)}\|_{H^{s_0+1+|\gamma_\ell|-p_\ell}}^2 \\ &= C_1 \prod_\ell \|\partial_t^{m_\ell} v^{(\ell)}\|_{H^{s_0+1+|\gamma_\ell+q_\ell|-m_\ell}}^2 = C_1 \prod_\ell \|\partial_t^{m_\ell} v^{(\ell)}\|_{H^{s_*-m_\ell}}^2. \end{aligned}$$

□

Appendix B: Classical results for degenerate equations are inherently linear

Linear degenerate elliptic second-order equations and systems have the form

$$\sum_{j,k=d_0}^d A^{j,k}(x)u_{x_jx_k} + \sum_{j=d_0}^d B^j(x)u_{x_j} + C(x)u = F(x), \quad \text{where } d_0 \in \{0, 1\}. \quad (\text{B12})$$

Equation (B12) includes equations of the form (1) in which the coefficients do not depend on the dependent variable u , because the time variable t in (1) can be relabeled as x_0 and the non-strict Legendre-Hadamard condition

$$\sum_{j,k} \xi_j \xi_k w^T A^{j,k} w \geq 0 \quad \text{for all real numbers } \{\xi_j\} \text{ and vectors } w \quad (\text{B13})$$

assumed for the coefficients of the second-order terms in most of the work discussed below allows $A^{j,k}$ to vanish when at least one of j or k equals 0, i.e., allows the second-order terms to include only spatial derivatives.

The results for linear degenerate elliptic equations and systems that have at least a slight resemblance to results in this paper appeared in [6, Section 9.2], and are also presented in [8, Section 1.9]. In particular [6, Theorem 6, p. 486], an H^m space-time estimate is obtained for solutions to the initial-value problem for linear systems of the form (1), where the index m of the Sobolev space depends on the smoothness of the coefficients, inhomogeneous term, and initial data. The proof uses [6, Lemma 3.1, p. 460], whose proof is in turn based on the identities [6, (3.4)–(3.5)]. In those identities the operator applied to u on the left side of (B12) is expressed as the product M^*L of two first-order differential operators, whose coefficients therefore contain the coefficients appearing in (B12). The identity [6, (3.5)] that re-expresses two terms on the right side of [6, (3.4)] contains the double commutator $[[L, A], A]$, where in the proof of [6, Theorem 6] the operator A is a differential operator of order m . The analogous identity for the remaining two terms on the right side of [6, (3.4)], which is not written explicitly, contains the similar expression $[[M, A], A]$. Thus, the H^m estimate for the solution requires a C^{2m} estimate for the coefficients, or at the very least an H^{2m} estimate for those coefficients if estimates for Sobolev norms of products are used. Hence if the coefficients depended on the solution then the bound obtained for the H^m norm of the solution would depend on the H^{2m} norm of that solution, so no closed estimate is obtained in the nonlinear case. This shows that the limitation to linear systems is inherent in the results of [6].

Since [6, Lemma 3.1, p. 460] or its consequences are used to obtain the results in [7], the limitation to linear systems is inherent in those results as well. Similarly, the C^1 estimate of Oleinik (e.g. [8, Theorem 1.8.1, pp. 76-77]) requires a bound on the C^2 norm of the coefficients, which again restricts the result to linear equations. Both these results are also limited to single equations rather than systems. Earlier results of Fichera (e.g. [8, Section 1.2-3]) only obtained L^p bounds, and so are also limited to linear equations. Finally, although strictly parabolic systems of the form B12 can be expressed as symmetric positive linear differential equations [21, p. 338], the strict parabolicity is essential in order to satisfy the positivity requirement, so the results there do not apply to the degenerate equations considered in this paper.

I thank the anonymous reviewer of a previous version of this paper for pointing out some of the references discussed in this section.

