NETWORKS AND HETEROGENEOUS MEDIA doi:10.3934/nhm.2014.9.353
©American Institute of Mathematical Sciences
Volume **9**, Number **2**, June **2014** pp. **353–382**

TWO-SCALE HOMOGENIZATION OF NONLINEAR REACTION-DIFFUSION SYSTEMS WITH SLOW DIFFUSION

Alexander Mielke

Weierstrass Institute
Mohrenstraße 39
10117 Berlin, Germany
and
Institut für Mathematik, Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin-Adlershof, Germany

Sina Reichelt and Marita Thomas

Weierstrass Institute
Mohrenstraße 39
10117 Berlin, Germany

(Communicated by Dag Lukkassen)

Abstract. We derive a two-scale homogenization limit for reaction-diffusion systems where for some species the diffusion length is of order 1 whereas for the other species the diffusion length is of the order of the periodic microstructure. Thus, in the limit the latter species will display diffusion only on the microscale but not on the macroscale. Because of this missing compactness, the nonlinear coupling through the reaction terms cannot be homogenized but needs to be treated on the two-scale level. In particular, we have to develop new error estimates to derive strong convergence results for passing to the limit.

1. **Introduction.** The theory of periodic homogenization is concerned with partial differential equations with periodically oscillating coefficients with small period ε and describes ways for finding a homogenized partial differential equation of which the solutions are the weak limits for $\varepsilon \to 0$ of the original solutions. We refer to the books [25, 28, 49] for general introductions and surveys. An important step in homogenization theory was the introduction of two-scale convergence in [41], which allows for the treatment of more general equations. While the original notion of two-scale convergence in [41, 1, 9] can be called weak two-scale convergence, it is crucial that one can also introduce a notion of strong two-scale convergence, see [53, 54, 35]. It is this strong convergence which can be used to study fully nonlinear problems like nonsmooth elastoplasticity (cf. [39, 55, 48, 22, 20]) and to allow for efficient numerical approximation [30].

2010 *Mathematics Subject Classification.* 35B25, 35A01, 35K57, 35K65, 35M10.

Key words and phrases. Two-scale convergence, folding and unfolding, coupled reaction-diffusion equations, nonlinear reaction, degenerating diffusion, Gronwall estimate.

This research was supported by the *C*ollaborative Research Center 910: Control of self-organizing nonlinear systems: Theoretical methods and concepts of application via the project *A*5: Pattern formation in systems with multiple scales.

The present work applies the ideas of strong two-scale convergence to nonlinear reaction-diffusion systems, which for fixed $\varepsilon > 0$ are simple semilinear parabolic systems. The difficulty arises in the limit of $\varepsilon \to 0$, since we allow some of the diffusion constants to scale with ε. Hence, we lose the compactness from the diffusion terms, but the nonlinearities in the reaction term can only be treated by strong convergence. The latter will be obtained by exploiting the fact that strong two-scale convergence can be measured in the norm topology of the two-scale functions and that these errors can be controlled by suitable Gronwall estimates.

To be more precise, we consider the following inhomogeneous system of coupled reaction-diffusion equations:

$$\begin{aligned} u_t^\varepsilon(t,x) &= \operatorname{div}(\mathbb{D}_1(x,\tfrac{x}{\varepsilon})\nabla u^\varepsilon(t,x)) + f_1(t,x,\tfrac{x}{\varepsilon},u^\varepsilon(t,x),v^\varepsilon(t,x)),\\ v_t^\varepsilon(t,x) &= \operatorname{div}(\varepsilon^2\mathbb{D}_2(x,\tfrac{x}{\varepsilon})\nabla v^\varepsilon(t,x)) + f_2(t,x,\tfrac{x}{\varepsilon},u^\varepsilon(t,x),v^\varepsilon(t,x)), \end{aligned} \qquad (1.1\mathrm{P}_\varepsilon^{\mathrm{cp}})$$

for $t \geq 0$ and on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$. We also add no-flux boundary conditions on $\partial\Omega$. Here, $\mathbb{D}_i \in \mathbb{R}^{m_i\times m_i}$ are diffusion tensors and f_i are reaction terms acting on the vector of concentrations $u^\varepsilon \in \mathbb{R}^{m_1}$, resp. $v^\varepsilon \in \mathbb{R}^{m_2}$ referring to m_1, resp. m_2 different species. The scaling ε^2 of $\mathbb{D}_2$ takes into account that the species related to the concentration vector v^ε diffuse much slower than those related to u^ε. Therefore, we call v^ε the slow diffusive variable and u^ε the "classically" diffusing one. We also call $(1.1\mathrm{P}_\varepsilon^{\mathrm{cp}})_1$ the non-degenerating part, while $(1.1\mathrm{P}_\varepsilon^{\mathrm{cp}})_2$ is called the degenerating part.

The coupling of the variables $(u^\varepsilon, v^\varepsilon)$ occurs via the reaction terms (f_1, f_2). To focus on the difficulties of the homogenization limit $\varepsilon \to 0$, we avoid any questions concerning global existence or positivity of the concentrations by assuming Lipschitz continuity, namely

$$(f_1, f_2) \text{ is differentiable and globally Lipschitz continuous in } (u^\varepsilon, v^\varepsilon), \qquad (1.2)$$

see (2.5L)&($2.6C_\infty$) for a precise statement of the assumptions and Example 2.3 for a discussion. But as we explain in the following, the assumption of nonlinearity complicates the limit passage $\varepsilon \to 0$ in $(1.1\mathrm{P}_\varepsilon^{\mathrm{cp}})$. In particular, confining the analysis to given data $(\mathbb{D}_1, \mathbb{D}_2)$ and (f_1, f_2) being spatially $O(\varepsilon)$-periodic in the x/ε-component, we show that for $\varepsilon \to 0$, the limit model is a two-scale model, given for $t \geq 0$ on $(x, y) \in \Omega \times \mathcal{Y}$ by

$$\begin{aligned} u_t(t,x) &= \operatorname{div}(\mathbb{D}_{\mathrm{eff}}(x)\nabla u(t,x)) + f_{\mathrm{eff}}(t,x,u(t,x),V(t,x,*)),\\ V_t(t,x,y) &= \operatorname{div}_y(\mathbb{D}_2(x,y)\nabla_y V(t,x,y)) + f_2(t,x,y,u(t,x),V(t,x,y)). \end{aligned} \qquad (1.3\mathrm{P}_0^{\mathrm{cp}})$$

Above, $\mathbb{D}_{\mathrm{eff}}$ denotes the classical effective diffusion tensor, cf. [2, 1, 26]. However, the reaction term f_{eff} is a macroscopic, one-scale function, but depending on the microscopic function $V(t, x, *)$, namely

$$f_{\mathrm{eff}}(t,x,u(t,x),V(t,x,*)) := \int_{\mathcal{Y}} f_1(t,x,y,u(t,x),V(t,x,y))\,\mathrm{d}y,$$

where $\mathcal{Y} = \mathbb{R}^d/_{\mathbb{Z}^d}$ denotes the so-called *periodicity cell*, which can be obtained from the unit cell $Y = [-1/2, 1/2)^d$ by identifying opposite faces of $\overline{Y}$. In contrast, the effective data $\mathbb{D}_2$ and f_2 in $(1.3\mathrm{P}_0^{\mathrm{cp}})_2$ are indeed two-scale functions, which additionally depend on $y \in \mathcal{Y}$. We use the term *effective model* rather than *homogenized model*, since in our case it is not possible to reduce the two-scale model $(1.3\mathrm{P}_0^{\mathrm{cp}})$ to a homogenized macroscopic one-scale model. Nevertheless, our effective model is independent of ε and hence it can be analyzed (or discretized) without any singular perturbation through a small parameter, cf. [30, 36, 4, 5].

