Mathematische Annalen

Reaction-diffusion systems with supercritical nonlinearities revisited

Anna Kostianko[1,2] · Chunyou Sun[1] · Sergey Zelik[1,2,3]

© The Author(s) 2021

Abstract

We give a comprehensive study of the analytic properties and long-time behavior of solutions of a reaction-diffusion system in a bounded domain in the case where the nonlinearity satisfies the standard monotonicity assumption. We pay the main attention to the supercritical case, where the nonlinearity is not subordinated to the linear part of the equation trying to put as small as possible amount of extra restrictions on this nonlinearity. The properties of such systems in the supercritical case may be very different in comparison with the standard case of subordinated nonlinearities. We examine the global existence and uniqueness of weak and strong solutions, various types of smoothing properties, asymptotic compactness and the existence of global and exponential attractors.

Mathematics Subject Classification 35B40 · 35B45 · 35K10

Communicated by Y. Giga.

✉ Sergey Zelik
  s.zelik@surrey.ac.uk

  Anna Kostianko
  anna.kostianko@surrey.ac.uk

  Chunyou Sun
  sunchy@lzu.edu.cn

[1] School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, People's Republic of China

[2] Department of Mathematics, University of Surrey, Guildford GU2 7XH, United Kingdom

[3] Keldysh Institute of Applied Mathematics, Moscow, Russia

Contents

1 Introduction

We study the following reaction-diffusion system in a bounded domain $\Omega \subset \mathbb{R}^d$ with smooth boundary:

$$\partial_t \mathbf{u} = \mathbf{D}\Delta_x \mathbf{u} - \mathbf{f}(\mathbf{u}) + \mathbf{g}, \ \ \mathbf{u}\big|_{t=0} = \mathbf{u_0}, \ \mathbf{u}\big|_{\partial\Omega} = 0 \tag{1.1}$$

(endowed with the Dirichlet boundary conditions). Here $\mathbf{u} = (u_1, \cdots, u_k)$ is an unknown vector-valued function, $\mathbf{D} := (d_{ij})_{i,j=1}^k$ is a given constant diffusion matrix and $\mathbf{f}(\mathbf{u}) := (f_1(\mathbf{u}), \cdots, f_k(\mathbf{u}))$ and $\mathbf{g} := (g_1, \cdots, g_k)$ are given nonlinearity and external forces respectively.

Equations of the form (1.1) model various classical phenomena in modern science (e.g., heat conduction, chemical kinetics, various quantum effects (Ginzburg-Landau equations), mathematical biology (Fitz-Hugh-Nagumo or Keller-Segel equations (if we allow the nonlinearity to depend also on $\nabla_x \mathbf{u}$), etc.) and have been intensively studied from both mathematical and applied points of view, see [3,8,10,15,24,30,34,39, 41,43] and references therein. In a sense, this is the most studied and somehow simplest model example of an evolutionary PDE which may generate non-trivial dynamics.

Since the analytic properties of the *linear* system (1.1) are completely understood, the analogous properties for the nonlinear equation depend strongly on whether or not we are able to treat the term $\mathbf{f}(\mathbf{u})$ as a perturbation. As usual, if we want to have global existence of a solution, we need to find the proper *a priori estimates*, usually with the help of energy functionals or some "wisely" chosen Lyapunov type functionals. This, in turn, requires some restrictions on the function $\mathbf{f}$ and matrix $\mathbf{D}$ (to prevent the finite-time blow up of solutions). Then, if the found a priori estimates are strong enough to treat the nonlinearity as a perturbation (the so-called *subcritical* case), the analytic properties of the nonlinear equation is usually the same as for the dominating linear one and more or less complete theory is available. In contrast to this, in the *supercritical* case, the nonlinearity is strong enough to destroy the nice properties of the underlying linear equation, for instance, to produce the finite-time blow up of initially smooth solutions (despite the fact that the "energy" remains bounded and dissipative, see

[6] for such a phenomena in complex Ginzburg-Landau equation, [36] for chemical kinetics equations or [25] for chemotaxis models). Usually, the sub/super criticality of the considered equation is determined by the *growth rate* of the nonlinearity **f**(**u**) which depends on a priori estimates available (through the choice of the phase space for the problem) and the space dimension (through Sobolev embedding theorems). Thus, the typical picture for equation (1.1) is the following: we have the so-called critical growth exponent $p = p_{crit} > 1$ and an extra condition

$$|\mathbf{f}(\mathbf{u})| \leq C(1 + |\mathbf{u}|^p), \ \mathbf{u} \in \mathbb{R}^k \tag{1.2}$$

on the nonlinearity and the equation is subcritical if $p < p_{crit}$, critical if $p = p_{crit}$ and supercritical if $p > p_{crit}$, see [3,8,39,41] for more details.

Note that in the scalar case $k = 1$, equation (1.1) possesses the maximum/comparison principle as well as a global Lyapunov functional (which can be obtained by formal multiplication of the equation by $\partial_t u$ and integration in x) and these two properties simplify greatly the study of the equation. Indeed, one can take $L^\infty(\Omega)$ as a natural phase space and compare the solution u of (1.1) with the corresponding solutions of ODEs

$$y'_\pm(t) = -f(y_\pm(t)) \pm \|g\|_{L^\infty}, \ \ y(0) = \pm\|u(0)\|_{L^\infty}.$$

This gives the global existence of the solution $u(t)$ under the following Osgood type conditions which are somehow close to the optimal ones, namely, the global solvability holds if there exists a positive smooth function ψ such that

$$f(u) \operatorname{sgn}(u) \geq -\psi(|u|), \ \ \forall u \in \mathbb{R} \ \text{ and } \int_0^\infty \frac{dz}{\psi(z)} = \infty,$$

see [26,38] and references therein. Moreover, the dissipativity (in, say, $L^2(\Omega)$ as well as in $L^\infty(\Omega)$) will hold if we take $\psi \equiv C$, i.e.,

$$f(u) \operatorname{sgn}(u) \geq -C, \ \ \forall u \in \mathbb{R}, \tag{1.3}$$

see, e.g., [3,39,41]. The natural class of nonlinearities are the polynomials $f(u) = a_{2n+1}u^{2n+1} + \cdots + a_0, n \in \mathbb{N}$, with the "right" sign of the leading order coefficient: $a_{2n+1} > 0$. In this case, the restriction (1.2) on the growth rate of f is not necessary and we have well-posedness and dissipativity of the considered scalar equation no matter how fast the growth of the non-linearity is.

In contrast to this, in the case of "wrong sign" $a_{2n+1} < 0$, we always have blow up in finite time at least for some initial data and the properties of solutions $u(t)$ strongly depend on the growth exponent p in (1.2). Since this case is out of the scope of our paper, we will not discuss it here and refer the interested reader to [38] (see also reference therein) for more details.

Unfortunately, the universal conditions on **f** and **D** which would allow to avoid the finite-time blow up and give the dissipativity in nice phase spaces are known in the scalar case $k = 1$ only, so many different classes of sufficient conditions are suggested

for the case of systems strongly depending on the area of science where the considered system comes from. For instance, from the point of view of chemical kinetics, it is natural to assume that $\mathbf{D}$ is diagonal with non-negative entries and $\mathbf{f}(\mathbf{u})$ satisfies the balance law

$$\sum_{i=1}^{k} f_i(\mathbf{u}) \leq 0 \tag{1.4}$$

which mimics the mass conservation law for the concentrations u_i of reagents (which usually belong to the non-negative cone in $\mathbb{R}^k$). The natural energy here is the L^1-norm of the solution $\mathbf{u}(t)$ (the total mass is conserved or at least non-increasing), see [12,22,34,37,38] and the references therein for more details. We note that in the supercritical case the solutions may blow up in finite time despite the conservation of total mass, see [36].

Alternatively, the so-called invariant region technique is often used in order to get the global existence of solutions of (1.1). The idea of this technique is to find a bounded invariant region $\mathbf{R}$ in the phase space of the problem and consider only the solutions $\mathbf{u}(t)$ satisfying $\mathbf{u}(0) \in \mathbf{R}$. Since the trajectory cannot leave the invariant region, this gives the global existence of solutions for such initial data and their boundedness. Usually this method is applied for the case where $\mathbf{R}$ is a segment

$$\mathbf{R} := \{\mathbf{u} \in L^{\infty}(\Omega), \ a_i \leq u_i \leq b_i, \ i = 1, \cdots, k\}, \ a_i, b_i \in \mathbb{R}$$

and $\mathbf{D}$ is a diagonal matrix. Necessary and sufficient conditions for such a region to be invariant can be found, e.g., in [40].

Clearly, assumptions (1.4) are not appropriate for many other types of equations of the form (1.1), for instance, for complex Ginzburg Landau or Fitz-Hugh-Nagumo equations (and the invariant region technique is hardly applicable for the non-diagonal matrices $\mathbf{D}$), so other types of assumptions should be used instead. The most widespread (especially in the literature related with the attractor theory, see [3,8,39,41]) is the following *dissipativity* condition:

$$\mathbf{f}(\mathbf{u}).\mathbf{u} \geq -C, \ \mathbf{u} \in \mathbb{R}^k \tag{1.5}$$

which is a straightforward extension of (1.3) to the case of systems and which is usually accompanied by the assumption that $\mathbf{D}$ has a positive symmetric part. These assumptions are related with the so-called L^2-energy identity

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{u}(t)\|_{L^2}^2 + (\mathbf{D}\nabla_x\mathbf{u}, \nabla_x\mathbf{u}(t)) + (\mathbf{f}(\mathbf{u}(t)), \mathbf{u}(t)) = (\mathbf{g}, \mathbf{u}(t)) \tag{1.6}$$

which can be formally obtained by multiplying equation (1.1) by $\mathbf{u}$ and integrating over x and which gives (due to these assumptions) the dissipative control of the L^2-norm of $\mathbf{u}(t)$, see Lemma 3.1. Here and below $(\mathbf{u}, \mathbf{v}) := \int_{\Omega} \mathbf{u}(x).\mathbf{v}(x)\,dx$ and

$$(\mathbf{D}\nabla_x\mathbf{u}, \nabla_x\mathbf{u}) := \sum_{i=1}^{d}(\mathbf{D}\partial_{x_i}\mathbf{u}, \partial_{x_i}\mathbf{u}).$$

As one can see from (1.6), assumption (1.5) may be slightly weakened by adding the term $-k|\mathbf{u}|^2$ with a sufficiently small k (depending on the first eigenvalue of the Laplacian in Ω and matrix $\mathbf{D}$) to the right-hand side. This term may be essential, e.g., when the non-linearities with sub-linear growth rate are considered, since in this paper we mainly concentrate on fast growing non-linearities, we restrict ourselves to the case $k = 0$ only.

More important is that the critical exponent which corresponds to this energy control (and the choice $H = L^2(\Omega)$ as a phase space):

$$\bar{p}_{crit} := 1 + \frac{4}{d}$$

is rather restrictive (the most natural cubic nonlinearity is supercritical in 3D case) and not much can be said in general about the supercritical case where the uniqueness of solutions may be lost and finite-time blow up of the L^∞-norm may occur (see [6] for the numerical blow up evidence in 3D complex Ginzburg-Landau equation, see also [8,35] and references therein for study the long-time behavior of solutions without uniqueness using the multi-valued or trajectory approaches). We also mention here the so-called *anisotropic dissipativity* assumption:

$$\sum_{i=1}^{k} f_i(\mathbf{u})u_i|u_i|^{l_i} \geq -C,$$

where $l = (l_1, \cdots, l_k)$ is a sufficiently large vector, introduced in [19]. This restriction accompanied by the assumption that $\mathbf{D}$ is diagonal gives $p_{crit} = \infty$ if $l = l(d)$ is large enough.

A natural alternative is to use the so-called *monotonicity* assumption:

$$\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\xi.\xi \geq -K|\xi|^2, \; \forall \xi, \mathbf{u} \in \mathbb{R}^k \tag{1.7}$$

for some $K \in \mathbb{R}$. Here and below

$$\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}) = \mathbf{f}'(\mathbf{u}) := (\partial_{u_j} f_i(\mathbf{u}))_{i,j=1}^k$$

is the Frechet derivative (=Jacobi matrix) of the function $\mathbf{f}(\mathbf{u})$. This assumption is also very widespread in the literature related with attractors and is naturally related with the H^1-energy identity:

$$\frac{1}{2}\frac{d}{dt}\|\nabla_x\mathbf{u}(t)\|_{L^2}^2 + (\mathbf{D}\Delta_x\mathbf{u}(t), \Delta_x\mathbf{u}(t)) + \\ + (\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}(t))\nabla_x\mathbf{u}(t), \nabla_x\mathbf{u}(t)) = -(\mathbf{g}, \Delta_x\mathbf{u}(t)) \tag{1.8}$$

which is obtained by formal multiplication of (1.1) by $-\Delta_x\mathbf{u}$ and integration over x. Together with (1.7) this gives the dissipative control of the H^1-norm of the solution,

see Lemma 3.2 for the details. The critical growth exponent associated with this H^1-energy control is

$$\widetilde{p}_{crit} = 1 + \frac{4}{d-2}, \ d > 2 \tag{1.9}$$

can be found in many works, see [3,8] and references therein. However, as pointed out in [46], the monotonicity assumption (1.7) gives for free the control of H^2-norm of the solution $\mathbf{u}(t)$ together with the L^2-norm of $\mathbf{f}(\mathbf{u}(t))$, namely, we have a priori estimates for the solutions in the nonlinear space

$$\mathbb{D} := \{\mathbf{u} \in H^2(\Omega) \cap H_0^1(\Omega), \ \mathbf{f}(\mathbf{u}) \in L^2(\Omega)\},$$

due to the control of the L^2-norm of $\partial_t \mathbf{u}(t)$, see Lemma 3.3 and Corollary 3.4 below, and this gives us much better value of the critical exponent:

$$p_{crit} = 1 + \frac{4}{d-4}, \ d > 4. \tag{1.10}$$

As far as we know, up to the moment, this is the best growth restriction which guarantees (of course, under the monotonicity assumption (1.7)) the global existence of smooth solutions and which is widely used nowadays not only for reaction-diffusion equations, but for many other related problems (like Cahn-Hilliard equations, see [35] and references therein; strongly damped wave equations, see [11,27] and reference therein, etc.). By these reasons, we will use formula (1.10) throughout the paper as a definition of the critical exponent for the considered case.

We also note that the monotonicity assumption (1.7) gives the uniqueness of weak solutions (= solutions in the energy phase space $H = L^2(\Omega)$, see Sect. 5 below) even in the supercritical case which, in turn, allows to get a lot of information about the solutions and their long-time behavior in the supercritical case as well. The theory of equations (1.1) in the critical or supercritical cases is of a great current interest, see for instance [8–10,14,35,47,48] and references therein. However, in most cases rather essential extra restrictions on the nonlinearity $\mathbf{f}$ are posed like the following two sided estimate:

$$C(|\mathbf{u}|^{p-1} - 1) \geq \nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}) \geq -K + \alpha |\mathbf{u}|^{p-1}, \ \mathbf{u} \in \mathbb{R}^k, \ C, \alpha > 0 \tag{1.11}$$

which really simplifies the situation, but automatically excludes some interesting new phenomena which may appear in a general case.

The aim of the present paper is to give a comprehensive study of weak and strong solutions (=solutions in the phase space $\mathbb{D}$) of problem (1.1) as well as their long-time behavior in the supercritical case $p > p_{crit}$ with dissipative (assumption (1.5) is fulfilled) and monotone (assumption (1.7) is satisfied) nonlinearities trying to avoid/minimize further restrictions on $\mathbf{f}$.

Strong and weak solutions of problem (1.1) have been constructed in [46] (see also Theorem 4.3 of Section 4 and Theorem 5.2 of Sect. 5). The construction of strong solutions is more or less standard. However, in contrast to the usual case where two sided conditions like (1.11) are posed, the phase space $\mathbb{D}$ is *non-linear* and it not clear even

whether or not smooth functions are dense in it, so some accuracy is required. In [46] this difficulty has been overcome using Galerkin approximations and the monotone operators theory, in the present paper, we prefer to give more explicit approximation scheme which does not use at least in a straightforward way the monotone operators technique and which is more convenient for what follows, see Sect. 5 for the details.

The absence of any two-sided control for the growth of the nonlinearity **f** also leads to extra difficulties on the level of weak solutions. Indeed, since we cannot guarantee that $\mathbf{f}(\mathbf{u}) \in L^1$, we cannot treat the equation in the sense of distributions and have to use variational inequalities, see Definition 5.1. In addition, even the parabolic smoothing property (whether or not a weak solution becomes strong at the next time moment) becomes non-trivial and has been posed in [46] as an open problem.

Our first main result gives the positive answer on this question in the case where the nonlinearity has a polynomial growth rate.

Theorem 1.1 *Let the nonlinearity* **f** *satisfy the assumptions* (1.2) *(for some* $p > 0$*),* (1.5) *and* (1.7)*, the diffusion matrix have positive symmetric part and* $\mathbf{g} \in L^2(\Omega)$*. Then any weak solution* $\mathbf{u}(t)$ *starting from* $\mathbf{u}(0) \in H = L^2(\Omega)$ *belongs to* $\mathbb{D}$ *for any* $t > 0$*. In addition, the strong solutions of* (1.1) *are dissipative in* $\mathbb{D}$*-norm as well.*

The proof of this theorem is based on estimation of $\mathbf{f}(\mathbf{u})$ in Lebesgue spaces $L^q(\Omega)$ with $0 < q < 1$ and is given in Sect. 6.

Our next result shows that the critical growth exponent can be slightly improved.

Theorem 1.2 *Let the assumptions of Theorem* 1.1 *hold and let, in addition, the growth exponent* p *of the nonlinearity satisfy*

$$p < p_{crit} + \varepsilon = 1 + \frac{4}{d-4} + \varepsilon, \ \ d > 4$$

for some small positive $\varepsilon = \varepsilon(\mathbf{D})$*. Then any weak solution* $\mathbf{u}(t)$ *of problem* (1.1) *starting from* $\mathbf{u}(0) \in L^2(\Omega)$ *belongs to* $L^\infty(\Omega)$ *for any* $t > 0$*. In particular, finite-time blow up of smooth solutions is impossible and the actual regularity of a solution* $\mathbf{u}(t)$ *is restricted by the regularity of* Ω*,* **f** *and* **g** *only. In the case where this data is* C^∞*-smooth, the corresponding solution* $\mathbf{u}(t)$ *will be also* C^∞ *for any* $t > 0$*.*

We now turn to the attractors. The existence of a global attractor for problem (1.1) in H has been verified in [46], however, the question about strong attraction in $\mathbb{D}$ has been remained open. Our next result gives a positive answer on this question under the extra restriction

$$|\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})| \leq C(|\mathbf{f}(\mathbf{u})| + 1 + |\mathbf{u}|), \ \ \mathbf{u} \in \mathbb{R}^k. \tag{1.12}$$

Theorem 1.3 *Let the assumptions of Theorem* 1.1 *hold and let, in addition,* **f** *satisfy* (1.12)*. Then the solution semigroup* $S(t)$ *associated with problem* (1.1) *possesses a compact global attractor* $\mathcal{A}$ *in the phase space* $\mathbb{D}$*.*

The proof of this theorem is given in Sect. 7 and is based on the energy type arguments. We expect that assumption (1.12) is technical and can be removed, but it is strongly

related with the validity of the integration by parts formula

$$(\mathbf{f}(\mathbf{u}), \Delta_x \mathbf{u}) = -(\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}) \nabla_x \mathbf{u}, \nabla_x \mathbf{u}), \quad \mathbf{u} \in \mathbb{D},$$

see the discussion in Sect. 9 below.

Finally, we study the finite-dimensionality of the constructed global attractor $\mathcal{A}$ in $\mathbb{D}$ and the existence of the so-called exponential attractor (see [16–18,35] and also Sect. 8 for more details).