References

1. Bonafede, S., Cirmi, G.R., Tedeev, A.F.: Finite speed of propagation for the porous media equation. *SIAM J. Math. Anal.* **29**(6), 1381–1398 (1998). <https://doi.org/10.1137/S0036141096298072>
2. Cieślak, T.: Quasilinear nonuniformly parabolic system modelling chemotaxis. *J. Math. Anal. Appl.* **326**(2), 1410–1426 (2007). <https://doi.org/10.1016/j.jmaa.2006.03.080>
3. Wang, J., Warnecke, G.: Existence and uniqueness of solutions for a non-uniformly parabolic equation. *J. Differ. Equ.* **189**(1), 1–16 (2003). [https://doi.org/10.1016/S0022-0396\(02\)00059-1](https://doi.org/10.1016/S0022-0396(02)00059-1)
4. Chemin, J.-Y., Desjardins, B., Gallagher, I., Grenier, E.: Fluids with anisotropic viscosity. *M2AN Math. Model. Numer. Anal.* **34**(2), 315–335 (2000). <https://doi.org/10.1051/m2an:2000143>
5. Tsyganov, E., Hoff, D.: Systems of partial differential equations of mixed hyperbolic-parabolic type. *J. Differ. Equ.* **204**(1), 163–201 (2004). <https://doi.org/10.1016/j.jde.2004.03.010>
6. Kohn, J.J., Nirenberg, L.: Non-coercive boundary value problems. *Commun. Pure Appl. Math.* **18**, 443–492 (1965). <https://doi.org/10.1002/cpa.3160180305>
7. Kohn, J.J., Nirenberg, L.: Degenerate elliptic-parabolic equations of second order. *Commun. Pure Appl. Math.* **20**, 797–872 (1967). <https://doi.org/10.1002/cpa.3160200410>
8. Oleřnik, O.A., Radkevič, E.V.: Second Order Equations with Nonnegative Characteristic Form. Translated from the Russian by Paul C. Fife, p. 259. Plenum Press, New York-London (1973)
9. Lieberman, G.M.: Second Order Parabolic Differential Equations. World Scientific Publishing Co., Inc., River Edge, p. 439 (1996). <https://doi.org/10.1142/3302>
10. DiBenedetto, E.: Degenerate Parabolic Equations. Universitext. Springer, New York, p. 387 (1993). <https://doi.org/10.1007/978-1-4612-0895-2>
11. Klainerman, S., Majda, A.: Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids. *Commun. Pure Appl. Math.* **34**(4), 481–524 (1981). <https://doi.org/10.1002/cpa.3160340405>
12. Schochet, S., Xu, X.: Towards Uniform Existence and Convergence Theorems for Three-scale Systems of Hyperbolic PDEs with General Initial Data. preprint (2021)
13. Taylor, M.E.: Pseudodifferential Operators and Nonlinear PDE. Progress in Mathematics, vol. 100. Birkhäuser Boston, Inc., Boston, p. 213 (1991). <https://doi.org/10.1007/978-1-4612-0431-2>
14. Schochet, S.: Fast singular limits of hyperbolic PDEs. *J. Differ. Equ.* **114**(2), 476–512 (1994). <https://doi.org/10.1006/jdeq.1994.1157>
15. Trias, F.X., Gorobets, A., Oliva, A.: A simple approach to discretize the viscous term with spatially varying (eddy-)viscosity. *J. Comput. Phys.* **253**, 405–417 (2013). <https://doi.org/10.1016/j.jcp.2013.07.021>
16. Meng, Y., Wang, Y.-G.: Highly oscillatory waves in quasilinear hyperbolic-parabolic coupled equations. *J. Math. Phys.* **58**(8), 081504–29 (2017). <https://doi.org/10.1063/1.4996887>
17. Cheng, B., Ju, Q., Schochet, S.: Three-scale singular limits of evolutionary PDEs. *Arch. Ration. Mech. Anal.* **229**(2), 601–625 (2018). <https://doi.org/10.1007/s00205-018-1233-5>
18. Hubbard, J.H., West, B.H.: Differential Equations: A Dynamical Systems Approach. Texts in Applied Mathematics, vol. 5. Springer, New York, p. 350 (1995). <https://doi.org/10.1007/978-1-4612-4192-8>
19. Majda, A.: Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables. Applied Mathematical Sciences, vol. 53. Springer, New York, p. 159 (1984). <https://doi.org/10.1007/978-1-4612-1116-7>
20. Schochet, S.: Symmetric hyperbolic systems with a large parameter. *Commun. Partial Differ. Equ.* **11**(15), 1627–1651 (1986). <https://doi.org/10.1080/03605308608820478>
21. Friedrichs, K.O.: Symmetric positive linear differential equations. *Commun. Pure Appl. Math.* **11**, 333–418 (1958). <https://doi.org/10.1002/cpa.3160110306>