In Section 2 we discuss the existence and uniqueness of a weak solution for both the original one-scale problem ($1.1\mathrm{P}^{\mathrm{cp}}_{\varepsilon}$) and the effective problem ($1.3\mathrm{P}^{\mathrm{cp}}_{0}$) simultaneously, making use of a suitable abstract setting. To perform the limit passage ($1.1\mathrm{P}^{\mathrm{cp}}_{\varepsilon}$) $\xrightarrow{\varepsilon\to 0}$ ($1.3\mathrm{P}^{\mathrm{cp}}_{0}$) we develop the necessary tools of two-scale convergence in Section 3. Since the equations in ($1.1\mathrm{P}^{\mathrm{cp}}_{\varepsilon}$) feature different scalings, their limit passage can be carried out using different methods. The limit in the non-degenerating equation $(1.1\mathrm{P}^{\mathrm{cp}}_{\varepsilon})_1$ can be obtained with the classical theory of G- or two-scale convergence and by exploiting the compact embedding of $H^1(\Omega)$ into $L^2(\Omega)$ for handling the nonlinear reaction terms f_1. For homogenization results based on two-scale convergence in the case of non-degenerating quasilinear parabolic PDEs we refer to e.g. [40, 13, 56, 16, 27, 47] and for monotone parabolic operators and multiscale-convergence in space and time, to [19, 43] and references therein.

Compared to the treatment of $(1.1\mathrm{P}^{\mathrm{cp}}_{\varepsilon})_1$, the limit passage in the degenerating equation $(1.1\mathrm{P}^{\mathrm{cp}}_{\varepsilon})_2$ is much more involved and needs special attention. This is why we first elaborate this limit passage separately in Section 4 and then merge it with the classical procedure to pass to the limit with the full system in Section 5. Let us explain the difficulties coming along with the degenerating equation and the nonlinear reaction terms.

Since the ellipticity of the diffusion tensor in $(1.1\mathrm{P}^{\mathrm{cp}}_{\varepsilon})_2$ degenerates for $\varepsilon \to 0$, the general theory of G-convergence (see e.g. [37]) is not suited here. But the concept of two-scale convergence, introduced in [41], is applicable, if $\mathbb{D}_i^{\varepsilon}$ and f_i^{ε}, $i = 1, 2$, are ε-periodic in $x \in \Omega$. Hereby, we operate with the equivalent definition of two-scale convergence formulated in [6, 35, 7] via the *periodic unfolding operator* $\mathcal{T}_{\varepsilon} : L^2(\Omega) \to L^2(\mathbb{R}^d \times \mathcal{Y})$, see (3.2). With the aid of $\mathcal{T}_{\varepsilon}$, weak and strong L^2–two-scale convergence can be defined in terms of weak and strong L^2–convergence of two-scale functions, see Definition 3.3.

For the degenerating equation $(1.3\mathrm{P}^{\mathrm{cp}}_{0})_2$ we will make use of the following *compactness result*, cf. [1, 6], Theorem 3.5: If $(v^{\varepsilon})_{\varepsilon} \subset H^1(\Omega)$ satisfies the a priori bound

$$\exists\, C \geq 0\ \forall\, \varepsilon > 0 : \quad \|v^{\varepsilon}\|_{L^2(\Omega)} + \varepsilon \|\nabla v^{\varepsilon}\|_{L^2(\Omega)} \leq C, \tag{1.4}$$

then there exists a two-scale function $V \in L^2(\Omega; H^1(\mathcal{Y}))$, and, up to a subsequence, we have weak two-scale convergence in the following sense

$$\mathcal{T}_{\varepsilon}(v^{\varepsilon}) \rightharpoonup V^{\mathrm{ex}} \quad \text{and} \quad \mathcal{T}_{\varepsilon}(\varepsilon \nabla v^{\varepsilon}) \rightharpoonup (\nabla_y V)^{\mathrm{ex}} \quad \textit{weakly in } L^2(\mathbb{R}^d \times \mathcal{Y}). \tag{1.5}$$

For a function $A \in L^2(\Omega \times \mathcal{Y})$, $A^{\mathrm{ex}} \in L^2(\mathbb{R}^d \times \mathcal{Y})$ denotes its extention with 0 outside of Ω, which, throughout this work, will be assumed to have Lipschitz boundary. Moreover, $H^1(\mathcal{Y}) \subset H^1(Y)$ is the subspace of functions with periodic boundary values.

The weak two-scale convergence from (1.5) is well suited to pass to the limit $\varepsilon \to 0$ in linear equations of the type $v_t^{\varepsilon} = \mathrm{div}(\varepsilon^2 \mathbb{D} \nabla v^{\varepsilon}) + f^{\varepsilon} \cdot v^{\varepsilon}$, see [1, 44, 31, 34]. But since we deal with nonlinear reaction terms, we need the strong convergence of $\mathcal{T}_{\varepsilon}(v^{\varepsilon}) \to V^{\mathrm{ex}}$ in $L^2(\mathbb{R}^d \times \mathcal{Y})$, which does not follow from (1.4)&(1.5). Therefore the limit passage in the semilinear equation $(1.1\mathrm{P}^{\mathrm{cp}}_{\varepsilon})_2$ is not straightforward and needs additional assumptions. For example, in [24], an effective system is rigorously derived for a degenerating equation with nonlinear reaction terms that are not directly coupled, i.e. $f_1^{\varepsilon}(u^{\varepsilon}, v^{\varepsilon}) = f_1(u^{\varepsilon})$ and $f_2^{\varepsilon}(u^{\varepsilon}, v^{\varepsilon}) = f_2(v^{\varepsilon})$. Assuming further that f_2 is the gradient of a λ-convex potential ϕ, a homogenization result is established using methods from convex analysis, solely based on *weak* two-scale convergence; strong two-scale convergence is not investigated. Moreover, in light of the general gradient structures for reaction-diffusion systems in [32, 33] it is clear that the assumption that f_2 has a potential ϕ is only reasonable for v_2 being scalar. For nonlinear problems, without the property of compactness, it is necessary to use concepts based on strong two-scale convergence, see e.g. [51, 35, 22, 52]. In a similar spirit, [11] uses two-scale correctors to prove strong convergence with explicit convergence rates. However, there the assumptions are rather strong, e.g. $\nabla_x V \in L^2(\Omega; H^1(\mathcal{Y}))$ and continuity w.r.t. $y \in \mathcal{Y}$ of all functions in $(1.3\mathrm{P}_0^{\mathrm{cp}})_2$. Quantitative homogenization results also exist for attractors for nonlinear reaction-diffusion systems, e.g. [17, 18].

In some sense, we are following a similar strategy as in [11], but we do not need any additional regularity in the x or y variables. Our main result Theorem 4.1 shows that the weak solutions v^ε and V of $(1.1\mathrm{P}_\varepsilon^{\mathrm{cp}})_2$ and $(1.3\mathrm{P}_0^{\mathrm{cp}})_2$, respectively, satisfy

$$\| \mathcal{T}_\varepsilon v^\varepsilon(t) - V^{\mathrm{ex}}(t)\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} \xrightarrow{\varepsilon\to 0} 0 \quad \text{uniformly in } [0,T]. \tag{1.6}$$

We neither assume that v^ε admits an asymptotic expansion in ε nor that V is continuous in space, and yet we prove (1.6) rigorously. If V were spatially continuous, then (1.6) would be equivalent to $\|v^\varepsilon(t) - [V]^\varepsilon(t)\|_{L^2(\Omega)} \xrightarrow{\varepsilon\to 0} 0$, where we set $[V]^\varepsilon(x) := V(x, x/\varepsilon)$ as in [11], see Remark 5.4.