Theorem 1.4 *Let the assumptions of Theorem 1.3 hold and let, in addition, some extra convexity assumptions on the function* $\mathbf{u} \to |\mathbf{f}(\mathbf{u})|$ *be posed (see formula* (8.3)*). Then the solution semigroup* $S(t)$ *associated with equation* (1.1) *possesses an exponential attractor* $\mathcal{M}$ *in H and, in particular, the fractal dimension of the global attractor* $\mathcal{A}$ *in H is finite.*

The finite-dimensionality of the global attractor $\mathcal{A}$ has been established in [46] under the similar assumptions using the so-called method of l-trajectories developed in [32, 33]. In Sect. 8 we suggest an alternative more transparent method for constructing of an exponential attractor which does not utilize l-trajectories and works directly in the phase space.

Note that the results of the paper are heavily based on the monotonicity arguments, so we cannot consider more or less general non-linearities $\mathbf{f}(\mathbf{u}, \nabla_x \mathbf{u})$ which contain on the spatial gradient of $\mathbf{u}$. However, there is an important particular case where our theory works with very minor changes, namely, where

$$\mathbf{f}(\mathbf{u}, \nabla_x \mathbf{u}) := \mathbf{f}(\mathbf{u}) + \mathbf{L} \nabla_x \mathbf{u}$$

and $\mathbf{L}$ is a linear map (a mild dependence of $\mathbf{L}$ on $\mathbf{u}$ is also allowed). Some results in this direction are obtained in [47]. We also note that many of our results can be extended to the case of quasilinear systems where the Laplacian $\Delta_x \mathbf{u}$ is replaced, for instance, by the so-called p-Laplacian for which the theory of monotone operators is applicable, see [3,5,43,45] and references therein. We return to this type of problems somewhere else.

The paper is organized as follows.

Notation and spaces which will be used throughout the paper are introduced in Sect. 2 and the standard a priori estimates for the solutions $\mathbf{u}(t)$ of problem (1.1) are recalled in Sect. 3.

The existence of strong solutions for problem (1.1) is verified in Sect. 4 based on special approximations of the nonlinearity $\mathbf{f}$. The definition of a weak solution of problem (1.1) in the sense of variational inequalities as well as the proof of its global existence and uniqueness is given in Sect. 5. Moreover, the existence of a global attractor $\mathcal{A}$ for the solution semigroup $S(t)$ is also verified there.

The weak to strong smoothing property, see Theorem 1.1, is verified in Sect. 6. The further regularity of strong solutions is obtained in Sect. 7. In particular, the proofs of Theorems 1.2 and 1.3 are given there. Some results about the partial regularity of the elliptic problem associated with equations (1.1) which have an independent interest are

obtained in Appendix A. The existence of an exponential attractor $\mathcal{M}$, see Theorem 1.4, is given in Sect. 8.

Finally, Sect. 9 discusses natural extensions of the developed theory to other classes of dissipative PDEs, in particular, to fractional reaction-diffusion systems and (fractional) Cahn-Hilliard type equations. At the end of this section we also discuss some important (at least from our point of view) open problems for further investigation.

2 Assumptions and preliminaries

Throughout the paper we consider the following reaction-diffusion system in a bounded domain $\Omega \subset \mathbb{R}^d$:

$$\partial_t \mathbf{u} = \mathbf{D}\Delta_x \mathbf{u} - \mathbf{f}(\mathbf{u}) + \mathbf{g}, \quad \mathbf{u}\big|_{\partial\Omega} = 0, \quad \mathbf{u}\big|_{t=0} = \mathbf{u}_0. \tag{2.1}$$

Here $\mathbf{u} = (u_1, \cdots, u_k)$ is an unknown vector-valued function, $\mathbf{D}$ is a constant diffusion matrix satisfying

$$\mathbf{D} + \mathbf{D}^* > 0, \tag{2.2}$$

$\mathbf{g} \in L^2(\Omega)$ is a given external force and the nonlinearity $\mathbf{f}$ is assumed to satisfy the following conditions:

$$\begin{cases} 1.\ \mathbf{f} \in C^1(\mathbb{R}^k, \mathbb{R}^k), \\ 2.\ \mathbf{f}(\mathbf{u}).\mathbf{u} \geq -C, \ \mathbf{u} \in \mathbb{R}^k, \\ 3.\ \nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}) \geq -K, \ \mathbf{u} \in \mathbb{R}^k, \end{cases} \tag{2.3}$$

where C and K are some fixed constants, $\mathbf{u}.\mathbf{v}$ stands for the standard inner product in $\mathbb{R}^k$ and $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}) \geq -K$ means $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\xi.\xi \geq -K|\xi|^2$ for all $\xi \in \mathbb{R}^k$.

For any $l \in \mathbb{N}$ and any $1 \leq p \leq \infty$ we denote by $W^{l,p}(\Omega)$ the Sobolev space of distributions $u \in \mathcal{D}'(\Omega)$ such that u and all its partial derivatives up to order l inclusively belong to the Lebesgue space $L^p(\Omega)$. As usual, for non-integer values of l, we define $W^{l,p}(\Omega) := B^l_{p,p}(\Omega)$ using real interpolation ($B^l_{p,p}$ is a classical Besov space, see e.g., [42]). Moreover, the symbol $W^{l,p}_0(\Omega)$ stands for the closure of $C^\infty_0(\Omega)$ in $W^{l,p}(\Omega)$ and the space $W^{-l,p}(\Omega)$ is defined as a dual space to $W^{l,q}_0(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$, with respect to the standard inner product in $H = L^2(\Omega)$. To simplify the notation, we will write $H^l(\Omega)$ instead of $W^{l,2}(\Omega)$. We will also systematically use the Sobolev spaces $W^{l,p}(\Omega, \mathbb{R}^k)$ of vector-valued functions $\mathbf{u} : \Omega \to \mathbb{R}^k$ and will write for brevity $W^{l,p}(\Omega)$ instead of $W^{l,p}(\Omega, \mathbb{R}^k)$ if this does not lead to misunderstandings.

In a sequel, we will also use the space $L^p(\Omega)$ with $0 < p < 1$ and use the standard notation

$$\|\mathbf{u}\|_{L^p(\Omega)} := \left(\int_\Omega |\mathbf{u}(\mathbf{x})|^p \, dx \right)^{1/p}$$

simply ignoring the fact that it is not a norm. Recall that the topology in this space is defined by the metric $d_p(\mathbf{u}, \mathbf{v}) := \|\mathbf{u} - \mathbf{v}\|^p_{L^p}$.

We say that the function $\mathbf{u}(t, x)$ is a *strong* solution of (2.1) if

$$\mathbf{u} \in C_{\mathrm{w}}(0, T; H^2(\Omega) \cap H_0^1(\Omega)), \quad \mathbf{f}(\mathbf{u}) \in C_{\mathrm{w}}(0, T; L^2(\Omega)) \tag{2.4}$$

and equation (2.1) is satisfied in the sense of distributions (here and below $C(a, b; X)$ means the space of X-valued functions defined and continuous on a *closed* interval $[a, b]$ and the lower index "w" stands for the continuity in weak topology). In particular, for strong solutions we require that the initial data $\mathbf{u}_0 \in \mathbb{D}$, where

$$\mathbb{D} := \{\mathbf{u} \in H^2(\Omega) \cap H_0^1(\Omega), \ \mathbf{f}(\mathbf{u}) \in L^2(\Omega)\},$$
$$\|\mathbf{u}\|_{\mathbb{D}}^2 := \|\mathbf{u}\|_{H^2}^2 + \|\mathbf{f}(\mathbf{u})\|_{L^2}^2. \tag{2.5}$$

Note that in general $\mathbb{D}$ is *not a linear* space and this causes a lot of extra difficulties in comparison with the case of linear phase space. We define the topology in the space $\mathbb{D}$ using the embedding

$$j : \mathbb{D} \to [H^2(\Omega) \cap H_0^1(\Omega)] \times L^2(\Omega), \quad j(\mathbf{u}) = \{\mathbf{u}, \mathbf{f}(\mathbf{u})\}.$$

In particular, the sequence $\mathbf{u}_n \to \mathbf{u}$ strongly in $\mathbb{D}$ if $\mathbf{u}_n \to \mathbf{u}$ in $H^2(\Omega)$ and $\mathbf{f}(\mathbf{u}_n) \to \mathbf{f}(\mathbf{u})$ in $L^2(\Omega)$. Analogously, we say that $\mathbf{u}_n \to \mathbf{u}$ *weakly* in $\mathbb{D}$ if $\mathbf{u}_n \to \mathbf{u}$ weakly in $H^2(\Omega)$ and $\mathbf{f}(\mathbf{u}_n) \to \mathbf{f}(\mathbf{u})$ weakly in $L^2(\Omega)$. We will also write for brevity $C_{\mathrm{w}}(0, T; \mathbb{D})$ instead of (2.4).

3 A priori estimates

In this section, we give a number of more or less standard estimates for strong solutions of problem (2.1) which will be justified later. We start with the dissipative estimate in the space $H = L^2(\Omega)$.

Lemma 3.1 *Let* $\mathbf{g} \in H$, *assumptions* (2.2)–(2.3) *hold and let* $\mathbf{u}$ *be a sufficiently smooth solution of* (2.1). *Then, the following estimate is valid:*

$$\|\mathbf{u}(T)\|_{L^2}^2 + \int_T^{T+1} \|\mathbf{u}(t)\|_{H^1}^2 \, dt + \int_T^{T+1} |\mathbf{f}(\mathbf{u}).\mathbf{u}| \, dt \leq$$
$$\leq C e^{-\alpha T} \|\mathbf{u}_0\|_{L^2}^2 + C(\|\mathbf{g}\|_{L^2}^2 + 1), \tag{3.1}$$

where the positive constants C and α are independent of t and $\mathbf{u}_0$.

Proof We multiply equation (2.1) by $\mathbf{u}$ and integrate over x. This gives

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{u}\|_{L^2}^2 + (\mathbf{D}\nabla_x \mathbf{u}, \nabla_x \mathbf{u}) + (\mathbf{f}(\mathbf{u}), \mathbf{u}) = (\mathbf{g}, \mathbf{u}),$$

where $(\mathbf{u}, \mathbf{v}) := \int_\Omega \mathbf{u}(x).\mathbf{v}(x)\, dx$ is a standard inner product in H and

$$(\mathbf{D}\nabla_x\mathbf{u}, \nabla_x\mathbf{u}) := \sum_{i=1}^{d}(\mathbf{D}\partial_{x_i}\mathbf{u}, \partial_{x_i}\mathbf{u}) = \int_{x\in\Omega} \mathrm{Tr}(\mathbf{D}\nabla_x\mathbf{u}(\nabla_x\mathbf{u})^*)\, dx.$$

Using the dissipativity assumption $\mathbf{f}(\mathbf{u}).\mathbf{u} \geq -C$ and positivity of the matrix $\mathbf{D}$ together with the Friedrichs (Poincare) inequality, we arrive at

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{u}\|_{L^2}^2 + \alpha\|\mathbf{u}\|_{L^2}^2 + \alpha\|\nabla_x\mathbf{u}\|_{L^2}^2 + (|\mathbf{f}(\mathbf{u}).\mathbf{u}|, 1) \leq C(\|\mathbf{g}\|_{L^2}^2 + 1)$$

for some positive constants C and α. The Gronwall inequality applied to this relation gives (3.1) and finishes the proof of the lemma.   □

The next lemma gives the analogous dissipative estimate for the H^1-norm of the solution.

Lemma 3.2 *Let the assumptions of Lemma 3.1 hold and $\mathbf{u}$ be a sufficiently regular solution of (2.1). Then, the following estimate is valid:*

$$\|\mathbf{u}(T)\|_{H^1}^2 + \int_T^{T+1} \|\mathbf{u}(t)\|_{H^2}^2\, dt +$$
$$+ \int_T^{T+1} (|\nabla_\mathbf{u}\mathbf{f}(\mathbf{u})\nabla_x\mathbf{u}(t).\nabla_x\mathbf{u}(t)|, 1)\, dt \leq Ce^{-\alpha T}\|\mathbf{u}_0\|_{H^1}^2 + C(\|\mathbf{g}\|_{L^2}^2 + 1),$$
$$\tag{3.2}$$

where the positive constants C and α are independent of t and $\mathbf{u}_0$.

Proof We multiply equation (2.1) by $-\Delta_x\mathbf{u}$ and integrate over x to get

$$\frac{1}{2}\frac{d}{dt}\|\nabla_x\mathbf{u}\|_{L^2}^2 + (\mathbf{D}\Delta_x\mathbf{u}, \Delta_x\mathbf{u}) + (\nabla_\mathbf{u}\mathbf{f}(\mathbf{u})\nabla_x\mathbf{u}, \nabla_x\mathbf{u}) = -(\mathbf{g}, \Delta_x\mathbf{u}).$$

Using the inequality $\nabla_\mathbf{u}\mathbf{f}(\mathbf{u}) \geq -K$ and positivity of matrix $\mathbf{D}$ again, we arrive at

$$\frac{d}{dt}\|\nabla_x\mathbf{u}\|_{L^2}^2 + \alpha\|\nabla_x\mathbf{u}\|_{L^2}^2 + \alpha\|\Delta_x\mathbf{u}\|_{L^2}^2 +$$
$$+ (|\nabla_\mathbf{u}\mathbf{f}(\mathbf{u})\nabla_x\mathbf{u}.\nabla_x\mathbf{u}|, 1) \leq C(\|\mathbf{g}\|_{L^2}^2 + \|\nabla_x\mathbf{u}\|_{L^2}^2). \tag{3.3}$$

Applying the Gronwall inequality to this relation and using (3.1) in order to control the integral of $\|\nabla_x\mathbf{u}\|_{L^2}^2$, we end up with (3.2) and finish the proof of the lemma.   □

Let now $\boldsymbol{\theta} = \partial_t\mathbf{u}$. Then this function solves

$$\partial_t\boldsymbol{\theta} = \mathbf{D}\Delta_x\boldsymbol{\theta} - \nabla_\mathbf{u}\mathbf{f}(\mathbf{u})\boldsymbol{\theta}, \;\; \boldsymbol{\theta}\big|_{t=0} = \mathbf{D}\Delta_x\mathbf{u}_0 - \mathbf{f}(\mathbf{u}_0) + \mathbf{g}, \; \boldsymbol{\theta}\big|_{\partial\Omega} = 0. \tag{3.4}$$

The next lemma gives the L^2-estimate for the time derivative $\boldsymbol{\theta}$.

Lemma 3.3 *Let the assumptions of Lemma 3.1 hold and let* **u** *be a sufficiently regular solution of equation* (2.1). *Then the following estimate is valid:*

$$\|\boldsymbol{\theta}(T)\|_{L^2}^2 + \int_T^{T+1} \|\boldsymbol{\theta}(t)\|_{H^1}^2 \, dt +$$

$$+ \int_T^{T+1} (|\nabla_\mathbf{u}\mathbf{f}(\mathbf{u})\boldsymbol{\theta}(t).\boldsymbol{\theta}(t)|, 1) \, dt \le C\|\mathbf{u}_0\|_{\mathbb{D}}^2 e^{K_1 T} + Ce^{K_1 T}(\|\mathbf{g}\|^2 + 1), \quad (3.5)$$

where positive constants C and K_1 are independent of t and $\mathbf{u}_0$.

Proof We multiply equation (3.4) by $\boldsymbol{\theta}$ and use assumption $\nabla_\mathbf{u}\mathbf{f}(\mathbf{u}) \ge -K$ and positivity of matrix **D** to get

$$\frac{d}{dt}\|\boldsymbol{\theta}\|_{L^2}^2 + \alpha\|\nabla_x\boldsymbol{\theta}\|_{L^2}^2 + (|\nabla_\mathbf{u}\mathbf{f}(\mathbf{u})\boldsymbol{\theta}.\boldsymbol{\theta}|, 1) \le 2K\|\boldsymbol{\theta}\|_{L^2}^2.$$

Applying the Gronwall inequality to this relation, we get the desired estimate and finish the proof of the lemma. □

As a corollary of this lemma, we get the key control for the norm of the solution in the space $\mathbb{D}$.

Corollary 3.4 *Let the assumptions of Lemma 3.1 hold and let* **u** *be a sufficiently regular solution of* (2.1). *Then the following estimate is valid:*

$$\|\mathbf{u}(T)\|_{\mathbb{D}}^2 \le Ce^{K_1 T}\|\mathbf{u}_0\|_{\mathbb{D}}^2 + Ce^{K_1 T}(1 + \|\mathbf{g}\|_{L^2}^2), \quad (3.6)$$

where the positive constants C and K_1 are independent of T and $\mathbf{u}_0$.

Proof We rewrite equation (2.1) as an elliptic problem

$$\mathbf{D}\Delta_x\mathbf{u}(T) - \mathbf{f}(\mathbf{u}(T)) = \tilde{\mathbf{g}}(T) := \boldsymbol{\theta}(T) - \mathbf{g} \quad (3.7)$$

for every fixed T. Multiplying then this equation by $\Delta_x\mathbf{u}(T)$ (without integration in time!) and using the control for $\boldsymbol{\theta}(T)$ obtained above (together with the elliptic regularity estimate for the Laplacian and the assumption $\nabla_\mathbf{u}\mathbf{f}(\mathbf{u}) \ge -K$), we arrive at the estimate

$$\|\mathbf{u}(T)\|_{H^2}^2 \le C(\|\mathbf{g}\|_{L^2}^2 + \|\boldsymbol{\theta}(T)\|_{L^2}^2 + \|\nabla_x\mathbf{u}(T)\|_{L^2}^2).$$

Expressing the L^2-norm of $\mathbf{f}(\mathbf{u})$ from equation (3.7), we get

$$\|\mathbf{u}(T)\|_{\mathbb{D}}^2 \le C(\|\mathbf{g}\|_{L^2}^2 + \|\boldsymbol{\theta}(T)\|_{L^2}^2 + \|\nabla_x\mathbf{u}(T)\|_{L^2}^2).$$

Estimating the right-hand side of this inequality by (3.5) and (3.2) and using that

$$\boldsymbol{\theta}(0) = \partial_t\mathbf{u}(0) = \mathbf{D}\Delta_x\mathbf{u}(0) - \mathbf{f}(\mathbf{u}(0)) + \mathbf{g},$$

we arrive at the desired estimate and finish the proof of the corollary. □

Remark 3.5 Note that, in contrast to estimates for the L^2 and H^1 norms of the solution $\mathbf{u}(t)$, the obtained estimate for the $\mathbb{D}$-norm of $\mathbf{u}(t)$ is *not* dissipative and even grows exponentially in time. We will remove this drawback later (under some extra assumptions on $\mathbf{f}$).

We conclude this section by establishing the global Lipschitz continuity with respect to the initial data which plays a crucial role in constructing weak solutions for (2.1).

Lemma 3.6 *Let the assumptions of Lemma 3.1 hold and let $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ be two sufficiently regular solutions of equation (2.1). Then the following estimate is valid:*

$$\|\mathbf{u}_1(T) - \mathbf{u}_2(T)\|_{L^2}^2 +$$
$$+ \int_0^T \|\mathbf{u}_1(t) - \mathbf{u}_2(t)\|_{H^1}^2 \, dt \leq Ce^{K_1 T} \|\mathbf{u}_1(0) - \mathbf{u}_2(0)\|_{L^2}^2, \qquad (3.8)$$

where the positive constants C and K_1 are independent of T, $\mathbf{u}_1$ and $\mathbf{u}_2$.

Proof Indeed, let $\mathbf{v}(t) = \mathbf{u}_1(t) - \mathbf{u}_2(t)$. Then this function solves

$$\partial_t \mathbf{v} = \mathbf{D}\Delta_x \mathbf{v} - [\mathbf{f}(\mathbf{u}_1) - \mathbf{f}(\mathbf{u}_2)]. \qquad (3.9)$$

Multiplying this equation by $\mathbf{v}$, using that, due to the monotonicity assumption $\nabla_\mathbf{u} \mathbf{f}(\mathbf{u}) \geq -K$,

$$[\mathbf{f}(\mathbf{u}_1) - \mathbf{f}(\mathbf{u}_2)].[\mathbf{u}_1 - \mathbf{u}_2] \geq -K|\mathbf{v}|^2,$$

and arguing as in the proof of Lemma 3.3, we arrive at the desired estimate. $\square$

4 Existence of strong solutions

Although the construction of a solution from a priori estimates obtained above is more or less standard, it is a bit delicate here since the phase space $\mathbb{D}$ is in general *nonlinear* and, particularly, it is not clear whether or not smooth functions are dense in $\mathbb{D}$. By this reason, we sketch the proof here.

We expect that the existence result can be also obtained using the monotone operators theory (e.g., the standard Yosida approximations), but we prefer to give an alternative, a bit more transparent proof. Our idea is to approximate the nonlinearity $\mathbf{f}$ by a sequence $\mathbf{f}_n$ of functions of which are globally Lipschitz continuous without destroying assumptions (2.3). Then, on the one hand, the existence of solutions for such $\mathbf{f}_n$ is well-known and, on the other hand, as not difficult to see, all estimates obtained above will be uniform with respect to n. Thus, it will only remain to pass to the limit $n \to \infty$. We start with the approximation of $\mathbf{f}$.

Lemma 4.1 *Let the function $\mathbf{f}$ satisfy assumptions (2.3). Then, there exists a sequence of functions $\mathbf{f}_n \in C^1(\mathbb{R}^k, \mathbb{R}^k)$ such that*

$$1. \ \mathbf{f}_n(\mathbf{u}).\mathbf{u} \geq -C, \quad 2. \ \nabla_\mathbf{u} \mathbf{f}(\mathbf{u}) \geq -K \qquad (4.1)$$

uniformly with respect to n. Moreover,

$$\mathbf{f}_n \to \mathbf{f} \tag{4.2}$$

in $C_{loc}(\mathbb{R}^k, \mathbb{R}^k)$ and

$$|\nabla_{\mathbf{u}}\mathbf{f}_n(\mathbf{u})| \le C_n, \quad \mathbf{u} \in \mathbb{R}^k, \tag{4.3}$$

where the constant C_n may depend on n.

Sketch of the proof Let us first introduce a smooth scalar convex function $\Psi(z)$ (we may also require that $\Psi'(s) \ge 0$ as $s \ge 0$) in such a way that $\nabla_{\mathbf{u}}\Psi(|\mathbf{u}|^2)$ grows faster than $|\mathbf{f}(\mathbf{u})|$ as $|\mathbf{u}| \to \infty$. Then, for every fixed $\varepsilon > 0$, we consider the function $\bar{\mathbf{f}}_\varepsilon(\mathbf{u}) = \mathbf{f}(\mathbf{u}) + \varepsilon\nabla_{\mathbf{u}}\Psi(|\mathbf{u}|^2)$. Since the second term will dominate the first one if $|\mathbf{u}|^2$ is large enough, introducing the first smooth cut-off function $\theta_R(|\mathbf{u}|^2)$ such that $\theta_R = 1$ is $|\mathbf{u}| \le R$ and zero if $|\mathbf{u}| \ge 2R$, we may find $R = R_\varepsilon \ge \frac{1}{\varepsilon}$ such that the function

$$\bar{\mathbf{f}}_\varepsilon(\mathbf{u}) := \theta_{R_\varepsilon}(|\mathbf{u}|^2)\mathbf{f}(\mathbf{u}) + \varepsilon\nabla_{\mathbf{u}}\Psi(|\mathbf{u}|^2)$$

satisfies assumptions (2.3) uniformly with respect to $\mathbf{u} \in \mathbb{R}^k$. Note that $\bar{\mathbf{f}}_\varepsilon(\mathbf{u}) = \varepsilon\nabla_{\mathbf{u}}\Psi(|\mathbf{u}|^2) = 2\varepsilon\Psi'(|\mathbf{u}|^2)\mathbf{u}$ for $|\mathbf{u}|^2 \ge 2R_\varepsilon$ and we may make it linear for $|\mathbf{u}|^2 \ge 3R_\varepsilon$ by cutting-off Ψ'' on the interval $2R_\varepsilon \le |\mathbf{u}|^2 \le 3R_\varepsilon$. For instance, we may introduce another smooth cut-off function $\tilde{\theta}_{R_\varepsilon}(\tau) = 1$ as $\tau \le 2R_\varepsilon$ and zero if $\tau \ge 3R_\varepsilon$ and define

$$\Psi'_{R_\varepsilon}(s) := \int_0^s \tilde{\theta}_{R_\varepsilon}(\tau)\Psi''(\tau)d\tau + \Psi'(0).$$

This gives

$$\mathbf{f}_\varepsilon(\mathbf{u}) = \theta_{R_\varepsilon}(|\mathbf{u}|^2)\mathbf{f}(\mathbf{u}) + \varepsilon\nabla_{\mathbf{u}}\Psi_{R_\varepsilon}(|\mathbf{u}|^2).$$

Taking finally $\varepsilon = \varepsilon_n = \frac{1}{n}$, we get the desired approximating sequence. □

We now introduce the approximating system for (2.1)

$$\partial_t\mathbf{u} = \mathbf{D}\Delta_x\mathbf{u} - \mathbf{f}_n(\mathbf{u}) + \mathbf{g}, \quad \mathbf{u}\big|_{t=0} = \mathbf{u}_0^n, \quad \mathbf{u}\big|_{\partial\Omega} = 0, \tag{4.4}$$

where the functions $\mathbf{f}_n$ are constructed in Lemma 4.1. However, the choice of the approximating initial data $\mathbf{u}_0^n$ requires some accuracy. Indeed, we cannot just fix $\mathbf{u}_0^n = \mathbf{u}_0$ since $\|\mathbf{f}_n(\mathbf{u}_0)\|_{L^2}$ will not be uniformly bounded and, as a result, we may lose the estimate of the $\mathbb{D}$-norm of the limit solution. Instead, we define $\mathbf{u}_0^n$ as a solution of the following auxiliary elliptic problem:

$$\mathbf{D}\Delta_x\mathbf{v} - \mathbf{f}_n(\mathbf{v}) - K\mathbf{v} = \mathbf{G} := \mathbf{D}\Delta_x\mathbf{u}_0 + \mathbf{f}(\mathbf{u}_0) - K\mathbf{u}_0, \quad \mathbf{v}\big|_{\partial\Omega} = 0. \tag{4.5}$$

The next lemma gives useful properties of the solutions of this auxiliary problem.

Lemma 4.2 *Let the functions $\mathbf{f}_n$ be as above and $\mathbf{u}_0 \in \mathbb{D}$. Then, for every fixed n, problem (4.5) has a unique solution $\mathbf{v} = \mathbf{u}_0^n$. Moreover, $\|\mathbf{u}_0^n\|_{H^2}$ and $\|\mathbf{f}_n(\mathbf{u}_0^n)\|_{L^2}$ are uniformly bounded as $n \to \infty$ and*

$$\mathbf{u}_0^n \rightharpoonup \mathbf{u}_0, \quad \mathbf{f}_n(\mathbf{u}_0^n) \rightharpoonup \mathbf{f}(\mathbf{u}_0) \tag{4.6}$$

in the spaces H^2 and L^2 respectively.

Proof Indeed, the existence and uniqueness of a solution for (4.5) is obvious since $\mathbf{f}_n(\mathbf{u}) + K\mathbf{u}$ are monotone and globally Lipschitz continuous. Let us prove uniform bounds. Indeed, multiplying (4.5) by $\Delta_x \mathbf{v} = \Delta_x \mathbf{u}_0^n$ and using the monotonicity, we get the estimate

$$\|\mathbf{u}_0^n\|_{H^2}^2 \le C\|\mathbf{G}\|_{L^2}^2 \le C\|\mathbf{u}_0\|_{\mathbb{D}}^2,$$

so $\mathbf{u}_0^n$ is uniformly bounded in H^2. Expressing $\mathbf{f}_n(\mathbf{v})$ from equation (4.5), we see that $\mathbf{f}_n(\mathbf{u}_0^n)$ are also uniformly bounded.

Let us verify the convergence. Since $\mathbf{u}_0^n$ is uniformly bounded, passing to a subsequence if necessary, we may assume that $\mathbf{u}_0^n \rightharpoonup \mathbf{w}$ as $n \to \infty$ and $\mathbf{u}_0^n \to \mathbf{w}$ strongly to $\mathbf{w}$ in H^1. Then, we have the convergence $\mathbf{u}_0^n(x) \to \mathbf{w}(x)$ almost everywhere. Moreover, from this convergence and Lemma 4.1, we may conclude that $\mathbf{f}_n(\mathbf{u}_0^n(x)) \to \mathbf{f}(\mathbf{w}(x))$ almost everywhere. Since $\mathbf{f}_n(\mathbf{u}_0^n)$ are uniformly bounded in L^2, passing to a subsequence again, we infer that $\mathbf{f}_n(\mathbf{u}_0^n) \rightharpoonup \mathbf{f}(\mathbf{w})$. Passing after that to the weak limit $n \to \infty$ in equations (4.5), we see that the limit function $\mathbf{w}$ solves

$$\mathbf{D}\Delta_x \mathbf{w} - \mathbf{f}(\mathbf{w}) - K\mathbf{w} = \mathbf{G} = \mathbf{D}\Delta_x \mathbf{u}_0 - \mathbf{f}(\mathbf{u}_0) - K\mathbf{u}_0, \quad \mathbf{w}\big|_{\partial\Omega} = 0. \tag{4.7}$$

Finally, since the solution $\mathbf{w} \in \mathbb{D}$ of equation (4.7) is unique (again due to the monotonicity of $\mathbf{f}(\mathbf{u}) + K\mathbf{u}$), we conclude that $\mathbf{w} = \mathbf{u}_0$. The uniqueness also gives that passing to a subsequence was not necessary and the whole sequence $\mathbf{u}_0^n$ converges to $\mathbf{u}_0$. $\qquad\square$

We are now ready to state and prove the main result of this section.

Theorem 4.3 *Let the nonlinearity $\mathbf{f}$ and matrix $\mathbf{D}$ satisfy assumptions (2.3) and (2.2), $\mathbf{g} \in L^2$ and $\mathbf{u}_0 \in \mathbb{D}$. Then, problem (2.1) possesses a unique strong solution $\mathbf{u}(t) \in \mathbb{D}$ which satisfies all estimates formally obtained in Section 3.*

Proof We approximate the desired solution $\mathbf{u}$ by the approximate solutions $\mathbf{u}_n(t)$ of problems (4.4), where $\mathbf{f}_n$ and $\mathbf{u}_0^n$ are chosen as in Lemmas 4.1 and 4.2. Then, since $\mathbf{f}_n$ is globally Lipschitz continuous, the existence and uniqueness of a solution $\mathbf{u}_n(t)$ of (4.4) is straightforward. At the next step, we need to check that all estimates of Section 3 are indeed uniform with respect to n (the justification of all these estimates for the case of sublinear growth rate is also obvious). This is obvious for Lemmas 3.1 and 3.2 (as well as for Lemma 3.6) since $\mathbf{u}_0^n \to \mathbf{u}_0$ *strongly* in H^1 and $\mathbf{f}_n$ satisfy (2.3) uniformly with respect to n. Thus, we only need to look on the estimates related with time differentiation and $\mathbb{D}$ norms. The key role in these estimates is played by the

L^2-norm of time derivative $\boldsymbol{\theta}_n(t) := \partial_t \mathbf{u}_n(t)$ and the L^2-norm of it at time moment t is estimated by its L^2-norm at time $t = 0$. But due to our construction

$$\boldsymbol{\theta}_n(0) = \mathbf{D}\Delta_x \mathbf{u}_0^n - \mathbf{f}_n(\mathbf{u}_0^n) + \mathbf{g} = \mathbf{D}\Delta_x \mathbf{u}_0 - \mathbf{f}(\mathbf{u}_0) + \mathbf{g} + K(\mathbf{u}_0^n - \mathbf{u}_0). \quad (4.8)$$

Therefore, according to Lemma 4.2,

$$\|\boldsymbol{\theta}_n(0)\|_{L^2}^2 \leq C(\|\mathbf{u}_0\|_{\mathbb{D}}^2 + \|\mathbf{g}\|_{L^2}^2 + 1). \quad (4.9)$$

By this reason, the analogue of estimate (3.6) on the level of approximations reads

$$\|\partial_t \mathbf{u}_n(t)\|_{L^2}^2 + \|\mathbf{u}_n(t)\|_{H^2}^2 + \|\mathbf{f}_n(\mathbf{u}_n(t))\|_{L^2}^2 \leq Ce^{K_1 t}(\|\mathbf{u}_0\|_{\mathbb{D}}^2 + \|\mathbf{g}\|_{L^2}^2 + 1), \quad (4.10)$$

where positive constants C and K_1 are independent of t, $\mathbf{u}_0$ and n.

When the uniform estimates are obtained, we may pass to the limit $n \to \infty$ in equations (4.4) and construct the desired strong solution $\mathbf{u}(t)$ of the limit problem (2.1) (the passage to the limit in the nonlinear term is done exactly as in Lemma 4.2). The uniqueness of a solution is an immediate corollary of Lemma 3.6 (which does not require justification on the level of strong solutions). Finally, passing to the limit in the corresponding estimates for $\mathbf{u}_n$, we prove that the limit solution $\mathbf{u}(t)$ satisfies indeed all of the estimates of Section 3. The only non-immediate thing is the passage to the limit in the terms like $(\nabla_{\mathbf{u}}\mathbf{f}_n(\mathbf{u}_n)\nabla_x \mathbf{u}_n, \nabla_x \mathbf{u}_n)$ (and in the analogous term containing $\partial_t \mathbf{u}_n$) since we do not have any control of the integral norms of $\nabla_{\mathbf{u}}\mathbf{f}_n(\mathbf{u}_n)$. However, the passage to the limit could be performed here using the condition $\nabla_{\mathbf{u}}\mathbf{f}_n(\mathbf{u}_n) \geq -K$, the fact that $\nabla_{\mathbf{u}}\mathbf{f}_n(\mathbf{u}_n(t,x)) \to \nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}(t,x))$ almost everywhere and the convexity arguments. Namely, under these conditions, we may establish that

$$\int_T^{T+1} (\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}(t))\nabla_x \mathbf{u}(t), \nabla_x \mathbf{u}(t)) \, dt \leq$$
$$\leq \liminf_{n \to \infty} \int_T^{T+1} (\nabla_{\mathbf{u}}\mathbf{f}_n(\mathbf{u}_n(t))\nabla_x \mathbf{u}_n(t), \nabla_x \mathbf{u}_n(t)) \, dt, \quad (4.11)$$

see e.g [4], Theorem 5.4. Thus, the theorem is proved.   □

5 Weak solutions, dissipativity and attractors

In the last section, we have proved the global existence and uniqueness of *strong* solutions of (2.1). Thus, the solution semigroup

$$S(t) : \mathbb{D} \to \mathbb{D}, \quad S(t)\mathbf{u}_0 := \mathbf{u}(t) \quad (5.1)$$

is well-defined. Moreover, according to Lemma 3.6, this semigroup is globally Lipschitz continuous in the L^2-metric:

$$\|S(t)\mathbf{u}_0^1 - S(t)\mathbf{u}_0^2\|_{L^2}^2 \leq Ce^{K_1 t}\|\mathbf{u}_0^1 - \mathbf{u}_0^2\|_{L^2}^2, \quad \mathbf{u}_0^1, \mathbf{u}_0^2 \in \mathbb{D}. \quad (5.2)$$

Thus, we can extend this semigroup by continuity from $\mathbb{D}$ to its closure in L^2 which obviously coincides with the whole L^2 (since $C^\infty \subset \mathbb{D}$). Thus, the semigroup

$$\widehat{S}(t)\mathbf{u}_0 := \lim_{n\to\infty} S(t)\mathbf{u}_0^n, \quad \mathbf{u}_0^n \in \mathbb{D}, \quad \mathbf{u}_0 = \lim_{n\to\infty} \mathbf{u}_0^n \tag{5.3}$$

is well-defined. Moreover, the limit in (5.3) can be considered in the space $C(0, T; L^2(\Omega))$, so the trajectories $\mathbf{u}(t) := \widehat{S}(t)\mathbf{u}_0$ automatically belong to the space $C(0, T; L^2(\Omega))$ for all $T \geq 0$.

Our next step is to understand in what sense the trajectory $\mathbf{u}(t)$ thus constructed satisfies the initial equation (2.1). Note that, for the general $\mathbf{u}_0 \in L^2$, we do not know whether or not $\mathbf{f}(\mathbf{u}) \in L^1$, so we cannot treat it in the distributional sense. Indeed, the only control related with $\mathbf{f}(\mathbf{u})$ which we have up to now follows from estimate (3.1) and claims that $\mathbf{f}(\mathbf{u}).\mathbf{u} \in L^1$ (being pedantic, this is proved for the strong solutions only, but it can be easily extended to weak solutions using the Fatou lemma). Unfortunately, this is not enough to control the L^1-norm of the function $\mathbf{f}$ itself, so we cannot treat the term $\mathbf{f}(\mathbf{u})$ in a distributional sense.

Instead, we use the ideas from the monotone operator theory and variational inequalities, see e.g., [5]. Namely, following [21], we take an arbitrary test function

$$\mathbf{v} \in C_\mathrm{w}(0, T; \mathbb{D}) \cap C_\mathrm{w}^1(0, T; L^2(\Omega)), \quad \forall T \in \mathbb{R}_+, \tag{5.4}$$

and multiply formally equation (2.1) by $\mathbf{u}(t) - \mathbf{v}(t)$ and integrate over $[0, T]$. Then, integrating by parts and using that

$$-(\mathbf{D}\Delta_x\mathbf{u} - \mathbf{D}\Delta_x\mathbf{v}, \mathbf{u} - \mathbf{v}) \geq 0, \quad (\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{v}), \mathbf{u} - \mathbf{v}) \geq -K\|\mathbf{u} - \mathbf{v}\|_{L^2}^2,$$

we end up with

$$\frac{1}{2}\|\mathbf{u}(T) - \mathbf{v}(T)\|_{L^2}^2 - \frac{1}{2}\|\mathbf{u}(0) - \mathbf{v}(0)\|_{L^2}^2 +$$
$$+ \int_0^T (\partial_t\mathbf{v}(t), \mathbf{u}(t) - \mathbf{v}(t))\, dt \leq$$
$$\leq \int_0^T (\mathbf{D}\Delta_x\mathbf{v}(t) - \mathbf{f}(\mathbf{v}(t)) + \mathbf{g}, \mathbf{u}(t) - \mathbf{v}(t))\, dt + K\int_0^T \|\mathbf{u}(t) - \mathbf{v}(t)\|_{L^2}^2\, dt. \tag{5.5}$$

The advantage of this approach is that the variational inequality (5.5) makes sense for all $\mathbf{u} \in C(0, T; L^2(\Omega))$ and, therefore, can be used to define a weak solution $\mathbf{u}(t)$ of problem (2.1).

Definition 5.1 A function $\mathbf{u} \in C(0, T; L^2(\Omega))$, $T \in \mathbb{R}_+$ is a weak solution of problem (2.1) if $\mathbf{u}(0) = \mathbf{u}_0$ and the variational inequality (5.5) holds for every $T \in \mathbb{R}_+$ and every test function v satisfying (5.4).

We are now ready to state the key result of this section.

Theorem 5.2 *Let* $\mathbf{g} \in L^2(\Omega)$, *the diffusion matrix* $\mathbf{D}$ *enjoy* $\mathbf{D} + \mathbf{D}^* > 0$ *and* $\mathbf{f}$ *satisfy assumptions* (2.3). *Then, for every* $\mathbf{u}_0 \in L^2(\Omega)$, *there exists a unique weak solution* $\mathbf{u}(t)$ *of problem* (2.1) *and this solution has the form* $\mathbf{u}(t) = \widehat{S}(t)\mathbf{u}_0$, *where the extension* $\widehat{S}(t)$ *of the solution semigroup* $S(t)$ *is defined by* (5.3).