The general strategy for proving this strong convergence is explained in an abstract way in Section 4.2. For the difference $W^\varepsilon(t) = \mathcal{T}_\varepsilon v^\varepsilon(t) - V^{\mathrm{ex}}(t)$ we derive a Gronwall estimate of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\|W^\varepsilon(t)\|_2^2 \le L\|W^\varepsilon(t)\|_2^2 + \Delta_1^\varepsilon(t) + \Delta_2^\varepsilon(t) + \Delta_3^\varepsilon(t) + \Delta_4^\varepsilon(t),$$

where the error terms Δ_i^ε are shown to converge pointwise to 0. To derive this estimate, we reformulate the weak formulations of $(1.1\mathrm{P}_\varepsilon^{\mathrm{cp}})_2$ and $(1.3\mathrm{P}_0^{\mathrm{cp}})_2$ via the unfolding operator $\mathcal{T}_\varepsilon$, the folding operator $\mathcal{F}_\varepsilon$, and the gradient folding operator $\mathcal{G}_\varepsilon$, see (3.5) and Definition 3.7. The rough idea is to subtract these prepared weak equations and to test with the difference W^ε. While the errors Δ_3^ε and Δ_4^ε measure the standard approximation errors of the linear diffusion terms and the nonlinear reaction terms, respectively, there occur two additional and more difficult errors to be controlled. The error Δ_1^ε arises through the fact, that we need to invert the unfolding operator $\mathcal{T}_\varepsilon$ in two different ways, namely first by the classical folding operator $\mathcal{F}_\varepsilon$ and second by the gradient folding operator $\mathcal{G}_\varepsilon$, see Theorem 3.9 in Section 3.4, where we follow ideas in [35, 22].

However, the most difficult error term Δ_2^ε arises from the fact that we cannot test with W^ε directly, since the unfolding operator $\mathcal{T}_\varepsilon$ is incapable to directly implant Y-periodicity, as $\mathcal{T}_\varepsilon v^\varepsilon \in L^2(\mathbb{R}^d; H^1(Y)) \nsubseteq L^2(\mathbb{R}^d; H^1(\mathcal{Y}))$ for all $v^\varepsilon \in H^1(\Omega)$, see Theorem 3.2. But as a consequence of the compactness result (1.5), Y-periodicity will be a characteristic feature of the admissible functions for the limit problem $(1.3\mathrm{P}_0^{\mathrm{cp}})_2$. It is a well-known fact (cf. [41, Thm. 3], [1, Prop. 1.14], [53, Thm. 6.1]) that the two-scale limit V is Y-periodic, although the unfolded sequence $\mathcal{T}_\varepsilon v^\varepsilon$ is in general not Y-periodic, see in particular [6, Prop. 3] and [10, Thm. 5.2] for a proof in the periodic unfolding formulation. In other words, the unfolding operator $\mathcal{T}_\varepsilon$ is incapable to directly generate Y-periodicity, but automatically ensures the recovery of periodicity in the weak two-scale limit. Since this effect plays a crucial role in our analysis, we term it the

$$\begin{aligned}&\mathcal{T}_\varepsilon\text{-}\textit{property of recovered periodicity:} \text{ while } \mathcal{T}_\varepsilon u^\varepsilon \in L^2(\mathbb{R}^d; H^1(Y)) \text{ only,}\\ &\text{we have } U^{\mathrm{ex}} = \operatorname*{w\text{-}lim}_{\varepsilon\to 0} \mathcal{T}_\varepsilon u^\varepsilon \in L^2(\mathbb{R}^d; H^1(\mathcal{Y})) \subsetneq L^2(\mathbb{R}^d; H^1(Y)).\end{aligned} \tag{1.7}$$

In principle, our method is strong enough to supply quantitative error estimates as in [11], but this will be subject of future work.

As a further technical issue let us mention that a priori $v_t^\varepsilon(t) \in H^1(\Omega)^*$, merely, whereas the operator $\mathcal{T}_\varepsilon$ is well-defined for integrable functions, only. To avoid technicalities we therefore improve the time-regularity of the weak solutions in Proposition 2.2 by imposing the differentiability on (f_1, f_2) and additional regularity on the initial datum $(u_0^\varepsilon, v_0^\varepsilon)$, see (2.9). For the reader's convenience we here provide a list of spaces used throughout this work:

$H = L^2(\Omega)$, $X = H^1(\Omega)$	function spaces in the abstract setting, see (2.2);
$X_\varepsilon = (X, \\| \cdot \\|_{X_\varepsilon})$	ε-weighted function space for the species v^ε, see (4.5);
$\mathbb{H} = L^2(\Omega \times \mathcal{Y})$, $\mathbb{X} = L^2(\Omega; H^1(\mathcal{Y}))$	function spaces for the limit problem (4.4P$_0$), see (4.6);
$\overline{H}$, $\overline{X}_\varepsilon$, $\overline{\mathbb{H}}$, $\overline{\mathbb{X}}$	function spaces for the coupled systems, see (5.4).

2. **Assumptions and existence of weak solutions.** Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary $\Gamma = \partial\Omega$, which ensures the validity of the Sobolev embedding theorems, and let $T > 0$ be fixed. We abbreviate the time-space cylinder $(0, T) \times \Omega$ with Ω_T and analogously we write Γ_T for $(0, T) \times \Gamma$. Moreover let $\vec{n}$ denote the outer unit normal vector of Ω. The focus of the paper are nonlinear reaction-diffusion equations of the type

$$\begin{aligned} u_t &= \mathcal{A}u + f(u) && \text{in } \Omega_T, \\ 0 &= (\mathbb{D}\nabla u) \cdot \vec{n} && \text{on } \Gamma_T, \\ u(0) &= u_0 && \text{in } \Omega. \end{aligned} \tag{2.1P}$$

Here $\mathcal{A}$ denotes an elliptic differential operator of the form $\mathcal{A}u = \operatorname{div}(\mathbb{D}\nabla u)$. For the application we have in mind, $u : [0, T] \times \Omega \to \mathbb{R}^m$ denotes the concentration, $\mathbb{D} : \Omega \to \mathbb{R}^{(m\times d)\times(m\times d)}$ the diffusion tensor, and $f : \Omega \times \mathbb{R}^m \to \mathbb{R}^m$ the reaction term.

Both systems, (1.1P$_\varepsilon^{\text{cp}}$) and (1.3P$_0^{\text{cp}}$), can be reformulated in terms of (2.1P). In this section, we present a mathematical setting that accounts for both systems and that is independent of $\varepsilon > 0$ and $y \in \mathcal{Y}$. We introduce the notion of weak solutions in Section 2.1 and give results concerning the existence of weak solutions and improved time-regularity in Section 2.2.

2.1. **Weak formulation and data qualification.** Let X and H denote two given Hilbert spaces. We denote with X^* the dual space of X and with $\langle \cdot, \cdot \rangle_{X^*,X}$ the associated dual pairing. We assume that H can be identified with its dual, i.e. $H = H^*$, and we write $(\cdot, \cdot)_H$ for the scalar product on H. Assume that X is dense and continuously embedded in H, then we obtain the evolution triple $X \subset H \subset X^*$. If not indicated otherwise, we set

$$X = H^1(\Omega) \quad \text{and} \quad H = L^2(\Omega), \tag{2.2}$$

and we call X the *space of test functions*. We always abbreviate $L^2(\Omega; \mathbb{R}^m)$ with $L^2(\Omega)$. For the evolution triple $X \subset H \subset X^*$, the relevant space for our analysis is $L^2(0, T; X) \cap H^1(0, T; X^*)$. By [14, Thm. 3 p. 287], we have that $L^2(0, T; X) \cap H^1(0, T; X^*)$ is continuously embedded in $C^0([0, T]; H)$. We call $u \in L^2(0, T; X) \cap H^1(0, T; X^*)$ a *weak solution* to (2.1P), if u satisfies a.e. in $(0, T)$ the *weak formulation*

$$\langle u_t, \varphi \rangle_{X^*,X} = (-\mathbb{D}\nabla u, \nabla\varphi)_H + \langle f(u), \varphi \rangle_{X^*,X} \quad \text{for all } \varphi \in X \tag{2.3WF}$$

and it holds $u(0) = u_0$. Since we are, among others, interested in the homogenization of the reaction term f, we do not want to understand $f(\cdot,\cdot,u(\cdot,\cdot))$ as general distribution (which is sufficient for the existence of weak solutions), but as an integrable function. Thus, we assume the reaction $f : u \mapsto f(u)$ to be differentiable and globally Lipschitz continuous (and not just locally) which is not too restrictive in practice as Example 2.1 shows.