Proof Indeed, any *strong* solution is a weak solution of (2.1) (since all manipulations used in the derivation of the variational inequality (5.5) are justified on the level of strong solutions). Let now $\mathbf{u}(t) = \lim_{n\to\infty} \mathbf{u}_n(t)$, where $\mathbf{u}_n(t)$ are the strong solutions $\mathbf{u}_n(t) := S(t)\mathbf{u}_0^n$, $\mathbf{u}_0^n \in \mathbb{D}$ and $\mathbf{u}_0^n \to \mathbf{u}_0$ in L^2. The variational inequality for $\mathbf{u}_n$ reads

$$\frac{1}{2}\|\mathbf{u}_n(T) - \mathbf{v}(T)\|_{L^2}^2 - \frac{1}{2}\|\mathbf{u}_n(0) - \mathbf{v}(0)\|_{L^2}^2 +$$
$$+ \int_0^T (\partial_t \mathbf{v}(t), \mathbf{u}_n(t) - \mathbf{v}(t))\, dt \le$$
$$\le \int_0^T (\mathbf{D}\Delta_x \mathbf{v}(t) - \mathbf{f}(\mathbf{v}(t)) + \mathbf{g}, \mathbf{u}_n(t) - \mathbf{v}(t))\, dt + K \int_0^T \|\mathbf{u}_n(t) - \mathbf{v}(t)\|_{L^2}^2\, dt.$$
$$(5.6)$$

Using that $\mathbf{u}_n \to \mathbf{u}$ in $C(0, T; L^2(\Omega))$ and passing to the limit $n \to \infty$ in (5.6), we see that $\mathbf{u}(t)$ satisfies (5.5) and, therefore, $\mathbf{u}(t) := \widehat{S}(t)\mathbf{u}_0$ is the desired weak solution of (2.1).

Vice versa, let $\bar{\mathbf{u}}(t)$ be a weak solution of (2.1) and let $\mathbf{u}(t) := \widehat{S}(t)\bar{\mathbf{u}}(0)$. Then, there exists a sequence $\mathbf{u}_0^n \in \mathbb{D}$ such that $\mathbf{u}_0^n \to \bar{\mathbf{u}}(0)$ and the sequence of *strong* solutions $\mathbf{u}_n(t) = S(t)\mathbf{u}_0^n$ which converges as $n \to \infty$ to the weak solution $\mathbf{u}(t)$. We need to show that $\mathbf{u}(t) = \bar{\mathbf{u}}(t)$.

Indeed, by the definition of a strong solution, $\mathbf{u}_n(t)$ satisfies the assumptions of (5.4) and therefore can be used as a test function $\mathbf{v} = \mathbf{u}_n$ in the variational inequality (5.5) for the weak solution $\bar{\mathbf{u}}$. Taking $\mathbf{v} = \mathbf{u}_n$ in it and using that $\partial_t \mathbf{u}_n = \mathbf{D}\Delta_x \mathbf{u}_n - \mathbf{f}(\mathbf{u}_n) + \mathbf{g}$, we get

$$\frac{1}{2}\|\bar{\mathbf{u}}(T) - \mathbf{u}_n(T)\|_{L^2}^2 \le K \int_0^T \|\bar{\mathbf{u}}(t) - \mathbf{u}_n(t)\|_{L^2}^2\, dt + \frac{1}{2}\|\bar{\mathbf{u}}(0) - \mathbf{u}_n(0)\|_{L^2}^2.$$

Passing to the limit $n \to \infty$ in this inequality, we get

$$\|\bar{\mathbf{u}}(T) - \mathbf{u}(T)\|_{L^2}^2 \le 2K \int_0^T \|\bar{\mathbf{u}}(t) - \mathbf{u}(t)\|_{L^2}^2\, dt$$

and, since T is arbitrary, the Gronwall inequality gives that $\bar{\mathbf{u}}(t) = \mathbf{u}(t)$ for all t. Thus, the theorem is proved. $\square$

Remark 5.3 There is an alternative possibility to relate a weak solution $\mathbf{u}(t)$ with equation (2.1), namely, for every $\boldsymbol{\psi} \in C_0^\infty(\mathbb{R}^k, \mathbb{R}^k)$, the identity

$$\partial_t(\boldsymbol{\psi}(\mathbf{u})) = \nabla_\mathbf{u} \boldsymbol{\psi}(\mathbf{u})\mathbf{D}\Delta_x \mathbf{u} - \nabla_\mathbf{u} \boldsymbol{\psi}(\mathbf{u})\mathbf{f}(\mathbf{u}) + \nabla_\mathbf{u} \boldsymbol{\psi}(\mathbf{u})\mathbf{g} \qquad (5.7)$$

should be satisfied in the sense of distributions, see [46]. It is not difficult to verify that, indeed, any weak solution $\mathbf{u}(t)$ should satisfy this identity. The drawback of this approach is that it is unclear whether or not (5.7) is enough for the uniqueness.

Using identity (5.7) it is not difficult to show that under the additional restriction

$$|\mathbf{f}(\mathbf{u})| \leq C(|\mathbf{f}(\mathbf{u}).\mathbf{u}| + |\mathbf{u}| + 1) \tag{5.8}$$

which guarantees that $\mathbf{f}(\mathbf{u}) \in L^1([0, T] \times \Omega)$, any weak solution satisfies equation (2.1) in a sense of distributions. However, in contrast to the scalar case $k = 1$, in the case of systems (5.8) is an extra restriction which we prefer to avoid.

As a next step, we note that the weak solutions are dissipative. Indeed, passing to the limit in the estimate of Lemma 3.1 for strong solutions, we derive that

$$\|\widehat{S}(T)\mathbf{u}_0\|_{L^2}^2 + \int_T^{T+1} \|\widehat{S}(t)\mathbf{u}_0\|_{H^1}^2 \, dt \leq Ce^{-\alpha T}\|\mathbf{u}_0\|_{L^2}^2 + C(1 + \|\mathbf{g}\|_{L^2}^2) \tag{5.9}$$

which is a standard dissipative estimate for the semigroup $\widehat{S}(t)$. Analogously, passing to the limit in the estimate of Lemma 3.2, we get the dissipative estimate in H^1 for weak solutions

$$\|\widehat{S}(T)\mathbf{u}_0\|_{H^1}^2 + \int_T^{T+1} \|\widehat{S}(t)\mathbf{u}_0\|_{H^2}^2 \, dt \leq Ce^{-\alpha T}\|\mathbf{u}_0\|_{H^1}^2 + C(1 + \|\mathbf{g}\|_{L^2}^2). \tag{5.10}$$

In addition, estimates (5.9) and (5.10) give in a standard way the L^2-H^1 smoothing property for the semigroup $\widehat{S}(t)$, namely, the following estimate holds:

$$\|\widehat{S}(t)\mathbf{u}_0\|_{H^1}^2 \leq C\frac{t+1}{t}\left(e^{-\alpha t}\|\mathbf{u}_0\|_{L^2}^2 + 1 + \|\mathbf{g}\|_{L^2}^2\right). \tag{5.11}$$

These estimates ensure us that the ball $\mathcal{B}_R := \{\mathbf{u} \in H_0^1, \ \|\mathbf{u}\|_{H^1} \leq R\}$ will be a compact (in L^2) *absorbing* ball for the semigroup $\widehat{S}(t)$ if $R = R(\|\mathbf{g}\|_{L^2})$ is large enough. Remind that the latter means that for every bounded set B of L^2 there exists a time $T = T(B)$ such that

$$\widehat{S}(t)B \subset \mathcal{B}_R$$

for all $t \geq T$. This fact, in turn, allows us to establish the existence of a global attractor for the solution semigroup $\widehat{S}(t)$ in the phase space $H = L^2(\Omega)$. We recall that, by definition, a set $\mathcal{A} \subset H$ is a global attractor for a semigroup $\widehat{S}(t) : H \to H$ if the following conditions are satisfied:

(1) $\mathcal{A}$ is a *compact* subset of H;
(2) $\mathcal{A}$ is strictly invariant: $\widehat{S}(t)\mathcal{A} = \mathcal{A}$, for all $t \geq 0$;
(3) It attracts the images of bounded sets as $t \to \infty$, namely, for any bounded $B \subset H$ and any neighbourhood $\mathcal{O}(\mathcal{A})$, there exists $T = T(B, \mathcal{O})$ such that

$$\widehat{S}(t)B \subset \mathcal{O}(\mathcal{A})$$

for all $t \geq T$.

The next theorem may be considered as the second key result of this section.

Theorem 5.4 *Let the assumptions of Theorem 5.2 be satisfied. Then the weak solution semigroup $\widehat{S}(t)$ possesses a global attractor $\mathcal{A}$ in $H = L^2(\Omega)$ which is a bounded set of $H^1(\Omega)$ and possesses the following description:*

$$\mathcal{A} = \mathcal{K}\big|_{t=0}, \tag{5.12}$$

where $\mathcal{K} \subset C_b(\mathbb{R}, L^2(\Omega))$ is the set of all complete bounded trajectories of the semigroup $\widehat{S}(t)$:

$$\mathcal{K} := \{\mathbf{u} \in C_b(\mathbb{R}, L^2(\Omega)), \ \mathbf{u}(t+h) = \widehat{S}(t)\mathbf{u}(h), \ h \in \mathbb{R}, \ t \in \mathbb{R}_+\}. \tag{5.13}$$

Proof According to the abstract attractor existence theorem, see e.g., [3], we need to verify two assumptions:

(1) The operators $\widehat{S}(t) : H \to H$ are continuous for every fixed t;
(2) The semigroup $\widehat{S}(t)$ possesses a compact absorbing set in H.

The first assumption is guaranteed by Lemma 3.6 and the second one follows from estimate (5.11). Thus, the global attractor $\mathcal{A}$ exists. The fact that $\mathcal{A}$ is a bounded subset of H^1 follows from the fact that the attractor is always a subset of an absorbing set and the representation formula (5.12) is a standard corollary of the attractor existence theorem. Thus, the theorem is proved. □

6 Weak to strong smoothing property

In this section, we establish that any weak solution $\mathbf{u}(t)$ of problem (2.1) becomes strong for $t > 0$. The main difficulty here is the fact that we cannot in general estimate $|\mathbf{f}(\mathbf{u})|$ through $\mathbf{f}(\mathbf{u}).\mathbf{u}$ and, by this reason, we do not know whether or not $\mathbf{f}(\mathbf{u})$ and $\partial_t \mathbf{u}$ are distributions. This makes the situation with the parabolic smoothing property a bit more delicate than usual. We overcome this difficulty under the extra assumption that

$$|\mathbf{f}(\mathbf{u})| \leq C(1 + |\mathbf{u}|^p) \tag{6.1}$$

for some $p > 1$ by using the L^q-spaces with $q < 1$. Namely, we will use the fact that

$$\|\mathbf{f}(\mathbf{u})\|_{L^{2/p}} \leq C(\|\mathbf{u}\|_{L^2}^p + 1), \tag{6.2}$$

where, for $q < 1$, we denote by $\|\mathbf{v}\|_{L^q}$ exactly the same expression as for the case $q \geq 1$ (simply ignoring the fact that it is no more a norm). Thus, at least on the level of approximations, we may expect that, for $\boldsymbol{\theta} = \partial_t \mathbf{u}$,

$$\|\boldsymbol{\theta}(t)\|_{L^2(0,1;L^2(\Omega))+L^\infty(0,1;L^{2/p}(\Omega))} \leq C(\|\mathbf{u}_0\|_{H^1}^p + \|\mathbf{g}\|_{L^2}^p + 1) \tag{6.3}$$

and this can be used in order to establish the smoothing property for $\boldsymbol{\theta}$. Namely, the following theorem holds.

Theorem 6.1 *Let $d \geq 3$, the assumptions of Theorem 5.2 and (6.1) hold, and let $\mathbf{u}_0 \in H^1$ and $\mathbf{u}(t)$ be the corresponding weak solution of equation (2.1). Then, $\mathbf{u}(t) \in \mathbb{D}$ for all $t > 0$ and the following estimate is valid:*

$$\|\partial_t \mathbf{u}(t)\|_{L^2}^2 + \|\mathbf{u}(t)\|_{\mathbb{D}}^2 \leq Ct^{-N}\left(Q(\|\mathbf{u}_0\|_{H^1}) + Q(\|\mathbf{g}\|_{L^2})\right), \quad t \in (0, 1], \quad (6.4)$$

where the exponent $N = N(d, p)$, positive constant C and the monotone increasing function Q are independent of $\mathbf{u}_0$ and t. Moreover, the function $Q(z)$ can be chosen as a polynomial of z.

Proof We first note that it is enough to verify (6.4) for $\mathbf{u}_0 \in \mathbb{D}$ only when $\mathbf{u}(t)$ is a *strong* solution. Moreover, it is enough to obtain the estimate for $\partial_t \mathbf{u}$ only since the estimate for $\|\mathbf{u}(t)\|_{\mathbb{D}}$ will follow from the elliptic problem (3.7). Second, we approximate the strong solution $\mathbf{u}(t)$ by the solutions $\mathbf{u}_n(t)$ of auxiliary problems (4.4). Finally, analyzing the proof of Lemma 4.1, we see that assumption (6.1) allows us to pose the extra assumption

$$|\mathbf{f}_n(\mathbf{u})| \leq C(1 + |\mathbf{u}|^{p_1}) \quad (6.5)$$

for some $p_1 > p > 1$ and constant C independent of n (e.g., by taking $\Psi(\mathbf{u}) := |\mathbf{u}|^{p_1+1}$).

Let $\boldsymbol{\theta}_n(t) := \partial_t \mathbf{u}_n(t)$. Then, this function solves the equation

$$\partial_t \boldsymbol{\theta}_n = \mathbf{D}\Delta_x \boldsymbol{\theta}_n - \nabla_{\mathbf{u}} \mathbf{f}_n(\mathbf{u}_n)\boldsymbol{\theta}_n. \quad (6.6)$$

Multiplying this equation by $t^N \boldsymbol{\theta}_n(t)$ where $N > 1$ is a sufficiently big number, we end up with

$$\frac{d}{dt}(t^N \|\boldsymbol{\theta}_n(t)\|_{L^2}^2) + \alpha t^N \|\boldsymbol{\theta}_n(t)\|_{H^1}^2 - $$
$$-2K(t^N \|\boldsymbol{\theta}_n(t)\|_{L^2}^2) \leq Ct^{N-1}\|\boldsymbol{\theta}_n(t)\|_{L^2}^2. \quad (6.7)$$

We need to estimate the integral in the right-hand side. To this end, we fix some $s \in (0, 2)$ which will be specified below and write

$$\|\boldsymbol{\theta}_n\|_{L^2}^2 = (|\boldsymbol{\theta}_n|^2, 1) = (|\boldsymbol{\theta}_n|^s, |\boldsymbol{\theta}_n|^{2-s}) \leq$$
$$\leq C(|\Delta_x \mathbf{u}_n|^s, |\boldsymbol{\theta}_n|^{2-s}) + C(|\mathbf{g}|^s, |\boldsymbol{\theta}_n|^{2-s}) + C(|\mathbf{f}_n(\mathbf{u}_n(s))|^s, |\boldsymbol{\theta}_n|^{2-s}), \quad (6.8)$$

where we have used equation (4.4) in order to express $\boldsymbol{\theta}_n$ through $\mathbf{u}_n$. Let us estimate every term in the RHS separately. Applying the Hölder and Young inequalities to the first term, we get

$$t^{N-1}(|\Delta_x \mathbf{u}_n|^s, |\boldsymbol{\theta}_n|^{2-s}) \leq C\|\Delta_x \mathbf{u}_n\|_{L^2}^s \left(t^{\frac{N-1}{2-s}}\|\boldsymbol{\theta}_n\|_{L^2}\right)^{2-s} \leq$$

$$\leq C\|\Delta_x \mathbf{u}_n\|_{L^2}^2 + t^{2\frac{N-1}{2-s}}\|\boldsymbol{\theta}_n\|_{L^2}^2. \tag{6.9}$$

This gives us a good estimate if N is chosen in such a way that $2\frac{N-1}{2-s} \geq N$. The second term in the RHS of (6.8) can be estimated analogously to have

$$t^{N-1}(|\mathbf{g}|^s, |\boldsymbol{\theta}_n|^{2-s}) \leq C\|\mathbf{g}\|_{L^2}^2 + t^{2\frac{N-1}{2-s}}\|\boldsymbol{\theta}_n\|_{L^2}^2. \tag{6.10}$$

Let us now estimate the most complicated third term. To this end, we use the embedding theorem $H^1 \subset L^r$ where $\frac{1}{r} = \frac{1}{2} - \frac{1}{d}$ together with the Hölder inequality with exponents q_1 and q_2, $\frac{1}{q_1} + \frac{1}{q_2} = 1$, with $(2-s)q_2 = r$ to get

$$(|\mathbf{f}_n(\mathbf{u}_n)|^s, |\boldsymbol{\theta}_n|^{2-s}) \leq \|\mathbf{f}_n(\mathbf{u}_n)\|_{L^{sq_1}}^s \|\boldsymbol{\theta}_n\|_{H^1}^{2-s}.$$

Moreover, due to assumption (6.5), we have

$$\|\mathbf{f}_n(\mathbf{u}_n)\|_{L^{2/p_1}} \leq C(\|\mathbf{u}_n\|_{L^2} + 1)^{p_1} \leq C(\|\mathbf{u}_0\|_{L^2} + 1 + \|\mathbf{g}\|_{L^2})^{p_1},$$

where C is independent of n. We may also fix $sq_1 = \frac{2}{p_1}$ to get

$$(|\mathbf{f}_n(\mathbf{u}_n)|^s, |\boldsymbol{\theta}_n|^{2-s}) \leq C(\|\mathbf{u}_0\|_{L^2} + 1 + \|\mathbf{g}\|_{L^2})^{sp_1}\|\boldsymbol{\theta}_n\|_{H^1}^{2-s} \tag{6.11}$$

and end up with the following system for the exponents q_1, q_2 and s

$$\frac{1}{q_1} + \frac{1}{q_2} = 1, \quad \frac{1}{q_1} = \frac{sp_1}{2}, \quad \frac{1}{q_2} = (2-s)\left(\frac{1}{2} - \frac{1}{d}\right).$$

Solving this system, we get

$$s = \frac{4}{d(p_1-1)+2}, \quad q_1 = \frac{d(p_1-1)+2}{2p_1}$$

and we see that $0 < s < 2$ and $1 < q_1 < \infty$, so all of the exponents are in the prescribed range and (6.11) holds indeed. Applying the Young inequality, we arrive at

$$t^{N-1}(|\mathbf{f}_n(\mathbf{u}_n)|^s, |\boldsymbol{\theta}_n|^{2-s}) \leq Q_\varepsilon(\|\mathbf{u}_0\|_{L^2}) + Q_\varepsilon(\|\mathbf{g}\|_{L^2}) + \varepsilon t^{2\frac{N-1}{2-s}}\|\boldsymbol{\theta}_n\|_{H^1}^2, \tag{6.12}$$

where $\varepsilon > 0$ is arbitrary small and the polynomial monotone function Q_ε is independent of $\mathbf{u}_0$. Combining estimates (6.8), (6.9), (6.10) and (6.12) for estimating the RHS of (6.7) and fixing $\varepsilon > 0$ to be small enough and N satisfying $2\frac{N-1}{2-s} \geq N$, we arrive at

$$\frac{d}{dt}\left(t^N \|\boldsymbol{\theta}_n(t)\|_{L^2}^2\right) \leq K_1\left(t^N \|\boldsymbol{\theta}_n(t)\|_{L^2}^2\right) + C\|\Delta_x \mathbf{u}_n(t)\|_{L^2}^2$$

$$+ Q(\|\mathbf{u}_0^n\|_{H^1}) + Q(\|\mathbf{g}\|_{L^2}), \quad t \leq 1. \qquad (6.13)$$

Applying the Gronwall inequality to this relation and using (5.10) for estimating the integral of the H^2-norm of the solution, we end up with

$$t^N \|\boldsymbol{\theta}_n(t)\|_{L^2}^2 \leq Q(\|\mathbf{u}_0^n\|_{H^1}) + Q(\|\mathbf{g}\|_{L^2}), \quad t \in (0, 1] \qquad (6.14)$$

and passing to the limit $n \to \infty$, we derive the desired estimate for $\partial_t \mathbf{u}(t)$. Thus, the theorem is proved. □

Combining this result with the L^2 to H^1 smoothing property (5.11), we get the following result.