Uniform Ellipticity: The diffusion tensor $\mathbb{D} \in L^\infty(\Omega;\mathbb{R}^{(m\times d)\times(m\times d)})$ is measurable and uniformly elliptic, i.e.

$$\exists\, \mu > 0 : \quad \mathbb{D}(x)A : A \geq \mu|A|^2 \quad \text{for all } A \in \mathbb{R}^{m\times d},\ \text{a.a. } x \in \Omega. \tag{2.4μ}$$

Lipschitz continuity: The reaction $f : [0,T]\times\Omega\times\mathbb{R}^m \to \mathbb{R}^m$ is measurable on Ω for all $(t,A) \in [0,T]\times\mathbb{R}^m$ and $f(\cdot,x,\cdot) \in C^1([0,T]\times\mathbb{R}^m)$ for a.a. $x \in \Omega$. Moreover,

$$\exists\, L > 0 : \quad |f(t,x,A) - f(t,x,B)| \leq L|A-B| \text{ for all } t,x,A,B. \tag{2.5L}$$

Boundedness: It holds

$$\exists\, D_\infty \geq 0 : \quad |\mathbb{D}(x)A| \leq D_\infty|A| \quad \text{for all } A \in \mathbb{R}^{m\times d},\ \text{a.a. } x \in \Omega, \tag{2.6D_∞}$$

$$\exists\, C_\infty \geq 0 : \quad |f(t,x,0)| \leq C_\infty \quad \text{for a.a. } (t,x) \in \Omega_T. \tag{2.6C_∞}$$

Here $A : B = \mathrm{tr}(A^t B)$ and $\vec{a}\cdot\vec{b}$ denote the scalar product for matrices in $\mathbb{R}^{m\times d}$ and for vectors in $\mathbb{R}^m$, respectively; $|\cdot|$ denotes the induced (matrix resp. vector) norm. For the sets of parameters (μ, D_∞) and (L, C_∞) with $\mu, L > 0$ and $D_\infty, C_\infty \geq 0$, we introduce the classes of functions

$$\begin{aligned}\mathbb{M}(\Omega,\mu,D_\infty) &:= \{\mathbb{D} : \Omega \to \mathbb{R}^{(m\times d)\times(m\times d)} \mid \\ &\qquad \mathbb{D} \text{ satisfies } (2.4\mu) \text{ and } (2.6D_\infty) \text{ with } (\mu,D_\infty)\} \text{ and} \\ \mathbb{F}(\Omega,L,C_\infty) &:= \{f : \Omega_T\times\mathbb{R}^m \to \mathbb{R}^m \mid \\ &\qquad f \text{ satisfies } (2.5\text{L}) \text{ and } (2.6C_\infty) \text{ with } (L,C_\infty)\}.\end{aligned}$$

For our analysis it is not necessary that $\mathbb{D}$ is symmetric. The assumptions (2.5L) and (2.6C_∞) guarantee $(t,x) \mapsto f(t,x,u(t,x)) \in L^2(0,T;H)$ for all $u \in L^2(0,T;H)$. Indeed, (2.5L) with $B = 0$ and (2.6C_∞) give the growth-condition

$$|f(t,x,A)| \leq \max\{L, C_\infty\}(1+|A|) \quad \text{for all } A \in \mathbb{R}^m,\ \text{a.a. } (t,x) \in \Omega_T. \tag{2.7C_1}$$

The existence result (Theorem 2.1) and the homogenization result (Theorem 4.1) do not rely on the homogeneous Neumann boundary conditions in (2.1P). In the case of non-homogeneous Neumann boundary conditions, the boundary integral $\int_\Gamma g\cdot\varphi\,\mathrm{d}\sigma$ would appear as linear term on the right-hand side in (2.3WF). Other choices such as Dirichlet or periodic boundary conditions are admissible as well and then $X = H^1_0(\Omega)$ or $X = H^1_{\mathrm{per}}(\Omega)$, respectively, and (2.3WF) holds as it is. A Poincaré-type inequality is not needed.

2.2. **Existence of weak solutions and improved time-regularity.** We emphasize that we do not use the compactness of the embedding $X \subset H$, recall (2.2), because in Section 4 we choose spaces $\mathbb{X} \subset \mathbb{H}$ (see (4.6)) that do not embed compactly.

Theorem 2.1. *Assume that $\mathbb{D} \in \mathbb{M}(\Omega,\mu,D_\infty)$, $f \in \mathbb{F}(\Omega,L,C_\infty)$ and $u_0 \in H$. Then there exists for every given $T > 0$ a unique weak solution $u \in L^2(0,T;X)\,\cap$*

$H^1(0,T;X^*)$ *to problem* (2.1P). *Moreover, there exists a positive constant* C_a *such that it holds*

$$\|u\|_{C([0,T];H)} + \sqrt{\mu}\|\nabla u\|_{L^2(0,T;H)} + \|u_t\|_{L^2(0,T;X^*)} \le C_a\left(1+\tfrac{D_\infty}{\sqrt{\mu}}\right), \tag{2.8}$$

where C_a *depends on the given quantities* $\|u_0\|_H, T, L, C_\infty, |\Omega|$.

Proof. The existence of a unique weak solution is deduced by applying Banach's fixed-point theorem to $u^n \mapsto u^{n+1}$, where $u^{n+1} \in L^2(0,T;X) \cap H^1(0,T;X^*)$ is the unique weak solution of the linear equation $u_t^{n+1} = \operatorname{div}(\mathbb{D}\nabla u^{n+1}) + f(u^n)$ according to [50, Thm. 3.1], cf. [14, Thm. 2 p. 500]. (Similar existence results can be found in e.g. [42, Thm. 1.2 p. 184] and [23, Thm. 3.3.3].) In the following three steps of the proof, we derive (2.8).

Step 1. Let u be the unique weak solution of (2.1P). Testing (2.3WF) with $\varphi = u$ and using $\frac{\mathrm{d}}{\mathrm{d}t}(u,u)_H = 2\langle u_t, u\rangle_{X^*,X}$ as well as (2.7C_1) gives

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_H^2 &= (-\mathbb{D}\nabla u, \nabla u)_H + (f(u),u)_H \\ &\le -\mu\|\nabla u\|_H^2 + \|C_1(1+2|u|^2)\|_{L^1(\Omega)} \le c_1(1+\|u\|_H^2),\end{aligned}$$

where $c_1 = c_1(C_1,|\Omega|)$ and C_1 is from (2.7C_1). Applying Gronwall's lemma yields $\|u(t)\|_H^2 + 1 \le (\|u_0\|_H^2+1)\exp(2Tc_1)$ for all $t \in [0,T]$. Hence, there exists $c_2 = c_2(\|u_0\|_H, T, L, C_\infty, |\Omega|) \ge 0$, independent of t, such that $\|u\|_{C^0([0,T];H)} \le c_2$.

Step 2. Again testing (2.3WF) with $\varphi = u$ and integrating over $(0,t)$ for $0 < t \le T$, yields

$$\begin{aligned}\mu\int_0^t \|\nabla u\|_H^2 \,\mathrm{d}\tau &\le \int_0^t (\mathbb{D}\nabla u, \nabla u)_H \,\mathrm{d}\tau = \int_0^t \langle f(u) - u_\tau, u\rangle_{X^*,X}\,\mathrm{d}\tau \\ &\le \int_0^t -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}\|u(\tau)\|_H^2 + \|C_1(1+2|u|^2)\|_{L^1(\Omega)}^2\,\mathrm{d}\tau \\ &\le \tfrac{1}{2}\left(\|u(0)\|_H^2 - \|u(t)\|_H^2\right) + C(c_2).\end{aligned}$$

Since $t \in (0,T]$ was chosen arbitrarily, we obtain $\sqrt{\mu}\|\nabla u\|_{L^2(0,T;H)} \le c_3 < \infty$ and c_3 depends on the same set of parameters as c_2.

Step 3. Analogously to Step 2, we obtain by applying Hölder's inequality, (2.6D_∞), and (2.7C_1):

$$\begin{aligned}\|u_t\|_{L^2(0,T;X^*)}^2 &= \int_0^T \left(\sup_{\|\varphi\|_X=1} -(\mathbb{D}\nabla u, \nabla\varphi)_H + \langle f(u),\varphi\rangle_{X^*,X}\,\mathrm{d}x\right)^2 \mathrm{d}t \\ &\le \int_0^T \left(\sup_{\|\varphi\|_X=1} D_\infty\|\nabla u\|_H\|\varphi\|_X + C_1^2(|\Omega| + \|u\|_H)\|\varphi\|_X\right)^2 \mathrm{d}t \\ &\le T\left(D_\infty \tfrac{1}{\sqrt{\mu}} c_3 + C(c_2)\right)^2.\end{aligned}$$

Hence Step 1–3 imply the existence of a constant C_a, depending on $\|u_0\|_H, T, L, C_\infty$, and $|\Omega|$, such that (2.8) holds true. □

We complete Section 2 with Proposition 2.2 that gives improved time-regularity for weak solutions u of (2.1P), i.e. $u_t \in C^0([0,T];H) \subsetneq L^2(0,T;X^*)$. This is motivated by the fact that the folding and unfolding operators, defined in Section

3.2, are only well-defined for integrable functions. Finally Example 2.3 gives an exemplary system satisfying all the assumptions presented in this section.

Proposition 2.2 (Improved time-regularity)**.** *Let the assumptions of Theorem 2.1 hold true. We assume the additional regularity for the initial value:*

$$\mathcal{A}u_0 \in H. \tag{2.9}$$

Then we have for all weak solutions u to (2.1P) that $u \in H^1(0,T;X) \cap H^2(0,T;X^*)$ and

$$\|u\|_{C^1([0,T];H)} + \sqrt{\mu}\|\nabla u\|_{H^1(0,T;H)} + \|u_t\|_{H^1(0,T;X^*)} \leq C_a^* \left(1 + \tfrac{D_\infty}{\sqrt{\mu}}\right) \tag{2.10}$$

with the constant $C_a^* \geq 0$ depending on C_a from (2.8) and $\|\mathcal{A}u_0\|_H$.

Proof. We follow the idea of the proof to [50, Thm. 3.2]. By setting $w = u_t$ and recalling the definition of $\mathcal{A}$ from (2.1P), we obtain $w_t = u_{tt} = (\mathrm{div}(\mathbb{D}\nabla u) + f(u))_t = \mathrm{div}(\mathbb{D}\nabla w) + f_t(u) + \mathrm{D}f(u)\cdot w$. This leads to a reaction-diffusion equation of the type (2.1P), i.e.

$$w_t = \mathrm{div}(\mathbb{D}\nabla w) + \tilde{f}(w) \quad \text{in } \Omega_T. \tag{2.11}$$

Here $\tilde{f}(t,x,A) := f_t(t,x,u(t,x)) + \mathrm{D}f(t,x,u(t,x))A$ and $\mathrm{D}f$ denotes the derivative of f w.r.t. A with $|\mathrm{D}f(x,A)| \leq L$ for all $(t,x,A) \in \Omega_T \times \mathbb{R}^m$. It holds further that $t \mapsto \tilde{f}(t,x,w(t,x)) = f_t(t,x,u(t,x)) + \mathrm{D}f(x,u(t,x))w(t,x) \in L^2(0,T;H)$ for all $u, w \in L^2(0,T;H)$ due to the continuity of f_t and $\mathrm{D}f$ (in the third argument). With (2.9), the initial value for w in (2.11) satisfies

$$w(0) = u_t(0) = \mathrm{div}(\mathbb{D}\nabla u_0) + f(u_0) \in H.$$

Furthermore we have for a.e. $(t,x) \in \Omega_T$ that $|\tilde{f}(t,x,A) - \tilde{f}(t,x,B)| \leq L|A-B|$ for all $(A,B) \in \mathbb{R}^m \times \mathbb{R}^m$ and $\tilde{f}(t,x,0) = 0$. Regardless that $\tilde{f}$ is not differentiable with respect to time, it satisfies (2.5L) and (2.6C_∞) and hence the necessary assumptions of Theorem 2.1. Therefore $w \in L^2(0,T;X) \cap H^1(0,T;X^*)$ is the unique weak solution of (2.11) and hence $u \in H^1(0,T;X) \cap H^2(0,T;X^*) \subset C^1([0,T];H)$. □

The additional assumption (2.9) seems to be quite restrictive, on the initial value u_0 and on the diffusion tensor $\mathbb{D}$, but actually $\mathbb{D}$ can be as general as in Theorem 2.1. We interpret (2.9) as a restriction on the choice of the initial value u_0, while $\mathbb{D}$ is possibly discontinuous. Indeed for $\mathbb{D} \in \mathbb{M}(\Omega, \mu, D_\infty)$ and arbitrary $g \in H$, we can solve the static equation

$$\mathrm{div}(\mathbb{D}\nabla u_0) - u_0 = g \quad \text{in } \Omega \tag{2.12}$$

and we obtain (by the Lax-Milgram lemma) a unique weak solution $u_0 \in X$. In particular, we have $\mathcal{A}u_0 = \mathrm{div}(\mathbb{D}\nabla u_0) \in H$.

We emphasize that the improved time-regularity, and therefore the more restrictive assumptions on f (differentiability) and u_0 (as in (2.12)), are only needed for technical reasons, i.e. the application of $\mathcal{T}_\varepsilon$. In particular, we expect that Theorem 4.1 can be proved without this improved time-regularity, but then the proof would become more technical by e.g. using time-discretized approximations of u_t.

Assuming further structural assumptions on f and u_0, one can prove even $L^\infty(\Omega)$–estimates for the solutions, cf. e.g. [21, Thm. 4.2], [3, Lem. 1], [40, Lem. 3.1], and [45, Lem. 1.1]. Such boundedness is meaningful, when u_i denote chemical concentrations. In particular, it justifies the modification of the nonlinear reaction term outside a large ball and hence, the assumption of global Lipschitz continuity can be fulfilled easily.