Corollary 6.2 *Under the assumptions of Theorem 6.1, the weak solution semigroup $\widehat{S}(t)$ possesses the following smoothing property:*

$$\|\widehat{S}(t)\mathbf{u}_0\|_{\mathbb{D}}^2 \leq Ct^{-M} \left(Q(\|\mathbf{u}_0\|_{L^2}) + Q(\|\mathbf{g}\|_{L^2}) \right), \quad t \in (0, 1], \qquad (6.15)$$

where the positive constants C and $M = M(d, p)$ and monotone polynomial function Q are independent of t and $\mathbf{u}_0 \in L^2(\Omega)$.

Thus, under the assumption (6.1), any weak solution indeed becomes strong for $t > 0$ (the extra assumption $d \geq 3$ is not essential since for $d \leq 2$ the equation is subcritical and the smoothing property is obvious). This, in particular, gives the following result on the regularity of the global attractor.

Corollary 6.3 *Let the assumptions of Theorem 6.1 hold. Then the global attractor $\mathcal{A}$ of the solution semigroup $\widehat{S}(t)$ constructed in Theorem 5.4 is a bounded subset of $\mathbb{D}$:*

$$\|\mathcal{A}\|_{\mathbb{D}} \leq Q(\|\mathbf{g}\|_{L^2}) \qquad (6.16)$$

for some monotone increasing function Q.

Indeed, this assertion is an immediate corollary of (6.15) and the strict invariance of the global attractor.

7 Further regularity and strong attraction

In this section we study the problem whether or not the constructed strong solution $\mathbf{u}$ is actually more regular (than $\mathbf{u} \in C_w(0, T; \mathbb{D})$). We start with some partial result on the regularity of $\partial_t \mathbf{u}$ which does not require any extra assumptions on $\mathbf{f}$ and $\mathbf{g}$.

Proposition 7.1 *Let the assumptions of Theorem 4.3 hold. Then, there exists a positive number $r > 0$ depending only on the matrix $\mathbf{D}$ such that, for any strong solution $\mathbf{u}(t) \in \mathbb{D}$, the following estimate holds:*

$$t\|\partial_t \mathbf{u}(t)\|_{L^{r+2}}^{r+2} + \int_0^t s\|\nabla_x(|\partial_t \mathbf{u}(s)|^{\frac{r+2}{2}})\|_{L^2}^2 \, ds \leq C\|\partial_t \mathbf{u}(0)\|_{L^2}^{r+2}, \qquad (7.1)$$

where $0 \leq t \leq 1$ and the constant C is independent of t and **u**.

Proof Let $\theta := \partial_t \mathbf{u}$. Then this function satisfies equation (3.4). Let us multiply this equation by $\theta |\theta|^r$ and integrate in x. This gives

$$\frac{1}{r+2} \frac{d}{dt} \|\theta(t)\|_{L^{r+2}}^{r+2} - (\mathbf{D}\Delta_x\theta, \theta|\theta|^r) + (\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\theta, \theta|\theta|^r) = 0.$$

Integrating by parts in the second term, we get

$$-(\mathbf{D}\Delta_x\theta, \theta|\theta|^r) \geq$$
$$\geq (\mathbf{D}\nabla_x\theta, \nabla_x\theta|\theta|^r) - Cr(|\nabla_x\theta|^2, |\theta|^r) \geq (\alpha - Cr)(|\nabla_x\theta|^2, |\theta|^r)$$

for some positive α. Fixing now $r > 0$ small enough and estimating the term containing **f** using $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}) \geq -K$, we arrive at

$$\frac{1}{r+2} \frac{d}{dt} \|\theta(t)\|_{L^{r+2}}^{r+2} + \frac{\alpha}{2}(|\nabla_x\theta|^2, |\theta(t)|^r) \leq K\|\theta(t)\|_{L^{r+2}}^{r+2}.$$

Multiplying this estimate by t and integrating in time, we arrive at

$$t\|\theta(t)\|_{L^{r+2}}^{r+2} + \int_0^t s(|\nabla_x\theta(s)|^2, |\theta(s)|^r)\, ds \leq C \int_0^t \|\theta(s)\|_{L^{r+2}}^{r+2}\, ds \qquad (7.2)$$

for $0 \leq t \leq 1$. To estimate the right-hand side of this inequality, we use estimate (3.5) and Sobolev embedding theorem together with the Hölder inequality which gives that, for sufficiently small $r = r(d) > 0$,

$$\int_0^t \|\theta(s)\|_{L^{r+2}}^{r+2}\, ds \leq C \left(\|\theta\|_{L^\infty(0,t;L^2)} + \|\theta\|_{L^2(0,t;H^1)}\right)^{r+2} \leq C\|\theta(0)\|_{L^2}^{r+2}$$

and finishes the proof of the proposition. $\square$

The obtained extra regularity in time can be transformed to extra regularity in space assuming that the right-hand side **g** is slightly more regular.

Corollary 7.2 *Let the assumptions of Proposition 7.1 hold and let, in addition,*

$$\mathbf{g} \in L^q(\Omega) \qquad (7.3)$$

for some $q > 2$. Then, there exists $r = r(\mathbf{D}, q) > 0$ such that

$$t\|\nabla_x\mathbf{u}(t)\|_{L^{\frac{d(r+2)}{d-2}}}^{r+2} \leq C \left(\|\mathbf{u}(0)\|_{\mathbb{D}} + \|\mathbf{g}\|_{L^q}\right)^{r+2} \qquad (7.4)$$

for $0 \leq t \leq 1$. In particular, if the attractor $\mathcal{A}$ is a bounded set in $\mathbb{D}$ then it is also bounded in $W^{1, \frac{d(r+2)}{d-2}}(\Omega)$.

Indeed, due to Proposition 7.1, we control the L^{r+2}-norm of $\partial_t \mathbf{u}$. Rewriting problem (2.1) as an elliptic boundary value problem

$$\mathbf{D}\Delta_x \mathbf{u}(t) - \mathbf{f}(\mathbf{u}(t)) = \tilde{\mathbf{g}}(t) := \partial_t \mathbf{u}(t) - \mathbf{g},$$

we get also the control for the L^{r+2}-norm of $\tilde{\mathbf{g}}(t)$ (point-wisely in time). Applying the elliptic regularity result proved in Appendix (see Theorem A.1) to this equation, we arrive at the desired estimate (7.4).

The obtained partial regularity results allow us to establish the crucial L^∞-estimates for critical and slightly supercritical growth rate of the nonlinearity $\mathbf{f}$. Namely, the following result holds.

Theorem 7.3 *Let the assumptions of Proposition 7.1 hold and $d \geq 4$. Then there exists a constant $\varepsilon = \varepsilon(\mathbb{D}) > 0$ such that, for every nonlinearity $\mathbf{f}$ satisfying in addition (6.1) with the exponent p restricted by the assumption*

$$p < p_{crit} + \varepsilon, \quad p_{crit} := 1 + \frac{4}{d-4}, \tag{7.5}$$

and the external forces $\mathbf{g}$ satisfying (7.3) for some $q > \frac{d}{2}$, any weak solution $\mathbf{u}(t)$ of problem (2.1) possesses the following smoothing property:

$$\|\mathbf{u}(t)\|_{L^\infty} \leq Q_t(\|\mathbf{u}(0)\|_{L^2} + \|\mathbf{g}\|_{L^q}), \quad t \in (0, 1] \tag{7.6}$$

for some monotone function Q_t which depends on t, but is independent of $\mathbf{u}$ and $\mathbf{g}$.

Proof Note that, due to estimate (6.15) we may assume from the very beginning that $\mathbf{u}_0 \in \mathbb{D}$ and work with strong solutions only. The derivation of (7.6) can be done by the standard bootstrapping arguments by iterating the classical interior regularity result for the linear parabolic equation

$$\partial_t \mathbf{u} - \mathbf{D}\Delta_x \mathbf{u} = \mathbf{h}(t) \tag{7.7}$$

which we state in the following lemma. □

Lemma 7.4 *Let $\mathbf{u}$ be a weak solution of (7.7). Then the following interior regularity holds:*

$$\|\mathbf{u}\|_{L^\infty(T,T+1;W^{2-\mu,s})} \leq C_{T,s,\mu}\left(\|\mathbf{u}\|_{L^2(0,T+1;L^2)} + \|\mathbf{h}\|_{L^\infty(0,T+1;L^s)}\right), \tag{7.8}$$

where $\mu > 0$ is arbitrarily small, $1 < s < \infty$ and $T > 0$.

Sketch of the proof This estimate, in turn, can be easily deduced from the fact that this linear equation generates an analytic semigroup in $L^s(\Omega)$ (since the operator $A = -\mathbf{D}\Delta_x$ is sectorial in $L^s(\Omega)$ with the domain $W^{2,s}(\Omega) \cap W_0^{1,s}(\Omega)$, see e.g., [42]) or, alternatively, from the anisotropic maximal $L^q(L^s)$-regularity estimate for

parabolic equations. Let us sketch the first approach. Using the variation of constants formula

$$\mathbf{u}(t) = e^{-At}\mathbf{u}_0 + \int_0^t e^{-A(t-\tau)}\mathbf{h}(\tau)\, ds$$

together with the estimate

$$\|e^{-A(t-\tau)}\|_{\mathcal{L}(L^s, W^{2-\mu,s})} = \|e^{-A(t-\tau)}\|_{\mathcal{L}(L^s, D(A^{\frac{2-\mu}{2}}))} \leq \frac{C_s}{(t-\tau)^{\frac{2-\mu}{2}}},$$

we get that the solution of (7.7) with $\mathbf{u}_0(0) = 0$ satisfies

$$\|\mathbf{u}(t)\|_{W^{2-\mu,s}} \leq C_{\mu,s,q,T}\|\mathbf{h}\|_{L^q(0,T;L^s)}, \quad t \in [0,T]$$

for any $1 < s, q < \infty$ and $0 < \mu < 2$ satisfying $\frac{(2-\mu)q}{q-1} < 1$. The desired interior estimate is a standard corollary of this estimate. Indeed, applying the last estimate to the function $t^N \mathbf{u}(t)$ (where $N = N(d, s, q) > 1$ is big enough), we get

$$\|t^N \mathbf{u}\|_{L^\infty(0,T+1;W^{2-\mu,s})} \leq C\|t^{N-1}\mathbf{u}\|_{L^q(0,T+1;L^s)} + C\|\mathbf{h}\|_{L^\infty(0,T+1;L^s)}$$

and using the Sobolev embedding theorem and the Hölder inequality, we arrive at

$$\|t^{N-1}\mathbf{u}\|_{L^q(0,T+1;L^s)} \leq \frac{1}{2C}\|t^N \mathbf{u}\|_{L^\infty(0,T+1;W^{2-\mu,s})} + C'\|\mathbf{u}\|_{L^2(0,T+1;L^2)}.$$

Finally, inserting the last estimate to the previous one, we have

$$\|t^N \mathbf{u}\|_{L^\infty(0,T+1;W^{2-\mu,s})} \leq C\left(\|\mathbf{u}\|_{L^2(0,T+1;L^2)} + \|\mathbf{h}\|_{L^\infty(0,T+1;L^s)}\right)$$

which gives (7.8) and finishes the proof of the lemma. $\qquad\square$

We are now ready to continue the proof of the theorem. From the interior estimate (7.8) and Sobolev embedding theorem, we derive the iterative estimate

$$\|\mathbf{u}\|_{L^\infty(T_{k+1},1;L^{q_{k+1}})} \leq C_k\left(\|\mathbf{u}\|_{L^2(0,1;L^2)} + \|\mathbf{h}\|_{L^\infty(T_k,1;L^{s_k})}\right), \qquad (7.9)$$

where $T_{k+1} > T_k$ and

$$q_{k+1} := \min\left\{\infty, \frac{s_k d}{d - s_k(2-\mu)}\right\}.$$

In our situation $\mathbf{h}(t) = \mathbf{g} - \mathbf{f}(\mathbf{u}(t))$ and, due to our growth restriction (6.1), we have

$$\|\mathbf{h}\|_{L^{s_k}} \leq C(\|\mathbf{g}\|_{L^q} + 1 + \|\mathbf{u}\|_{L^{q_k}}^p),$$

where $s_k := \min\{q, q_k p^{-1}\}$. Thus, in order to prove the theorem, it is sufficient to verify that the sequence q_k defined via

$$q_0 = \frac{d(r+2)}{d-r-4}, \quad q_{k+1} = \frac{q_k d}{pd - q_k(2-\mu)}, \quad \mu \ll 1$$

will become larger than $q > \frac{d}{2}$ in finitely many steps (we have used here estimate (7.4) and the embedding $W^{1, \frac{d(r+2)}{d-2}} \subset L^{q_0}$ to initialize the iterations and the embedding $W^{2-\mu, q} \subset L^\infty$ which holds for sufficiently small μ due to the condition $q > \frac{d}{2}$).

Obviously this sequence will be monotone increasing if (and only if)

$$p - \frac{q_0}{d}(2-\mu) < 1.$$

Then it must converge to $+\infty$, so we only need to verify the last inequality. Using assumption (7.5) and the explicit formula for q_0, we only need the inequality

$$\frac{4}{d-4} + \varepsilon - (2-\mu)\frac{r+2}{d-r-4} < 0.$$

It remains to note that the last inequality is satisfied if $\mu \ll 1$ and $\varepsilon < \varepsilon_0 = \varepsilon_0(r)$ for some positive ε_0 if $r > 0$. This finishes the proof of the theorem.

Remark 7.5 The growth rate of the nonlinearity is no more important if the L^∞-estimate for the solutions is obtained, so further regularity can be obtained by bootstrapping exactly as in the subcritical case. Thus, under the growth restriction (7.5), the actual regularity of a solution is determined by the smoothness of Ω, $\mathbf{f}$ and $\mathbf{g}$ only (if all of them are C^∞-smooth, the solutions will be also C^∞-smooth). In other words, we may say that the critical growth exponent for $\mathbf{f}$ in our problem (2.1) is slightly *larger* than $p_{crit} = 1 + \frac{4}{d-4}$. We also note that the value $\varepsilon = \varepsilon(\mathbf{D})$ somehow measures how far the matrix $\mathbf{D}$ is from the scalar matrix. It is easy to show that $\varepsilon(\mathbf{D}) = \infty$ if $\mathbf{D}$ is scalar.

We now turn to the question of whether or not the attraction to $\mathcal{A}$ holds in the space $\mathbb{D}$. Since in this case we at least need the dissipativity of our semigroup in $\mathbb{D}$, we assume that $\mathbf{f}$ has a polynomial growth rate (i.e., that (6.1) is satisfied for some $p \in \mathbb{R}_+$). Of course, the most interesting here is the supercritical case when the assumption (7.5) is not satisfied. Unfortunately, we do not know the answer on this question in general and have to pose some extra restrictions which however look natural. Namely, we assume that the nonlinearity also satisfies

$$|\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})| \le C(1 + |\mathbf{f}(\mathbf{u})| + |\mathbf{u}|), \quad \mathbf{u} \in \mathbb{R}^k. \tag{7.10}$$

Then, the following result holds.

Theorem 7.6 *Let the assumptions of Proposition 7.1 hold and let also assumptions (2.3) and (7.10) be satisfied. Then the image $\widehat{S}(1)B_R$ of any closed ball B_R of radius*

R in H is a compact set in $\mathbb{D}$. In particular the global attractor $\mathcal{A}$ is compact in $\mathbb{D}$ and attracts in the strong topology of $\mathbb{D}$ as well.

Proof We only need to prove the compactness of $\hat{S}(1)B_R$, the rest is a corollary of the standard attractor's existence theorem. The fact that this set is closed is also standard and we left it to the reader. So, we will only check pre-compactness below.

The proof of this fact is a combination of parabolic regularity estimates which gives the pre-compactness of the set

$$\mathcal{B}_R := \{\partial_t \mathbf{u}(1), \ \mathbf{u}(t) = \widehat{S}(t)\mathbf{u}_0, \ \mathbf{u}_0 \in B_R\}$$

in $L^2(\Omega)$ and energy type estimates for the elliptic equation which then give the desired compactness in $H^2(\Omega)$.

Step 1. $\mathcal{B}_R$ is compact in $L^2(\Omega)$. We already know that $\mathcal{B}_R$ is a bounded set in $L^{2+r}(\Omega)$, due to Proposition 7.1 and Corollary 6.2. In order to get the desired compactness we will use the standard interpolation embedding:

$$W^{1-\kappa,1}(\Omega) \cap L^{r+2}(\Omega) \subset H^{(1-\kappa)\frac{r}{2(r+1)}}(\Omega), \tag{7.11}$$

see [42]. This embedding together with the compactness of the embedding $H^\varepsilon \subset L^2$ will give the desired result if we prove boundedness of $\mathcal{B}_R$ in $W^{1-\kappa,1}$ for some $0 < \kappa < 1$. To this end, we note that according to Corollary 6.2, $\mathbf{u}(t) \in \mathbb{D}$ for $t \in [1/2, 1]$ and is uniformly bounded there if $\mathbf{u}_0 \in B_R$. Thus, according to assumption (7.10), $\|\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}(t))\|_{L^2}$ is uniformly bounded. Since the L^{2+r}-norm of $\partial_t \mathbf{u}(t)$ is also bounded due to Proposition 7.1, we have the estimate

$$\|\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\partial_t \mathbf{u}\|_{L^s} \le C_R, \ \ t \in [1/2, 1], \ \ \frac{1}{s} = \frac{1}{2} + \frac{1}{r+2}.$$

Thus, applying the maximal L^s-regularity estimate (with $s > 1$, see e.g., [29]) to equation

$$\partial_t \boldsymbol{\theta} - \mathbf{D}\Delta_x \boldsymbol{\theta} = -\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}(t))\partial_t \mathbf{u}(t), \ \ \boldsymbol{\theta} := \partial_t \mathbf{u},$$

using the anisotropic Sobolev embedding

$$W^{1,s}(0,1;L^s) \cap L^s(0,1;W^{2,s}) \subset C(0,1;W^{2(1-\frac{1}{s}),s}),$$

and arguing as at the end of the proof of Lemma 7.4, we arrive at

$$\|\boldsymbol{\theta}(1)\|_{W^{2(1-\frac{1}{s}),s}} \le C(\|\partial_t \boldsymbol{\theta}\|_{L^s(3/4,1;L^s)} + \|\Delta_x \boldsymbol{\theta}\|_{L^s(3/4,1;L^s)}) \le$$
$$\le C(\|\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\partial_t \mathbf{u}\|_{L^s(1/2,1;L^s)} + \|\partial_t \mathbf{u}\|_{L^2(1/2;1;L^2)}) \le C'_R. \tag{7.12}$$

This estimate gives the desired boundedness of $\mathcal{B}_R$ in $W^{1-\kappa,1}(\Omega)$ and completes the first step of the proof.