Example 2.3 (A system with quadratic nonlinearity)**.** We consider a system with two species X_u and X_v, with densities $u, v \geq 0$ interacting through one reaction of the type $X_u \rightleftharpoons 2X_v$. Normalizing the densities suitably, the mass-action law leads to the system

$$u_t = \delta_u \Delta u + k\,(v^2 - u), \quad v_t = \delta_v \Delta v + 2k\,(u - v^2), \tag{2.13}$$

where $\delta_u, \delta_v > 0$ and the reaction coefficient k is given via $k(u,v) = \dfrac{k_0}{1 + \alpha u + \beta v}$. The numerator $k_0 > 0$ denotes the empirical reaction rate and the denominator $1 + \alpha u + \beta v$, for $0 < \alpha, \beta \ll 1$, leads to partial saturation of the reaction for large values of $u, v > 0$. The nonlinearity $f(u,v) = k(u,v) \begin{pmatrix} v^2 - u \\ 2(u - v^2) \end{pmatrix}$ is differentiable and globally Lipschitz continuous with constant $L = O(\max\{\frac{\alpha}{\beta^2}, \frac{1}{\beta}\})$. Hence f satisfies the assumptions (2.5L)–(2.6C_∞). In many applications (cf. e.g. [32, 33] for general reaction-diffusion systems based on the mass-action law) the reaction terms are given by polynomials and choosing suitable prefactors one obtains globally Lipschitz continuous $f \in \mathbb{F}(\Omega, L, C_\infty)$, e.g. the Shockley-Read-Hall term in semiconductor equations [29, Eq. (3.1.9)] or in Michaelis-Menten kinetics for enzymatic catalysis [38, pp. 175].

3. Two-scale convergence.

In the introduction, the original model (1.1$\mathrm{P}_\varepsilon^{\mathrm{cp}}$) is formulated on one scale, i.e. $x \in \Omega$, while the limit model (1.3$\mathrm{P}_0^{\mathrm{cp}}$) is defined on the two-scale space $(x,y) \in \Omega \times \mathcal{Y}$. Here the microscopic variable y captures periodic oscillations in x/ε and x denotes the macroscopic variable. In order to describe the convergence from (1.1$\mathrm{P}_\varepsilon^{\mathrm{cp}}$) to (1.3$\mathrm{P}_0^{\mathrm{cp}}$), we introduce the concept of two-scale convergence, which is designed for problems with underlying periodic microstructure. But before giving the definition (in Section 3.3), we introduce the concept of the *periodicity cell* $\mathcal{Y}$, the decomposition in macro- and microscopic scale (in Section 3.1) and the unfolding and folding operators $\mathcal{T}_\varepsilon$ and $\mathcal{F}_\varepsilon$ (in Section 3.2).

3.1. Microstructure and the periodicity cell $\mathcal{Y}$.

Following [7, Sec. 2.1], let $\Omega \subset \mathbb{R}^d$ be a bounded domain and let $Y \subset \mathbb{R}^d$ denote the *unit cell*. Here and afterwards, we set $Y = [-\frac{1}{2}, \frac{1}{2})^d$, but more general choices for Y are possible, see e.g. [35, Sec. 2.1], so that $\mathbb{R}^d$ is the disjoint union of translated cells $\lambda + Y$, where $\lambda \in \mathbb{Z}^d$. Furthermore, we distinguish the unit cell Y from the periodicity cell $\mathcal{Y}$, which is obtained by identifying the opposite faces of $\overline{Y}$, i.e. the torus

$$\mathcal{Y} := \mathbb{R}^d/_{\mathbb{Z}^d}.$$

But, in notation, we will not distinguish between elements of the unit cell $y \in Y$ and the ones of the periodicity cell $y \in \mathcal{Y}$. Using the mappings $[\cdot]_Y : \mathbb{R}^d \to \mathbb{Z}^d$ and $\{\cdot\}_Y : \mathbb{R}^d \to Y$ defined via the relation $x = [x]_Y + \{x\}_Y$, each point $x \in \mathbb{R}^d$ is uniquely decomposed into an element of the unit cell $\{x\}_Y \in Y$ and a lattice point $[x]_Y \in \mathbb{Z}^d$. A function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ is called Y*-periodic*, if $f(x) = f(\{x\}_Y)$ for a.a. $x \in \mathbb{R}^d$. Then we can identify every periodic function f with a function $\widetilde{f}$ on $\mathcal{Y}$. Whereas $L^p(\mathcal{Y})$ and $L^p(Y)$ can be identified, $H^1(\mathcal{Y}) = H^1_{\mathrm{per}}(Y)$ is a closed subspace of $H^1(Y)$.

For our problem (1.1$\mathrm{P}_\varepsilon^{\mathrm{cp}}$), we introduce the small length-scale parameter $\varepsilon > 0$ and we use the abbreviation $\mathcal{N}_\varepsilon(x) := \varepsilon \left[\frac{x}{\varepsilon}\right]_Y$ for the nodes of the microscopic cells $\{\varepsilon(\lambda + Y) \,|\, \lambda \in \mathbb{Z}^d\}$, which describe the macroscopic scale. The microscopic scale is given by $y = \{\frac{x}{\varepsilon}\}_Y \in Y$ so that we obtain for all $x \in \mathbb{R}^d$ the decomposition $x = \mathcal{N}_\varepsilon(x) + \varepsilon y$. Since the domain Ω is bounded and not the whole $\mathbb{R}^d$, we have to treat the cells close to the boundary $\partial\Omega$ with care so that cells intersecting $\partial\Omega$ are sorted out for each $\varepsilon > 0$ fixed. We set $\hat{\Omega}_\varepsilon := \mathrm{int}\big(\bigcup_{\lambda\in Z} \varepsilon(\lambda + Y)\big)$ with $Z := \{\lambda \in \mathbb{Z}^d \,|\, \varepsilon(\lambda + Y) \subset \Omega\}$. Hence $\hat{\Omega}_\varepsilon$ denotes (the interior of) the union of all microscopic cells $\varepsilon(\lambda + Y)$ strictly contained in Ω. For bounded domains Ω with Lipschitz boundary Γ, we have by [22, Eq. (2.3)] that

$$\mathrm{vol}(\Omega\backslash\hat{\Omega}_\varepsilon) \to 0. \tag{3.1}$$

3.2. **Folding and periodic unfolding operators.** Two-scale convergence is suited to describe convergences on different scales, namely the macroscopic scale, represented by $x \in \Omega$, and the microscopic scale for $y \in \mathcal{Y}$. Therefore the notion of a suitable embedding of the function space $L^2(\Omega)$ into the two-scale space $L^2(\mathbb{R}^d \times \mathcal{Y})$ is desirable in order to find a "natural" definition of two-scale convergence. Here, we call such a mapping *periodic unfolding operator*. Vice versa, for any two-scale function U defined on $\Omega \times \mathcal{Y}$ we seek a one-scale dependent u^ε defined on Ω, and we call a corresponding mapping from the two-scale space $L^2(\mathbb{R}^d \times \mathcal{Y})$ into $L^2(\Omega)$ *folding operator*.

Following [6, 7, 35], the *periodic unfolding operator* $\mathcal{T}_\varepsilon : L^2(\Omega) \to L^2(\mathbb{R}^d \times \mathcal{Y})$ is defined via

$$(\mathcal{T}_\varepsilon\, u)(x, y) := u^{\mathrm{ex}}(\mathcal{N}_\varepsilon(x) + \varepsilon y), \tag{3.2}$$

where $u^{\mathrm{ex}} \in L^2(\mathbb{R}^d)$ is obtained from u by extension with 0 outside of Ω. By definition, we have immediately the product rule

$$u \in L^2(\Omega),\ v \in L^2(\Omega) \quad\Longrightarrow\quad \mathcal{T}_\varepsilon(uv) = (\mathcal{T}_\varepsilon\, u)(\mathcal{T}_\varepsilon\, v) \in L^1(\mathbb{R}^d \times \mathcal{Y}). \tag{3.3}$$

Moreover, we obtain (see [10, p. 121]) the crucial identity

$$\int_\Omega u \,\mathrm{d}x = \int_{\mathbb{R}^d\times\mathcal{Y}} \mathcal{T}_\varepsilon\, u \,\mathrm{d}x\,\mathrm{d}y \quad \text{for all } u \in L^1(\Omega). \tag{3.4}$$

With $[\Omega\times\mathcal{Y}]_\varepsilon := \{\, (x, y) \in \mathbb{R}^d\times\mathcal{Y} \mid \mathcal{N}_\varepsilon(x)+\varepsilon y \in \Omega \,\}$ we have $\mathrm{supp}(\mathcal{T}_\varepsilon\, u) \subseteq \overline{[\Omega \times \mathcal{Y}]_\varepsilon}$, i.e. in general the support of a two-scale function $\mathcal{T}_\varepsilon\, u$ is not contained in $\Omega \times \mathcal{Y}$. For a proper definition of the reverse operation taking care of the overhanging supports, we follow the construction of the folding operator in [35], which involves the characteristic functions $\mathbb{1}_\Omega$ and $\mathbb{1}_\varepsilon := \mathcal{T}_\varepsilon\, \mathbb{1}_\Omega$ of Ω and $[\Omega\times\mathcal{Y}]_\varepsilon$, respectively. The *folding operator* $\mathcal{F}_\varepsilon : L^2(\mathbb{R}^d \times \mathcal{Y}) \to L^2(\Omega)$ is defined via

$$(\mathcal{F}_\varepsilon\, U)(x) := \left(\frac{1}{\varepsilon^d}\int_{\mathcal{N}_\varepsilon(x)+\varepsilon Y} \mathbb{1}_\varepsilon(\xi, \{\tfrac{x}{\varepsilon}\}_Y) \cdot U(\xi, \{\tfrac{x}{\varepsilon}\}_Y)\,\mathrm{d}\xi\right)\Bigg|_\Omega, \tag{3.5}$$

whereby $\mathrm{vol}(\mathcal{N}_\varepsilon(x) + \varepsilon Y) = \varepsilon^d$. We will use several properties of $\mathcal{T}_\varepsilon$ and $\mathcal{F}_\varepsilon$, see e.g. [35, Prop. 2.1]:

Proposition 3.1. *For all $\varepsilon > 0$ we have the following properties:*

(a) $\|\mathcal{T}_\varepsilon\, u\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} = \|u\|_{L^2(\Omega)}$ *and* $\mathrm{supp}(\mathcal{T}_\varepsilon\, u) \subset \overline{[\Omega \times \mathcal{Y}]_\varepsilon}$ *for all* $u \in L^2(\Omega)$.
(b) $\|\mathcal{F}_\varepsilon\, U\|_{L^2(\Omega)} \le \|U\|_{L^2(\mathbb{R}^d\times\mathcal{Y})}$ *for all* $U \in L^2(\mathbb{R}^d \times \mathcal{Y})$.
(c) $\mathcal{F}_\varepsilon \circ \mathcal{T}_\varepsilon = id_{L^2(\Omega)}$.
(d) $\mathcal{F}_\varepsilon$ *is the adjoint of* $\mathcal{T}_\varepsilon$, *i.e.* $\mathcal{F}_\varepsilon = \mathcal{T}_\varepsilon{}'$.

The following result states in which sense the periodic unfolding operator $\mathcal{T}_\varepsilon$ is compatible with differentiation and composition of functions.

Theorem 3.2 (Properties of $\mathcal{T}_\varepsilon$)**.**
(a) For $u \in H^1(\Omega)$ we have $\mathcal{T}_\varepsilon u \in L^2(\mathbb{R}^d; H^1(Y))$ and $\mathcal{T}_\varepsilon(\varepsilon \nabla u) = \nabla_y(\mathcal{T}_\varepsilon u)$.
(b) For $f \in \mathbb{F}(\Omega, L, C_0)$ and $u \in L^2(\Omega)$ we have $\mathcal{T}_\varepsilon[f(u)] = \mathcal{T}_\varepsilon f(\mathcal{T}_\varepsilon u)$.

Proof. For part (a), we refer to [10, Thm. 5.1]. Part (b) follows from (3.2), i.e.

$$\begin{aligned}\mathcal{T}_\varepsilon\left[f(u)\right](x,y) &= \begin{cases} f(\mathcal{N}_\varepsilon(x)+\varepsilon y, u(\mathcal{N}_\varepsilon(x)+\varepsilon y)), & \text{if } (x,y) \in [\Omega \times \mathcal{Y}]_\varepsilon \\ 0, & \text{if } (x,y) \in (\mathbb{R}^d \times \mathcal{Y}) \setminus [\Omega \times \mathcal{Y}]_\varepsilon \end{cases} \\ &= f(\mathcal{N}_\varepsilon(x)+\varepsilon y, u(\mathcal{N}_\varepsilon(x)+\varepsilon y)^{\mathrm{ex}})^{\mathrm{ex}} = \mathcal{T}_\varepsilon f(\mathcal{T}_\varepsilon u)(x,y).\end{aligned}$$

□

3.3. **Weak and strong two-scale convergence.** We are now in the position to give the definition of weak and strong two-scale convergence following again [35]. The notion of two-scale convergence was first introduced in [41] and coincides for bounded sequences with Definition 3.3(a), here below, and a more detailed comparison of the different definitions is given in [35, Sec. 2.3]. Since the construction of the periodic unfolding operator was quite technical, the definition of weak and strong two-scale convergence can now be stated easily:

Definition 3.3 (Weak and strong two-scale convergence)**.** For $(u^\varepsilon)_\varepsilon \subset L^2(\Omega)$
(a) we say that u^ε weakly two-scale converges to U in $L^2(\Omega \times \mathcal{Y})$ and we write "$u^\varepsilon \xrightharpoonup{2\mathrm{w}} U$ in $L^2(\Omega \times \mathcal{Y})$", if $\mathcal{T}_\varepsilon u^\varepsilon \rightharpoonup U^{\mathrm{ex}}$ weakly in $L^2(\mathbb{R}^d \times \mathcal{Y})$;
(b) we say that u^ε strongly two-scale converges to U in $L^2(\Omega \times \mathcal{Y})$ and we write "$u^\varepsilon \xrightarrow{2\mathrm{s}} U$ in $L^2(\Omega \times \mathcal{Y})$", if $\mathcal{T}_\varepsilon u^\varepsilon \to U^{\mathrm{ex}}$ strongly in $L^2(\mathbb{R}^d \times \mathcal{Y})$.

Note that the weak and strong convergence is asked to occur in $L^2(\mathbb{R}^d \times \mathcal{Y})$ and not in $L^2(\Omega \times \mathcal{Y})$. Otherwise a slightly different notion of convergence is generated, see e.g. [35, Ex. 2.3]. The unfolding operator $\mathcal{T}_\varepsilon : L^2(\Omega) \to L^2(\mathbb{R}^d \times \mathcal{Y})$ is defined for the class of Lebesgue-integrable functions, where boundary values play no role, so that in particular $L^2(\mathbb{R}^d \times \mathcal{Y}) = L^2(\mathbb{R}^d \times Y)$. In view of the *$\mathcal{T}_\varepsilon$-property of recovered periodicity* (1.7), we carefully distinguish the spaces $H^1(Y)$ and $H^1(\mathcal{Y}) = H^1_{\mathrm{per}}(Y)$, where the latter one is a closed subspace of $H^1(Y)$. We now collect various properties of two-scale convergence.