Step 2. Compactness in H^2. Let us consider a sequence of solutions $\mathbf{u}_n(t)$, $\mathbf{u}_n(0) \in B_R$ and find a subsequence which is convergent strongly in H^2 to some solution $\mathbf{u}(t)$. Due to the result of Step 1, we may assume without loss of generality that $\mathbf{u}_n(1) \to \mathbf{u}(1)$ weakly in H^2 and $\partial_t \mathbf{u}_n(1) \to \partial_t \mathbf{u}(1)$ *strongly in L^2*. In other words, we need to pass to the limit $n \to \infty$ in the semilinear elliptic equation

$$\mathbf{D}\Delta_x \mathbf{u}_n(1) - \mathbf{f}(\mathbf{u}_n(1)) = \mathbf{h}_n := \partial_t \mathbf{u}_n(1) - \mathbf{g}. \tag{7.13}$$

Without loss of generality we may assume also that $\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}) \geq 0$. We will utilize the so-called energy method. Assume at this moment that we are able to integrate by parts and get

$$(\mathbf{f}(\mathbf{u}), \Delta_x \mathbf{u}) = -(\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u})\nabla_x \mathbf{u}, \nabla_x \mathbf{u}), \quad \mathbf{u} \in \mathbb{D}, \tag{7.14}$$

this formula will be verified later at the end of the proof. Then, multiplying (7.13) by $\Delta_x \mathbf{u}_n$ and integrating over x, we get the energy identity

$$(\mathbf{D}\Delta_x \mathbf{u}_n, \Delta_x \mathbf{u}_n) + (\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}_n)\nabla_x \mathbf{u}_n, \nabla_x \mathbf{u}_n) = (\mathbf{h}_n, \Delta_x \mathbf{u}_n). \tag{7.15}$$

Our aim here is to pass to the limit $n \to \infty$ in this equality and compare it with the energy equality for the limit solution. Indeed, using the convexity arguments (similarly to (4.11)), we get

$$\begin{aligned} (\mathbf{D}\Delta_x \mathbf{u}(1), \Delta_x \mathbf{u}(1)) &\leq \liminf_{n\to\infty}(\mathbf{D}\Delta_x \mathbf{u}_n(1), \Delta_x \mathbf{u}_n(1)), \\ (\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u})\nabla_x \mathbf{u}, \nabla_x \mathbf{u}) &\leq \liminf_{n\to\infty}(\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}_n)\nabla_x \mathbf{u}_n, \nabla_x \mathbf{u}_n) \end{aligned} \tag{7.16}$$

and due to the strong convergence $\mathbf{h}_n \to \mathbf{h}$, we have

$$(\mathbf{h}, \Delta_x \mathbf{u}(1)) = \lim_{n\to\infty}(\mathbf{h}_n, \Delta_x \mathbf{u}_n(1)).$$

Then, the comparison with the limit energy identity

$$(\mathbf{D}\Delta_x \mathbf{u}, \Delta_x \mathbf{u}) + (\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u})\nabla_x \mathbf{u}, \nabla_x \mathbf{u}) = (\mathbf{h}, \Delta_x \mathbf{u}(1))$$

shows that we must have

$$\lim_{n\to\infty}(\mathbf{D}\Delta_x \mathbf{u}_n(1), \Delta_x \mathbf{u}_n(1)) = (\mathbf{D}\Delta_x \mathbf{u}(1), \Delta_x \mathbf{u}(1)).$$

Together with the weak convergence $\Delta_x \mathbf{u}_n(1) \to \Delta_x \mathbf{u}(1)$ this gives the strong convergence $\Delta_x \mathbf{u}_n(1) \to \Delta_x \mathbf{u}(1)$ in L^2 and, therefore, the strong convergence $\mathbf{u}_n(1) \to \mathbf{u}(1)$ in H^2. From the equation (7.13) we finally establish that $\mathbf{f}(\mathbf{u}_n(1)) \to \mathbf{f}(\mathbf{u}(1))$ also strongly. Thus, the compactness of $\hat{S}(1)B_R$ in $\mathbb{D}$ is proved. Thus, the theorem is proved by modulo of the integration by parts formula (7.14) which we prove in the following lemma. □

Lemma 7.7 *Let the nonlinearity $\mathbf{f}$ satisfy the assumptions of Theorem 7.6. Then integration by parts (7.14) is valid for every $\mathbf{u} \in \mathbb{D}$.*

Proof of the lemma We first establish the identity

$$(\mathbf{f}(\mathbf{u}), \operatorname{div} W) = -(\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\nabla_x\mathbf{u}, W) \tag{7.17}$$

for all vector fields $W \in C^\infty(\bar{\Omega})$. Due to our assumption (7.10) both parts of this equality make sense. The identity may be proved by approximating the function $\mathbf{f}$ by "good" functions $\mathbf{f}_n$ as in Lemma 4.1. Since $\mathbf{f}$ has a polynomial growth, we may take, say,

$$\Psi(z) = e^{\sqrt{z+1}},$$

and this allows us to keep also assumption (7.10) uniformly in n. Let $\mathbf{u}_n$ be the corresponding approximating functions for $\mathbf{u}$ constructed as in (4.5). Then, we first verify the integration by parts for $\mathbf{f}_n$ and $\mathbf{u}_n$ (which is trivial since everything is smooth) and after that pass to the limit $n \to \infty$ (which is also straightforward since as in Lemma 4.2, we have weak convergence $\mathbf{f}_n(\mathbf{u}_n) \to \mathbf{f}(\mathbf{u})$ in L^2 and, due to our assumption (7.10), we also have weak convergence $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}_n)\nabla_x\mathbf{u}_n$ to $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\nabla_x\mathbf{u}$ in $L^{1+\varepsilon}$ for small positive ε. Thus, the integration by parts (7.17) is verified for smooth vector fields W.

Note that the C^∞ smoothness assumption on the vector field W can be relaxed till

$$W \in H^1(\Omega) \cap L^q(\Omega)$$

(where q is large enough that $\frac{1}{q} + \frac{1}{1+\varepsilon} \leq 1$) by density arguments.

We now construct a sequence of Lipschitz continuous cut-off functions

$$\varphi_n(z) = \begin{cases} 1, & z \leq n, \\ 1 - \ln\frac{z}{n}, & z \in [n, en], \\ 0, & z > ne. \end{cases} \tag{7.18}$$

Then, the sequence $\varphi_n(z)$ is monotone increasing in n and is convergent point-wise to one. Moreover, the following estimate holds:

$$|\varphi'(z)z| \leq 1, \quad z \in \mathbb{R} \tag{7.19}$$

(there are no problems to construct similar *smooth* sequence, but we prefer to give relatively simple explicit expression). Then, we define a vector field $W = W_n$ as follows:

$$W_n(x) := \varphi_n(|\nabla_x\mathbf{u}|^2)\nabla_x\mathbf{u}.$$

Then, as simple calculation shows, $W \in L^\infty(\Omega)$ and, due to condition (7.19),

$$\|\nabla_x W_n\|_{L^2} \leq C\|\nabla_x^2\mathbf{u}\|,$$

where the constant C is independent of n. Thus, we may conclude that div $W_n \to \Delta_x \mathbf{u}$ weakly in $L^2(\Omega)$. Moreover, we may put W_n to the integration by parts formula (7.17) and get

$$(\mathbf{f}(\mathbf{u}), \text{div } W_n) = -(\varphi_n(|\nabla_x \mathbf{u}|^2)\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\nabla_x \mathbf{u}, \nabla_x \mathbf{u}).$$

It only remains to pass to the limit $n \to \infty$ here. Passing to the limit in the left-hand side is immediate and to pass to the limit $n \to \infty$ in the right-hand side, it is enough to note that $(\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}) + K)\nabla_x \mathbf{u}.\nabla_x \mathbf{u}$ is non-negative and belongs to $L^1(\Omega)$. The monotonicity of φ_n in n and its point-wise convergence to one allow us to apply the Levy monotone convergence theorem and get the desired result. Thus, the lemma is proved and the theorem is also proved. $\qquad\square$

Remark 7.8 We expect that the integration by parts formula (7.14) holds without the extra assumption (7.10), however, it is not clear how to verify it. Key difficulty here is that $\mathbb{D}$ is a nonlinear set and it is not easy to construct good smooth approximations for functions $\mathbf{u} \in \mathbb{D}$.

The first step in the proof of Theorem 7.6 can be also done using the energy type arguments. To this end one just needs to verify the energy identity

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\theta}(t)\|_{L^2}^2 + (\mathbf{D}\nabla_x\boldsymbol{\theta}(t), \nabla_x\boldsymbol{\theta}(t)) + (\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}(t))\boldsymbol{\theta}(t), \boldsymbol{\theta}(t)) = 0 \qquad (7.20)$$

which can be verified similarly to the proof of Lemma 7.7. In the case of reaction diffusion system (2.1) this is not necessary since Proposition 7.1 gives a simpler way to verify the compactness. However, it may be useful in the case of higher order equations where the technique of Proposition 7.1 may not work.

8 Finite dimensionality and exponential attractors

In this section we discuss the finite-dimensionality of the global attractor for problem (2.1) and the existence of the so-called exponential attractor. We recall that a set $\mathcal{M} \subset H$ is called an exponential attractor of the semigroup $\widehat{S}(t) : H \to H$ if the following conditions are satisfied:

1. $\mathcal{M}$ is compact in H;
2. $\mathcal{M}$ is semi-invariant: $\widehat{S}(t)\mathcal{M} \subset \mathcal{M}$;
3. It has a finite fractal dimension in H: $\dim_F(\mathcal{A}, H) < \infty$;
4. It attracts the images of bounded in H sets exponentially as time tends to infinity, i.e., for every bounded set B,

$$\text{dist}(\hat{S}(t)B, \mathcal{M}) \le Q(\|B\|_H)e^{-\alpha t}$$

for some positive α and monotone function Q which are independent of B.

It is well-known that the exponential attractor if exists always contains a global attractor, so the existence of $\mathcal{M}$ automatically implies the finite-dimensionality of a

global attractor. In contrast to global attractors, exponential attractors usually more robust with respect to perturbations and allow us to control the rate of attraction in terms of physical parameters of the system considered, but as a price to pay for that, an exponential attractor is not unique, see [16,18,35] for more details.

The existence of an exponential attractor is usually verified using the following abstract result for discrete semigroups $\hat{S}(n) := \hat{S}^n : H \to H$ generated by the map $\hat{S} : H \to H$.

Proposition 8.1 *Let H, V be two Banach spaces such that V is compactly embedded in H. Assume that there exists a bounded closed set $B \subset H$ and a map $\hat{S} : B \to B$ such that*

$$\|\hat{S}(\xi_1) - \hat{S}(\xi_2)\|_V \leq K\|\xi_1 - \xi_2\|_H, \ \ \xi_1, \xi_2 \in B. \tag{8.1}$$

Then the corresponding discrete semigroup $\hat{S}(n) : B \to B$ possesses an exponential attractor $\mathcal{M} \subset B$.

For the proof of this proposition, see [17,18].

In applications usually B is an absorbing ball of the considered continuous semigroup $\hat{S}(t) : H \to H$, $\hat{S} := \widehat{S}(T)$ for some properly chosen T and (8.1) is verified using the proper parabolic smoothing property for the equation on differences of two solutions. If the existence of a discrete exponential attractor $\mathcal{M}_d$ is established, the exponential attractor for the continuous semigroup can be constructed by the standard formula:

$$\mathcal{M} := \cup_{t \in [T,2T]} \hat{S}(t)\mathcal{M}_d$$

and in order to get its finite-dimensionality, we need to assume in addition that the semigroup is also Hölder continuous in time:

$$\|\hat{S}(t_1)\xi_1 - \hat{S}(t_2)\xi_2\|_H \leq L\left(\|\xi_1 - \xi_2\|_H + |t_1 - t_2|^\alpha\right), \tag{8.2}$$

for some $\alpha \in (0, 1]$ and all $t_i \in [T, 2T]$ and $\xi_i \in B$, see [18] for the details.

The main result of this section is the following theorem.

Theorem 8.2 *Let the nonlinearity $\mathbf{f}$ satisfy assumptions (2.3), (6.1) for some $p \in \mathbb{R}_+$, (7.10) and the following convexity property: there exist a convex function $\Psi : \mathbb{R}^k \to \mathbb{R}_+$ such that*

$$C_2(\Psi(\mathbf{u}) - 1 - |\mathbf{u}|^2) \leq |\mathbf{f}(\mathbf{u})|^2 \leq C_1(\Psi(\mathbf{u}) + |\mathbf{u}|^2 + 1), \ \ \mathbf{u} \in \mathbb{R}^k, \tag{8.3}$$

for some positive constants C_1 and C_2. Let also $\mathbf{g} \in L^2(\Omega)$ and $\mathbf{D}$ satisfy (2.2). Then problem (2.1) possesses an exponential attractor $\mathcal{M}$ in the space $H := L^2(\Omega)$ which is a compact set in $\mathbb{D}$.

Proof According to Corollary 6.2 a ball $B = B_R$ in $\mathbb{D}$ of a sufficiently large radius R is an absorbing set for the solution semigroup $\hat{S}(t) : H \to H$ associated with equation (2.1). Let us fix $T > 0$ big enough that $\hat{S}(T)B \subset B$ and set $\hat{S} := \hat{S}(T)$. Then, according to estimate (3.8), the semigroup $\hat{S}(t)$ is Lipschitz continuous with respect

to the initial data for every fixed t. Moreover, since $\partial_t \mathbf{u}(t)$ is bounded in the L^2-norm if $\mathbf{u}(0) \in \mathbb{D}$, this semigroup is also Lipschitz continuous in time, so condition (8.2) is satisfied with $\alpha = 1$. Therefore, in order to verify the existence of an exponential attractor, it is enough to check the smoothing property (8.1) for the properly chosen space V. To this end, we need to establish a number of smoothing estimates for the difference of solutions of equation (2.1).

Let $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ be two solutions of (2.1) starting from the absorbing ball B. Then their difference $\boldsymbol{\theta}(t) := \mathbf{u}_1(t) - \mathbf{u}_2(t)$ solves the equation

$$\partial_t \boldsymbol{\theta} = \mathbf{D}\Delta_x \boldsymbol{\theta} - \mathbf{L}(t)\boldsymbol{\theta}, \ \ \mathbf{L}(t) := \int_0^1 \nabla_{\mathbf{u}} \mathbf{f}(s\mathbf{u}_1(t) + (1-s)\mathbf{u}_2(t)) \, ds. \tag{8.4}$$

We recall that, due to the assumption $\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}) \geq -K$, multiplication of this equation on $\boldsymbol{\theta}$ gives the basic Lipschitz continuity estimate:

$$\|\boldsymbol{\theta}(T)\|_{L^2}^2 + \int_0^T \|\nabla_x \boldsymbol{\theta}(t)\|_{L^2}^2 \, dt \leq Ce^{KT}\|\boldsymbol{\theta}(0)\|_{L^2}^2, \tag{8.5}$$

see Lemma 3.6. Moreover, multiplying (8.4) by $\boldsymbol{\theta}|\boldsymbol{\theta}|^r$ and arguing exactly as in the proof of Proposition 7.1, we get the estimate

$$\|\boldsymbol{\theta}(T)\|_{L^{r+2}} \leq Ce^{KT}T^{-1}\|\boldsymbol{\theta}(0)\|_{L^2} \tag{8.6}$$

for some sufficiently small positive r depending only on the matrix $\mathbf{D}$.

In order to get smoothing estimate for $\boldsymbol{\theta}$, we argue as in Step 1 of the proof of Theorem 7.6. Namely, from (7.10) and (8.3), we conclude that

$$\begin{aligned}
|\nabla_{\mathbf{u}} \mathbf{f}(s\mathbf{u}_1 + (1-s)\mathbf{u}_2)|^2 &\leq \\
&\leq C(|\mathbf{f}(s\mathbf{u}_1 + (1-s)\mathbf{u}_2)|^2 + 1 + |\mathbf{u}_1|^2 + |\mathbf{u}_2|^2) \leq \\
&\leq C'(\Psi(s\mathbf{u}_1 + (1-s)\mathbf{u}_2) + 1 + |\mathbf{u}_1|^2 + |\mathbf{u}_2|^2) \leq \\
&\leq C'(\Psi(\mathbf{u}_1) + \Psi(\mathbf{u}_2) + 1 + |\mathbf{u}_1|^2 + |\mathbf{u}_2|^2) \leq \\
&\leq C''(|\mathbf{f}(\mathbf{u}_1)|^2 + |\mathbf{f}(\mathbf{u}_2)|^2 + |\mathbf{u}_1|^2 + |\mathbf{u}_2|^2 + 1), \ s \in (0,1) \tag{8.7}
\end{aligned}$$

and, therefore, since $\mathbf{f}(\mathbf{u}(t))$ is uniformly bounded in L^2-norm for our solutions $\mathbf{u}_1$ and $\mathbf{u}_2$, we have

$$\|\mathbf{L}(t)\|_{L^2} \leq C, \ \ t \in [0, T]. \tag{8.8}$$

This estimate, in turn, implies (together with (8.5), Sobolev embedding theorem and Hölder inequality) that

$$\|\mathbf{L}(t)\boldsymbol{\theta}\|_{L^s(0,T;L^s)} \leq C_T$$

for some $1 < s < 2$. Applying now the L^s interior regularity estimate to equation (8.4) and arguing as in the proof of Theorem 7.6, we get

$$\|\boldsymbol{\theta}(T)\|_{W^{2(1-\frac{1}{s}),s}} \leq C_T\|\boldsymbol{\theta}(0)\|_{L^2}$$

which together with the embedding (7.11) gives

$$\|\boldsymbol{\theta}(T)\|_{H^\varepsilon} \leq C_T \|\boldsymbol{\theta}(0)\|_{L^2}$$

for some positive exponent ε. Setting finally $V = H^\varepsilon(\Omega)$ we get the desired smoothing property (8.1) and finish the proof of the theorem. □

Remark 8.3 The finite-dimensionality of the global attractor $\mathcal{A}$ has been established under similar assumptions on **f** in [46] using the so-called method of l-trajectories, see also [32,33]. In the present work we suggest the simplified version of the proof which is based on multiplication of equation (8.4) on the quantities like $\boldsymbol{\theta}|\boldsymbol{\theta}|^r$. Although the proof becomes more transparent, it is slightly less general than the one suggested in [46] since this multiplication is suitable for reaction-diffusion systems and may not work for more general ones (e.g., higher order equations). In such cases one should return back to the method of l-trajectories.

9 Generalizations and concluding remarks

In this concluding section we briefly consider other types of equations for which the technique developed above works (with some minor changes which we will discuss) and state some interesting open problems. We start with the case of fractional Laplacians and the corresponding reaction-diffusion equations which are becoming more and more popular nowadays, see [1,2,23,31] and references therein for more details.

9.1 Fractional reaction-diffusion systems

Let us define $A := (-\Delta_x)^\alpha$, $0 < \alpha < 1$, in the domain Ω endowed with Dirichlet boundary conditions. In other words, we define A as a fractional power of the Laplacian $-\Delta_x$ in the domain Ω endowed with Dirichlet boundary conditions:

$$Au := \frac{1}{\Gamma(-\alpha)} \int_0^\infty (e^{t\Delta_x}u - u)\frac{dt}{t^{1+\alpha}},$$

where $\Gamma(z)$ is the standard Euler Γ-function, see [7,42] for more details, although we believe that similar results can be obtained for other types of fractional Laplacians.

Let us consider the following fractional reaction-diffusion system:

$$\partial_t \mathbf{u} + \mathbf{D}(-\Delta_x)^\alpha \mathbf{u} + \mathbf{f}(\mathbf{u}) = \mathbf{g}, \quad \mathbf{u}\big|_{\partial\Omega} = 0, \tag{9.1}$$

where the function **f** and the matrix **D** satisfy assumptions (2.3) and (2.2) respectively. In this case, the definition of the phase space $\mathbb{D}$ should be transformed as follows:

$$\mathbb{D}_\alpha := \{\mathbf{u} \in H_{\Delta_x}^{2\alpha}, \ \mathbf{f}(\mathbf{u}) \in L^2(\Omega)\}, \tag{9.2}$$

where $H_{\Delta_x}^{2\alpha} := D((-\Delta_x)^\alpha)$.

All of the estimates and results stated above for the case $\alpha = 1$ can be extended in a straightforward way to a general case $0 < \alpha < 1$. The only non-trivial place is the estimates of the terms like $((-\Delta_x)^\alpha \mathbf{u}, \mathbf{f}(\mathbf{u}))$ or $((-\Delta_x)^\alpha \mathbf{u}, \mathbf{u}|\mathbf{u}|^r)$. In the case when $\Omega = \mathbb{R}^d$ (or in the case of periodic BC), we have a nice explicit formula for such inner products which trivializes the required estimates (see e.g., [42]), namely,

$$(A\mathbf{u}, \mathbf{v}) = C_\alpha \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{(\mathbf{u}(x) - \mathbf{u}(y)).(\mathbf{v}(x) - \mathbf{v}(y))}{|x - y|^{d+2\alpha}} \, dx \, dy. \tag{9.3}$$

In particular, it gives the positivity of $(A\mathbf{u}, \mathbf{f}(\mathbf{u}))$ if $\nabla_\mathbf{u} \mathbf{f}(\mathbf{u}) \geq 0$. Fortunately, there is an extension of this formula to the case of bounded domains (see [7]), namely,

$$(A\mathbf{u}, \mathbf{v}) = \int_\Omega \int_\Omega (\mathbf{u}(x) - \mathbf{u}(y)).(\mathbf{v}(x) - \mathbf{v}(y)) K_{\Omega,\alpha}(x, y) \, dx \, dy +$$
$$+ \int_\Omega \mathbf{u}(x).\mathbf{v}(x) B_{\Omega,\alpha}(x) \, dx \tag{9.4}$$

for some non-negative functions $K_{\Omega,\alpha}$ and $B_{\Omega,\alpha}$. This formula allows us to get the same type of estimates as for the local case $\alpha = 1$.

Namely, analogously to Theorem 4.3, we have the existence and uniqueness of strong solutions $\mathbf{u} \in C_w(0, T; \mathbb{D}_\alpha)$ for any $\mathbf{u_0} \in \mathbb{D}_\alpha$ and $g \in L^2(\Omega)$ (we always assume in this section that assumptions (2.2) and (2.3) are satisfied), so the solution semigroup

$$S(t) : \mathbb{D}_\alpha \to \mathbb{D}_\alpha$$

is well-defined and globally Lipschitz continuous in $L^2(\Omega)$. Moreover, exactly as in §5, this semigroup can be extended by continuity to the semigroup $\widehat{S}(t)$ acting on whole phase space $L^2(\Omega)$. The corresponding trajectory $\mathbf{u}(t) := \widehat{S}(t)\mathbf{u_0}$ belongs to $C(0, T; L^2(\Omega))$ and can be interpreted as a unique weak solution of problem (9.1) in a sense of Definition 5.1 where $\mathbb{D}$ is replaced by $\mathbb{D}_\alpha$ and Δ_x by $-(-\Delta_x)^\alpha$ respectively. Furthermore, exactly as in §5, this semigroup possesses a global attractor $\mathcal{A}$ in $L^2(\Omega)$.

For the convenience of the reader, we state below the analogues of two key results for the fractional case which can be obtained analogously to the case $\alpha = 1$. The comprehensive study of this case out of the scope of this paper, so we leave the details to the reader. We return to this problem somewhere else.

Theorem 9.1 *Let the matrix* $\mathbf{D}$ *and the nonlinearity* $\mathbf{f}$ *satisfy* (2.2) *and* (2.3) *respectively and let, in addition, the nonlinearity* $\mathbf{f}$ *satisfy* (6.1) *with the exponent* p *restricted by the assumption*

$$p < p_{crit}(\alpha), \quad p_{crit}(\alpha) := 1 + \frac{4\alpha}{d - 4\alpha} \tag{9.5}$$

if $d \geq 4\alpha$ *and the external forces* $\mathbf{g} \in L^q(\Omega)$ *for some* $q > \frac{d}{2\alpha}$. *Then any weak solution* $\mathbf{u}(t)$ *of problem* (9.1) *starting from* $\mathbf{u}(0) \in H$ *possesses the following smoothing property:*

$$\|\mathbf{u}(t)\|_{L^\infty} \leq Q_t(\|\mathbf{u}(0)\|_{L^2} + \|\mathbf{g}\|_{L^q}), \quad t \in (0, 1] \tag{9.6}$$

for some monotone function Q_t depending on t, but independent of $\mathbf{u}$ and $\mathbf{g}$.

Remark 9.2 Note that this result is not very helpful if $0 < \alpha < \frac{1}{2}$ since the direct H^1-estimate which is obtained by multiplication of the equation by $-\Delta_x \mathbf{u}$ gives the control of the H^1-norm (of course, assuming in addition that $\mathbf{g} \in H^{1-\alpha}$) which is better than $H^{2\alpha}$-control finally obtained from $\mathbf{u} \in \mathbb{D}_\alpha$. However, it is useful for $\alpha \geq \frac{1}{2}$. In particular, in the case $0 < \alpha < \frac{3}{4}$, we may have the supercritical growth rate in the case of physical dimension $d = 3$ as well. So, main results become applicable for $d = 3$ as well. Note also that many of the results of our paper may be extended also to the case $\alpha > 1$ (e.g., to the Swift-Hohenberg type equations where $\alpha = 2$), but in this case we will not be able to multiply the equation by $A\mathbf{u}$ since the term $(A\mathbf{u}, \mathbf{f}(\mathbf{u}))$ will be out of control, so we may multiply it only on $\Delta_x \mathbf{u}$ and this gives the control of the $H^{\frac{1+\alpha}{2}}$-norm of $\mathbf{u}(t)$ (not $H^{2\alpha}$ as before).

We now state the result about exponential attractors for the supercritical case.

Theorem 9.3 *Let the nonlinearity $\mathbf{f}$ satisfy assumptions (2.3), (6.1) for some $p \in \mathbb{R}_+$, (7.10) and the following convexity property: there exist a convex function $\Psi : \mathbb{R}^k \to \mathbb{R}$ such that*

$$C_2(\Psi(\mathbf{u}) - 1 - |\mathbf{u}|^2) \leq |f(\mathbf{u})|^2 \leq C_1(\Psi(\mathbf{u}) + |\mathbf{u}|^2 + 1), \quad \mathbf{u} \in \mathbb{R}^k, \qquad (9.7)$$

for some positive constants C_1 and C_2. Let also $0 < \alpha < 1$, $\mathbf{g} \in L^2(\Omega)$ and $\mathbf{D}$ satisfy (2.2). Then problem (9.1) possesses an exponential attractor $\mathcal{M}$ in the space $H := L^2(\Omega)$ which is a boumded set in $\mathbb{D}_\alpha$.

Remark 9.4 We expect that we may add sufficiently small $\varepsilon > 0$ to the critical exponent p_{crit} in (9.5) (exactly as in the case of classical diffusion, see (7.5)) as well as to prove that the exponential attractor $\mathcal{M}$ is not only bounded, but also compact in $\mathbb{D}_\alpha$. However, to get these results, we need to verify the analogue of Theorem A.1 for the fractional Laplacian and this requires extra efforts in comparison with the local case $\alpha = 1$, so we prefer not to state these results here.

9.2 Cahn-Hilliard type systems

Let us consider the following fractional Cahn-Hilliard system in $\Omega \subset \mathbb{R}^d$:

$$\partial_t \mathbf{u} + (-\Delta_x)^\beta (\mathbf{D}(-\Delta_x)^\alpha \mathbf{u} + \mathbf{f}(\mathbf{u}) - \mathbf{g}) = 0 \qquad (9.8)$$

endowed by the Dirichlet boundary conditions. We assume here that $0 < \beta \leq 1$, $0 < \alpha \leq 1$. Note that $\alpha = \beta = 1$ corresponds to the classical Cahn-Hilliard system and $\beta = 0, \alpha = 1$ to the reaction-diffusion system considered above. See [1,35,41] and references therein for more details concerning classical and fractional CH-equations. It is natural to take $\mathbb{D}_\alpha$ as the phase space for this problem and rewrite it in the following form:

$$\partial_t(-\Delta_x)^{-\beta} \mathbf{u} + \mathbf{D}(-\Delta_x)^\alpha \mathbf{u} + \mathbf{f}(\mathbf{u}) = \mathbf{g}. \qquad (9.9)$$

Then we may utilize the monotonicity of the function $\mathbf{f}$ and apply the developed above theory to this equation (see also [35] for the case $\alpha = \beta = 1$). In this case, weak solutions are naturally defined in the space $H := H^{-\beta}(\Omega)$ and strong solutions live in $\mathbb{D}_\alpha$.

The key result on the existence of exponential attractors now reads.

Theorem 9.5 *Let the nonlinearity* $\mathbf{f}$ *satisfy assumptions* (2.3), (6.1) *for some* $p \in \mathbb{R}_+$, (7.10) *and the following convexity property: there exist a convex function* $\Psi : \mathbb{R}^k \to \mathbb{R}$ *such that*

$$C_2(\Psi(\mathbf{u}) - 1 - |\mathbf{u}|^2) \le |\mathbf{f}(\mathbf{u})|^2 \le C_1(\Psi(\mathbf{u}) + |\mathbf{u}|^2 + 1), \quad \mathbf{u} \in \mathbb{R}^k, \qquad (9.10)$$

for some positive constants C_1 *and* C_2. *Let also* $\mathbf{g} \in L^2(\Omega)$ *and* $\mathbf{D}$ *satisfy* (2.2). *Then problem* (9.8) *possesses an exponential attractor* $\mathcal{M}$ *in the space* $H := H^{-\beta}(\Omega)$ *which is a bounded set in* $\mathbb{D}_\alpha$.

We leave the rigorous proof of this theorem to the reader.

9.3 Open problems

We conclude this section by a discussion of some open questions and possible further improvements of the above developed theory.

Problem 1. We start with the already posed question about the validity of the integration by parts formula

$$(\mathbf{f}(\mathbf{u}), \Delta_x \mathbf{u}) = -(\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\nabla_x \mathbf{u}, \nabla_x \mathbf{u}) \qquad (9.11)$$

for every $\mathbf{u} \in \mathbb{D}$. We know that both parts of this equality are well-defined for any $\mathbf{u} \in \mathbb{D}$. However, since we do not know the density of smooth functions in $\mathbb{D}$, we cannot verify the identity in a standard way, so we need to use something else. We have proved this identity under the extra assumption (7.10) which allows us to control the Lebesgue norm of $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\nabla_x \mathbf{u}$ and simplifies the situation. Clarifying the situation with this integration by parts in general would be very useful for establishing energy equalities for many other equations containing monotone nonlinearities which, in turn, may give compactness of the corresponding global attractors. We were sure that (7.10) is technical, but surprisingly are unable to remove it (or find the proper reference).

Problem 2. Next problem is related with smoothness of weak/strong solutions of problem (2.1). We have established that under the assumption (6.1) that $\mathbf{f}$ has a polynomial growth rate, the problem possesses H to $\mathbb{D}$ smoothing property. It would be interesting to understand whether or not this polynomial growth restriction is really necessary for the smoothing (ideally, to construct a non-smoothing weak solution for problem (2.1)), say, with exponential or stronger nonlinearities. A natural idea here is to extend the proof of Theorem 6.1 to the case where $\partial_t \mathbf{u}$ belongs to some weaker spaces than $L^p(\Omega)$ with $p \ll 1$ using the technique of Orlich spaces. But more detailed analysis shows that this does not work already when $\ln(1 + |\mathbf{f}(\mathbf{u})|) \in L^1$, so we may

expect the existence of such exotic non-smoothing weak solutions for fast growing nonlinearities.

The phenomenon of delayed regularization is well-known in the class of nonlinear diffusion problems, see [44] and reference therein. For example, the equation

$$\partial_t u |\partial_t u|^p = \Delta_x u, \ u\big|_{\partial\Omega} = 0, \ \ p \geq 0$$

is well-posed in a natural energy phase space $\Phi = W_0^{1,2}(\Omega)$. However, the solutions of this equation *do not possess* the standard parabolic smoothing property if, say, $p > 4$ and $d = 3$. Indeed, the energy identity for this equation reads

$$\|\nabla_x u(T)\|_{L^2}^2 + \int_0^T \|\partial_t u(t)\|_{L^{p+2}}^{p+2} \, dt = \|\nabla_x u(0)\|_{L^2}^2,$$

so if $u(0) \notin L^{p+2}(\Omega)$, we have $u(T) \notin L^{p+2}(\Omega)$ for any finite $T > 0$. However, if we start from more regular phase space $\Psi := W_0^{1,2}(\Omega) \cap L^\infty(\Omega)$, we will have instantaneous further regularization, see [20]. The open question is whether or not something similar happens in the case of system (1.1) of reaction-diffusion equations with fast growing nonlinearity $\mathbf{f}$ satisfying (2.3).

Another related question is about generating singularities in finite time in equations like (2.1). It is known that general reaction-diffusion systems may generate singularities in higher norms even if the natural energy norm remains finite and dissipative, see e.g., [36] for RDS satisfying balance law (=mass conservation law), [25] for the case of reaction-diffusion with chemotaxis or [6] for Ginzburg-Landau equations in $\mathbb{R}^3$ (see also references therein). However, to the best of our knowledge, there are no such examples in the class of equation (2.1) with nonlinearities satisfying $\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}) \geq -K$. As we know, in this case the H^2-norm cannot blow up, so this is the question of possible blow up of higher norms and high space dimension $d > 4$.

Problem 3. Finally, about the finite-dimensionality of global attractors. The most popular scheme for proving this result is related with volume contraction technique, see [3,41] and references therein. Using this technique, we need to estimate l dimensional traces $\text{Tr}_l \, \mathcal{L}_u$, where

$$\mathcal{L}_{\mathbf{u}} \mathbf{v} = \mathbf{D} \Delta_x \mathbf{v} - \nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}) \mathbf{v}$$

is the linearized operator on the trajectory $\mathbf{u}(t)$ of the equation (2.1) lying on the attractor. *Formal* estimates of this quantity depend only on K (if the assumption $\nabla_{\mathbf{u}} \mathbf{f}(\mathbf{u}) \geq -K$ is posed) and are independent of the norm of $\mathbf{u}(t)$ and any norms of $\mathbf{f}(\mathbf{u})$.

However, to justify this method we need to verify the *differentiability* of the semigroup $\hat{S}(T)$ with respect to the initial data (at least the so-called uniform quasi-differentiability on the attractor, see [41]) and such a differentiability usually *does not* hold in supercritical cases.

This was the main reason to use the alternative scheme based on Proposition 8.1 for verifying the finite-dimensionality. In this scheme the differentiability is not required,

but as the price to pay, we get essentially worse estimates than expected since now the norm of $|\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})|$ is involved into all dimension estimates.

It would be interesting to remove this drawback and remove the dependence on $|\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})|$ from these estimates, e.g., by finding a "clever" choice of spaces H and V in Proposition 8.1. Up to the moment we know how to do this in a scalar case only, due to the possibility to multiply (8.4) by sgn v and using the Kato inequality. This in turn gives the estimate of the L^1-norm of $L(t)\theta$ through quantities depending only on K. To the best of our knowledge, nothing similar is known for the vector case.

Declarations

Appendix A. Nonlinear localization and elliptic regularity

In this appendix we consider the following semi-linear elliptic problem:

$$\mathbf{D}\triangle_x\mathbf{u} - \mathbf{f}(\mathbf{u}) = \mathbf{g}, \tag{A.1}$$

where the matrix $\mathbf{D}$ satisfies assumption (2.2) and $\mathbf{f}$ enjoys assumptions (2.3). Then, arguing as before, we get the H^2-elliptic regularity

$$\|\mathbf{u}\|_{H^2} + \|\mathbf{f}(\mathbf{u})\|_{L^2} \le C\left(1 + \|\mathbf{g}\|_{L^2}\right). \tag{A.2}$$

The question addressed here concerns an additional regularity under the extra assumption

$$\mathbf{g} \in L^q(\Omega), \quad q > 2. \tag{A.3}$$

A partial answer on this question is given in the following theorem.

Theorem A.1 *Let the above assumptions hold. Then there exists $\kappa = \kappa(\mathbf{D}) > 0$ such that, for any $d > 2$ and $2 < q < d - \frac{d(d-2)}{\kappa+d}$, the following estimate holds:*

$$\||D_x^2\mathbf{u}| \cdot |\nabla_x\mathbf{u}|^{r/2}\|_{L^2}^2 + \|\nabla_x\mathbf{u}\|_{L^{\frac{d(r+2)}{d-2}}}^{r+2} \le C\left(1 + \|\mathbf{g}\|_{L^q}^{r+2}\right), \tag{A.4}$$

where $r := \frac{d(q-2)}{d-q} \in (0, \kappa)$ *and* $D_x^2\mathbf{u}$ *stands for the collection of all second derivatives of the function* $\mathbf{u}$.

Proof We give below only the formal derivation of estimate (A.4) which can be justified in a standard way (e.g., by cutting off the nonlinearity as explained in Section 4, mollifying $\mathbf{g}$ and using the corresponding smooth solutions of the cut off equation to approximate the initial solution $\mathbf{u}$).

Step 1. We start with the simplest case of periodic boundary conditions where no difficult terms related with the boundary arise and we may integrate by parts freely. Without loss of generality, we may also assume that $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}) \geq 0$ and $\mathbf{f}(0) = 0$.

Let us multiply equation (A.1) by $\partial_{x_i}(\partial_{x_i}\mathbf{u}|\nabla_x\mathbf{u}|^r)$ and integrate over x. This gives

$$(\mathbf{D}\Delta_x\mathbf{u}, \partial_{x_i}^2\mathbf{u}|\nabla_x\mathbf{u}|^r) + (\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\partial_{x_i}\mathbf{u}, \partial_{x_i}\mathbf{u}|\nabla_x\mathbf{u}|^r) \leq$$
$$\leq (\mathbf{g}, \partial_{x_i}(\partial_{x_i}\mathbf{u}|\nabla_x\mathbf{u}|^r)) + Cr(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r). \tag{A.5}$$

Integrating by parts twice in the first term in the left-hand side and using the positivity of $\mathbf{D}$ and $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})$, we arrive at

$$\alpha(|\nabla_x(\partial_{x_i}\mathbf{u})|^2, |\nabla_x\mathbf{u}|^r) \leq (\mathbf{g}, \partial_{x_i}(\partial_{x_i}\mathbf{u}|\nabla_x\mathbf{u}|^r)) + Cr(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r). \tag{A.6}$$

Due to Sobolev embedding theorem, we have

$$\|\nabla_x\mathbf{u}\|_{L^{\frac{d(r+2)}{d-2}}}^{r+2} \leq C\|\nabla_x(|\nabla_x\mathbf{u}|^{\frac{r+2}{2}})\|_{L^2}^2 \leq C'(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r). \tag{A.7}$$

Taking now a sum in (A.6) with respect to $i = 1, \cdots, d$ and assuming that $r \leq \kappa = \kappa(\alpha)$ is small enough, we end up with

$$\alpha\left((|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r) + \|\nabla_x\mathbf{u}\|_{L^{\frac{d(r+2)}{d-2}}}^{r+2}\right) \leq$$
$$\leq C|(\mathbf{g}, |D_x^2\mathbf{u}||\nabla_x\mathbf{u}|^{r/2}|\nabla_x\mathbf{u}|^{r/2})|. \tag{A.8}$$

Finally, using the Hölder inequality with the exponents q, 2 and $\frac{2d(r+2)}{r(d-2)}$, we get

$$|(\mathbf{g}, |D_x^2\mathbf{u}|^2|\nabla_x\mathbf{u}|^{r/2}|\nabla_x\mathbf{u}|^{r/2})| \leq$$
$$\leq \frac{\alpha}{2}\left((|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r) + \|\nabla_x\mathbf{u}\|_{L^{\frac{d(r+2)}{d-2}}}^{r+2}\right) + C_\alpha\|\mathbf{g}\|_{L^q}^{r+2} \tag{A.9}$$

and finish the proof of the theorem in the case of periodic BC.

The case of Dirichlet BC is more delicate since hardly controllable boundary terms will appear if we try to integrate by parts in the first term of (A.5). To avoid them, we will use the nonlinear localization technique suggested in [27] (see also [28]) for more details.

Step 2. Interior estimates. We introduce a cut-off function $\theta \in C^2(\mathbb{R}^d)$ which vanishes near the boundary and equals to one in the δ-interior of the domain Ω ($\delta \ll 1$

is small positive) and satisfies

$$|\nabla_x \theta(x)| \leq C\theta(x)^{1/2},$$

such a function exists and at least C^1-smooth since the domain is smooth. Then, multiplying equation (A.1) by $\nabla_x \cdot (\theta \nabla_x \mathbf{u} |\nabla_x \mathbf{u}|^r)$ and arguing as in Step 1, we arrive at

$$(\theta |D_x^2 \mathbf{u}|^2, |\nabla_x \mathbf{u}|^r) \leq C_\delta \left(\|\mathbf{g}\|_{L^q}^{r+2} + \|\nabla_x \mathbf{u}\|_{L^{r+2}}^{r+2} \right). \tag{A.10}$$

Indeed, integration by parts is now allowed due to the factor θ vanishing near the boundary. Of course this will produce the extra terms containing $\nabla_x \theta$ at every step, but these terms are all under the control due to obvious estimate:

$$\begin{aligned}
&|(|D_x^2 \mathbf{u}| \, |\nabla_x \mathbf{u}|, |\nabla_x \theta| \, |\nabla_x \mathbf{u}|^r)| \\
&\leq \varepsilon(\theta |D_x^2 \mathbf{u}|^2, |\nabla_x \mathbf{u}|^r) + \\
&\quad + C_\varepsilon(|\nabla_x \theta|^2 \theta^{-1}, |\nabla_x \mathbf{u}|^{r+2}) \\
&\leq \varepsilon(\theta |D_x^2 \mathbf{u}|^2, |\nabla_x \mathbf{u}|^r) + C_\varepsilon \|\nabla_x \mathbf{u}\|_{L^{r+2}}^{r+2},
\end{aligned} \tag{A.11}$$

where $\varepsilon > 0$ is arbitrary.

Step 3. Boundary estimates: tangential derivatives. To treat the neighbourhood of the boundary, we introduce an x-dependent smooth orthonormal base $(\boldsymbol{\tau}_1(x), \cdots, \boldsymbol{\tau}_{d-1}(x), \mathbf{n}(x))$ near the boundary such that, when $x \in \partial\Omega$, $\boldsymbol{\tau}_i$ correspond to *tangential* directions and $\mathbf{n}(x)$ is outer normal. We also assume that these vector fields are cut-off outside a small neighbourhood of the boundary similarly to Step 2. Note also that in general such smooth vector fields exist near the boundary only locally, but we will ignore this fact assuming that they exist globally (one more localization is necessary in general). We also define the corresponding directional derivatives

$$\partial_{\boldsymbol{\tau}_i} := \boldsymbol{\tau}_i.\nabla_x, \quad \partial_{\mathbf{n}} := \mathbf{n}.\nabla_x.$$

Then, as follows from the orthogonality conditions,

$$|\nabla_x \mathbf{u}| = |\nabla_{(\boldsymbol{\tau},\mathbf{n})} \mathbf{u}|, \quad [\partial_{\boldsymbol{\tau}_i}, \partial_{\boldsymbol{\tau}_j}] = L.O.T., \quad \partial_{\boldsymbol{\tau}_i}^* = -\partial_{\boldsymbol{\tau}_i} + L.O.T.$$

where "L.O.T." means "lower order terms" and also we have analogous commutator formulas for normal derivative as well. In addition,

$$\Delta_x \mathbf{u} = \partial_{\mathbf{n}}^2 \mathbf{u} + \sum_{i=1}^{d-1} \partial_{\boldsymbol{\tau}_i}^2 \mathbf{u} + L.O.T.,$$
$$\nabla_x \cdot (\nabla_x \mathbf{u} |\nabla_x \mathbf{u}|^r) = \nabla_{(\boldsymbol{\tau},\mathbf{n})} \cdot (\nabla_{(\boldsymbol{\tau},\mathbf{n})} \mathbf{u} |\nabla_{(\boldsymbol{\tau},\mathbf{n})} \mathbf{u}|^r) + L.O.T. \tag{A.12}$$

Since all "L.O.T." are under the control analogously to estimate (A.11), we may replace x-derivatives by $(\boldsymbol{\tau}, \mathbf{n})$-derivatives and do calculations simply assuming that

they commute. At this step we multiply equation (A.1) by $-\partial_{\boldsymbol{\tau}_i}^*(\partial_{\boldsymbol{\tau}_i}\mathbf{u}|\nabla_{(\boldsymbol{\tau},\mathbf{n})}\mathbf{u}|^r)$ and integrate over x. Analogously to Step 1, this gives

$$
\begin{aligned}
(\mathbf{D}\partial_{\mathbf{n}}^2\mathbf{u} + \mathbf{D}\sum_{j=1}^{d-1}\partial_{\boldsymbol{\tau}_j}^2\mathbf{u}, \ \partial_{\boldsymbol{\tau}_i}^2\mathbf{u}|\nabla_{(\boldsymbol{\tau},\mathbf{n})}\mathbf{u}|^r) + \\
+(\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u})\partial_{\boldsymbol{\tau}_i}\mathbf{u}, \ \partial_{\boldsymbol{\tau}_i}\mathbf{u}|\nabla_x\mathbf{u}|^r) \leq \\
\leq \varepsilon(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r) + C_\varepsilon\left(\|\mathbf{g}\|_{L^q}^{r+2} + \|\nabla_x\mathbf{u}\|_{L^{r+2}}^{r+2}\right).
\end{aligned} \quad (A.13)
$$

The advantage of separating tangential and normal derivatives is the fact that

$$
\partial_{\boldsymbol{\tau}_i}\mathbf{u}\big|_{\partial\Omega} = \partial_{\boldsymbol{\tau}_i}^2\mathbf{u}\big|_{\partial\Omega} = 0,
$$

due to the Dirichlet boundary conditions, so we again may integrate by parts the expression in the left-hand side of (A.13) freely and, analogously to Step 1, get the following estimate (using also that $\nabla_{\mathbf{u}}\mathbf{f}(\mathbf{u}) \geq 0$):

$$
\begin{aligned}
(|D_{\boldsymbol{\tau},\boldsymbol{\tau}}^2\mathbf{u}|^2 + |D_{\mathbf{n},\boldsymbol{\tau}}^2\mathbf{u}|^2, \ |\nabla_x\mathbf{u}|^r) \leq \\
\leq \varepsilon(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r) + C_\varepsilon\left(\|\mathbf{g}\|_{L^q}^{r+2} + \|\nabla_x\mathbf{u}\|_{L^{r+2}}^{r+2}\right),
\end{aligned} \quad (A.14)
$$

where $D_{\boldsymbol{\tau},\boldsymbol{\tau}}^2$ and $D_{\mathbf{n},\boldsymbol{\tau}}^2$ are the collections of all tangential second derivatives $\partial_{\boldsymbol{\tau}_i}\partial_{\boldsymbol{\tau}_j}$ and the second derivatives of the form $\partial_{\mathbf{n}}\partial_{\boldsymbol{\tau}_i}$ respectively.

Step 4. Boundary estimates: normal derivatives. To estimate normal derivatives we multiply (A.1) by $\partial_{\mathbf{n}}^*(\partial_{\mathbf{n}}\mathbf{u}|\nabla_x\mathbf{u}|^r)$ and use that $\mathbf{f}(0) = 0$ in order to kill boundary terms related with nonlinearity. Together with the already obtained estimate (A.14) this gives

$$
\begin{aligned}
(\partial_{\mathbf{n}}^2\mathbf{u}, \ \partial_{\mathbf{n}}^2\mathbf{u}|\nabla_x\mathbf{u}|^r) \\
\leq \varepsilon(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r) + C_\varepsilon\left(\|\mathbf{g}\|_{L^q}^{r+2} + \|\nabla_x\mathbf{u}\|_{L^{r+2}}^{r+2}\right) \\
+|(|D_{\boldsymbol{\tau},\boldsymbol{\tau}}^2\mathbf{u}|, \ |\partial_{\mathbf{n}}^2\mathbf{u}||\nabla_x\mathbf{u}|^r)| \\
\leq \varepsilon(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r) + C_\varepsilon\left(\|\mathbf{g}\|_{L^q}^{r+2} + \|\nabla_x\mathbf{u}\|_{L^{r+2}}^{r+2}\right),
\end{aligned} \quad (A.15)
$$

where the constant $\varepsilon > 0$ can be chosen arbitrarily small.

Step 5. Combining all together. Combining the interior estimates obtained at Step 2 with the tangential and normal boundary estimates (A.14) and (A.15) (e.g., with the help of the proper partition of unity), we finally arrive at

$$
(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r) \leq C_\varepsilon(\|\mathbf{g}\|_{L^{r+2}}^{r+2} + \|\nabla_x\mathbf{u}\|_{L^{r+2}}^{r+2}) + C\varepsilon(|D_x^2\mathbf{u}|, |\nabla_x\mathbf{u}|^r)
$$

and after fixing $\varepsilon > 0$ small enough, we end up with

$$
(|D_x^2\mathbf{u}|^2, |\nabla_x\mathbf{u}|^r) \leq C(\|\mathbf{g}\|_{L^{r+2}}^{r+2} + \|\nabla_x\mathbf{u}\|_{L^{r+2}}^{r+2}).
$$

So, it only remains to estimate the L^{r+2}-norm of the gradient $\nabla_x \mathbf{u}$. To this end, it is enough to use (A.7) together with the obvious estimate

$$\|\nabla_x \mathbf{u}\|_{L^{r+2}}^{r+2} \leq \varepsilon \|\nabla_x \mathbf{u}\|_{L^{\frac{d(r+2)}{d-2}}}^{r+2} + C_\varepsilon \|\nabla_x \mathbf{u}\|_{L^2}^{r+2}$$

and estimate (A.2). This completes the proof of the theorem. $\qquad\square$

Remark A.2 As we see, the nonlinear localization uses the general strategy of the classical (linear) localization technique. However, it is more delicate since we need also to treat the nonlinear term $\mathbf{f}(\mathbf{u})$ which is now *not subordinated* to the linear ones, so we can multiply the equation only on the terms which can be estimated using the monotonicity assumption $\nabla_\mathbf{u} \mathbf{f}(\mathbf{u}) \geq 0$. Fortunately, the amount of such multipliers is enough to get the estimates similar to the linear case.

References

1. Akagi, G., Schimperna, G., Segatti, A.: Fractional Cahn-Hilliard, Allen-Cahn and Porous Medium Equations. J. Diff. Eqns **261**(6), 2935–2985 (2016)
2. Andreu-Vaillo, F., Mazon, J., Rossi, J., Toledo-Melero, J.: Nonlocal Diffusion Problems, Mathematical Surveys and Monographs. American Mathematical Society, Real Sociedad Matemática Espanola, vol. 165. Providence, RI Madrid (2010)
3. Babin, A., Vishik, M.: Attractors of Evolution Equations, Studies in Mathematics and its Applications, vol. 25. North-Holland Publishing Co., Amsterdam (1992)
4. Ball, J., Currie, J.: Null Lagrangians, Weak Continuity and Variational Problems of Arbitrary Order. J. Func. Anal. **41**, 135–174 (1981)
5. Brezis, H.: Operateurs Maximaux Monotones et Semigroupes de Contractions dans les Espaces de Hilbert. North-Holland, Amstredam (1973)
6. Budd, C., Rottschäfer, V., Williams, J.: Multibump, Blow-Up, Self-Similar Solutions of the Complex Ginzburg-Landau Equation. SIAM J. Applied Dynamical Systems **4**(3), 649–678 (2005)
7. Caffarelli, L., Stinga, P.: Fractional Elliptic Equations, Caccioppoli Estimates and Regularity. Ann. H. Poincare - AN **33**, 767–807 (2016)
8. Chepyzhov, V., Vishik, M.: Attractors for Equations of Mathematical Physics, American Mathematical Society Colloquium Publications. American Mathematical Society, vol.49. Providence, RI (2002)
9. Cholewa, J., Dlotko, T.: Parabolic Equations with Critical Nonlinearities. Top. Meth. Non. Anal. **21**, 311–324 (2003)
10. Cholewa, J., Dlotko, T.: Global Attractors in Abstract Parabolic Problems, LMS Lecture Note Series. Cambridge University Press, Cambridge (2019)
11. Chueshov, I.: Long-Time Dynamics of Kirchhoff Wave Models with Strong Nonlinear Damping. J. Diff. Eqns **252**(2), 1229–1262 (2012)
12. Canizo, J., Desvillettes, L., Fellner, K.: Improved duality estimates and applications to reaction-diffusion equations. Commun. PDEs **39**(4), 1185–1204 (2014)
13. Doering, C., Gibbon, J., Levermore, C.: Weak and Strong Solutions of the Complex Ginzburg-Landau Equation. Physica D **71**, 285–318 (1994)
14. Dlotko, T., Liang, T., Wang, Y.: Critical and Super-Critical Abstract Parabolic Equations. Discrete Contin. Dyn. Syst. Ser. B **25**(4), 1517–1541 (2020)
15. Du, Y., Ishii, H., Lin, W.: (Editors), Recent Progress on Reaction-Diffusion Systems and Viscosity Solutions, World Scientific Pub Co Inc, (2009)
16. Eden, A., Foias, C., Nicolaenko, B., Temam, R.: Exponential Attractors for Dissipative Evolution Equations. Wiley/Masson, Chichester/Paris (1994)
17. Efendiev, M., Miranville, A., Zelik, S.: Exponential Attractors for a Nonlinear Reaction-Diffusion System in $\mathbb{R}^3$. C. R. Acad. Sci. Paris **330**, 713–718 (2000)
18. Efendiev, M., Miranville, A., Zelik, S.: Exponential Attractors and Finite-Dimensional Reduction for Nonautonomous Dynamical Systems. Proc. R. Soc. Edinburg, Sect. A **135**, 703–730 (2005)

19. Efendiev, M., Zelik, S.: The Regular Attractor for the Reaction-Diffusion System with a Nonlinearity Rapidly Oscillating in Time and it's Averaging. Adv. Diff. Eqns **8**(6), 673–732 (2003)
20. Efendiev, M., Zelik, S.: Finite Dimensional Attractors and Exponential Attractors for Degenerate Doubly Nonlinear Equations. Math. Meth. Appl. Sci. **32**(13), 1638–1668 (2009)
21. Efendiev, M., Gajewski, H., Zelik, S.: The Finite Dimensional Attractor for a 4th Order System of Cahn-Hilliard Type with a Supercritical Nonlinearity. Adv. Diff. Eqns **7**(9), 1072–1100 (2002)
22. Fellner, K., Morgan, J., Tang, B.: Global classical solutions to quadratic systems with mass control in arbitrary dimensions. Ann. Inst. H. Poincaré Anal. Non Linéaire **37**(2), 281–307 (2020)
23. Gal, C., Warma, M.: Reaction-Diffusion Equations with Fractional Diffusion on Nonsmooth Domains with Various Boundary Conditions. Discrete Contin. Dyn. Syst. **36**, 1279–1319 (2016)
24. Hale, J.: Asymptotic Behavior of Dissipative Systems, Mathematical Surveys and Monographs, no. 25, Amer. Math. Soc., Providence, R. I., (1988)
25. Herrero, M., Velázquez, J.: A Blow-up Mechanism for a Chemotaxis Model. Ann. Scuola Norm. Sup. Pisa Cl. Sci. **24**(4), 633–683 (1998)
26. Hu, B.: Blow-up theories for semilinear parabolic equations. Lecture Notes in Mathematics, vol. 2018. Springer, Heidelberg (2011)
27. Kalantarov, V., Zelik, S.: Finite-Dimensional Attractors for the Quasi-Linear Strongly Damped Wave Equation. J. Diff. Eqns **247**(4), 1120–1155 (2009)
28. Kalantarov, V., Zelik, S.: Smooth Attractors for the Brinkman-Forchheimer Equations with Fast Growing Nonlinearities. Comm. Pure Appl. Anal. **11**(5), 2037–2054 (2012)
29. Ladyzhenskaya, O., Solonnikov, V., Uraltseva, N.: Linear and Quasi-Linear Equations of Parabolic Type, AMS, (1995)
30. Ladyzhenskaya, O.: Attractors of Nonlinear Evolution Problems with Dissipation, (Russian) Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 152 (1986), Kraev. Zadachi Mat. Fiz. i Smezhnye Vopr. Teor. Funktsii 18, 72–85, 182
31. Li, X., Shen, W., Sun, C.: Asymptotic Dynamics of Non-Autonomous Fractional Reaction-Diffusion Equations on Bounded Domains. Top. Meth. Non. Anal. **55**(1), 105–139 (2020)
32. Málek, J., Necas, J.: A Finite-Dimensional Attractor for Three-Dimensional Flow of Incompressible Fluids. J. Diff. Eqns **127**, 498–518 (1996)
33. Malek, J., Prazak, D.: Large Time Behavior via the Method of *l*-Trajectories. J. Diff. Eqns **181**(2), 243–279 (2002)
34. Martin, R., Pierre, M.: Nonlinear Reaction Diffusion Systems, in Nonlinear Equations in the Applied Sciences, W. F. Ames and C. Rogers, eds., Math. Sci. Engrg. 185, Academic Press, New York, (1991)
35. Miranville, A., Zelik, S.: Attractors for Dissipative Partial Differential Equations in Bounded and Unbounded Domains, Handbook of differential equations: evolutionary equations, pp. 103–200. Elsevier/North-Holland, Amsterdam IV, Handb. Differ. Equ. (2008)
36. Pierre, M., Schmidt, D.: Blowup in Reaction-Diffusion Systems with Dissipation of Mass. SIAM Review **42**(1), 93–106 (2000)
37. Pierre, M.: Global Existence in Reaction-Diffusion Systems with Control of Mass: a Survey. Milan J. Math **78**, 417–455 (2010)
38. Quittner, P., Souplet, P.: Superlinear parabolic problems. Blow-up, global existence and steady states, Second edition, Birkhauser Advanced Texts: Basler Lehrbucher, Birkhauser/Springer, Cham, (2019)
39. Robinson, J.: Infinite-Dimensional Dynamical Systems. Cambridge University Press, Cambridge (2001)
40. Smoller, J.: ShockWaves and Reaction-Diffusion Equations. Springer-Verlag, New York (1983)
41. Temam, R.: Infinite-Dimensional Dynamical Systems in Mechanics and Physics. Springer, New York (1997)
42. Triebel, H.: Interpolation Theory. Function Space, Differential Operators, North- Holland, Amsterdam-New York (1978)
43. Yagi, A.: Abstract Parabolic Evolution Equations and their Applications. Springer- Verlag, Heidelberg (2010)
44. Vázquez, J.: Smoothing and Decay Estimates for Nonlinear Diffusion Equations. Equations of Porous Medium Type, Oxford Lecture Ser. Math. Appl., vol. 33, Oxford University Press, (2006)
45. Zeidler, E.: Nonlinear Functional Analysis and its Applications. Nonlinear monotone operators, Springer-Verlag, New York, II/B (1990)
46. Zelik, S.: The Attractor for a Non-linear Reaction-Diffusion System with a Supercritical Non-linearity and its Dimension. Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. **24**, 1–25 (2000)

47. Zelik, S.: Attractors of Reaction-Diffusion Systems in Unbounded Domains and their Spatial Complexity. Comm. Pure Appl. Math. **56**(5), 584–637 (2003)
48. Zhong, Y., Sun, C.: Uniform Quasi-Differentiability of Semigroup to Nonlinear Reaction-Diffusion Equations with Supercritical Exponent. Acta Mathematica Scientia **37B**(2), 301–315 (2017)