Proposition 3.4. *For all $\varepsilon > 0$, we have the following properties:*
(a) $u^\varepsilon \xrightharpoonup{2\mathrm{w}} U$ in $L^2(\Omega \times \mathcal{Y}) \implies \|u^\varepsilon\|_{L^2(\Omega)}$ is bounded for all $\varepsilon > 0$.
(b) $u^\varepsilon \xrightharpoonup{2\mathrm{w}} U$ in $L^2(\Omega \times \mathcal{Y})$ and $v^\varepsilon \xrightarrow{2\mathrm{s}} V$ in $L^2(\Omega \times \mathcal{Y}) \implies (u^\varepsilon, v^\varepsilon)_{L^2(\Omega)} \to (U, V)_{L^2(\Omega \times \mathcal{Y})}$.
(c) For all $U \in L^2(\Omega \times \mathcal{Y})$ there exists a sequence $(u^\varepsilon)_\varepsilon$ so that $u^\varepsilon \xrightarrow{2\mathrm{s}} U$ in $L^2(\Omega \times \mathcal{Y})$. (for example $u^\varepsilon = \mathcal{F}_\varepsilon U^{\mathrm{ex}}$)
(d) $u^\varepsilon \to u$ in $L^2(\Omega) \implies u^\varepsilon \xrightarrow{2\mathrm{s}} u$ in $L^2(\Omega \times \mathcal{Y})$.
(e) $u^\varepsilon \xrightharpoonup{2\mathrm{w}} U$ in $L^2(\Omega \times \mathcal{Y}) \implies u^\varepsilon \rightharpoonup u$ in $L^2(\Omega)$, where $u(x) = \int_{\mathcal{Y}} U(x,y)\,\mathrm{d}y$.

We refer to [35, Prop. 2.4] for a proof of (a)–(d) and to [10, Thm. 3.3] for (e). The following theorem states the fundamental results for two-scale convergence, in particular parts (b) and (c) are crucial for the proofs of Theorem 4.1 and Theorem 5.1. We define

$$H^1_{\mathrm{av}}(\mathcal{Y}) := \left\{ u \in H^1(\mathcal{Y}) \mid \textstyle\int_{\mathcal{Y}} u(y)\,\mathrm{d}y = 0 \right\}..$$

Theorem 3.5 (Compactness)**.** *Let $(u^\varepsilon)_\varepsilon$ be a sequence of functions.*

(a) If $u^\varepsilon \in L^2(\Omega)$ and $\|u^\varepsilon\|_{L^2(\Omega)} \le C$, then there exists $U \in L^2(\Omega\times\mathcal{Y})$ and a subsequence ε' of ε such that it holds $u^{\varepsilon'} \xrightarrow{2w} U$ in $L^2(\Omega\times\mathcal{Y})$.

(b) If $u^\varepsilon \in H^1(\Omega)$ and $\|u^\varepsilon\|_{L^2(\Omega)} + \varepsilon\|\nabla u^\varepsilon\|_{L^2(\Omega)} \le C$, then there exists $U \in L^2(\Omega;H^1(\mathcal{Y}))$ and a subsequence ε' of ε such that $u^{\varepsilon'} \xrightarrow{2w} U$ & $\varepsilon'\nabla u^{\varepsilon'} \xrightarrow{2w} \nabla_y U$ in $L^2(\Omega\times\mathcal{Y})$.

(c) If $u^\varepsilon \in H^1(\Omega)$ and $\|u^\varepsilon\|_{H^1(\Omega)} \le C$, then there exists $u \in H^1(\Omega)$, a two-scale function $U_1 \in L^2(\Omega;H^1_{av}(\mathcal{Y}))$, and a subsequence ε' of ε such that $u^{\varepsilon'} \rightharpoonup u$ in $H^1(\Omega)$ and $\nabla u^{\varepsilon'} \xrightarrow{2w} \nabla u + \nabla_y U_1$ in $L^2(\Omega\times\mathcal{Y})$.

Proof. For the proof of (a), we refer to [41], alternatively one can apply Prop. 3.1(a) and Banach's selection principle. Items (b) and (c) are shown in e.g. [1, Prop. 1.14] or [10, Thm. 5.2, Thm. 5.4]. For other scalings such as ε^γ with $0 \le \gamma < \infty$, we refer to [44, Thm. 3.4]. □

We finish this subsection by stating two results, needed in the proof of Theorem 4.1, concerning the multiplication and composition of sequences in $L^2(\Omega\times\mathcal{Y})$.

Lemma 3.6 (Multiplication and composition of sequences in $L^2(\Omega\times\mathcal{Y})$)**.** *Let $\varepsilon > 0$.*

(a) Let $(U^\varepsilon)_\varepsilon \subset L^2(\Omega\times\mathcal{Y})$ with $U^\varepsilon \to U$ in $L^2(\Omega\times\mathcal{Y})$ and $(M^\varepsilon)_\varepsilon \subset L^\infty(\Omega\times\mathcal{Y})$ such that $\|M^\varepsilon\|_{L^\infty(\Omega\times\mathcal{Y})} \le C$ for some constant $C > 0$ and $M^\varepsilon(x,y) \to M(x,y)$ for almost every $(x,y) \in \Omega\times\mathcal{Y}$. Then $M^\varepsilon U^\varepsilon \to MU$ in $L^2(\Omega\times\mathcal{Y})$.

(b) Let $f^\varepsilon \in \mathbb{F}(\Omega, L, C_0)$, $F \in \mathbb{F}(\Omega\times\mathcal{Y}, L, C_0)$ and keep $t \in (0,T)$ fixed. If for all vectors $A \in \mathbb{R}^m$ it is $f^\varepsilon(t,A) \xrightarrow{2s} F(t,A)$ in $L^2(\Omega\times\mathcal{Y})$, then for all $U \in L^2(\mathbb{R}^d\times\mathcal{Y})$ we have $\mathcal{T}_\varepsilon f^\varepsilon(t,U) \to F^{\mathrm{ex}}(t,U)$ in $L^2(\mathbb{R}^d\times\mathcal{Y})$.

Proof. Ad (a): Extracting from $(U^\varepsilon)_\varepsilon$ a pointwise convergent subsequence, we find that $M^\varepsilon U^\varepsilon \to MU$ pointwise a.e. in $\Omega\times\mathcal{Y}$ for this subsequence. Moreover, since $|M^\varepsilon U^\varepsilon| \le C|U^\varepsilon|$ a.e. in $\Omega\times\mathcal{Y}$ by assumption, the sequence $(CU^\varepsilon)_\varepsilon$ serves as an L^2–convergent majorant. Thus, Pratt's theorem, see [12, Thm. 5.1 p. 260], a variant of the dominated convergence theorem, yields the strong L^2–convergence of the subsequence. Arguing by contradiction for a different subsequence and by the uniqueness of the limit we conclude the convergence of the *whole* sequence.

Ad (b): For shorter notation we omit indicating the t-dependence of the functions. We approximate $U \in L^2(\mathbb{R}^d\times\mathcal{Y})$ with a sequence of integrable step functions $\overline{U}_n = \sum_{i=1}^n \mathbb{1}_{U_i}\cdot A_i$, where $A_i \in \mathbb{R}^m$ and $\overline{\Omega}\times\mathcal{Y} \subset \bigcup_{i=1}^n U_i$. Hence $\overline{U}_n \to U^{\mathrm{ex}}$ in $L^2(\mathbb{R}^d\times\mathcal{Y})$ and it follows by assumption that $\mathcal{T}_\varepsilon f^\varepsilon(\overline{U}_n) = \sum_{i=1}^n \mathbb{1}_{U_i}\cdot\mathcal{T}_\varepsilon f^\varepsilon(A_i) \xrightarrow{\varepsilon\to 0} \sum_{i=1}^n \mathbb{1}_{U_i}\cdot F^{\mathrm{ex}}(A_i) = F^{\mathrm{ex}}(\overline{U}_n)$ in $L^p(\mathbb{R}^d\times\mathcal{Y})$. Exploiting moreover (2.5L), we obtain $\|\mathcal{T}_\varepsilon f^\varepsilon(U) - F^{\mathrm{ex}}(U)\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} \to 0$ by introducing suitable nils, i.e. $\mathcal{T}_\varepsilon f^\varepsilon(U) - F^{\mathrm{ex}}(U) = [\mathcal{T}_\varepsilon f^\varepsilon(U) - \mathcal{T}_\varepsilon f^\varepsilon(\overline{U}_n)] + [\mathcal{T}_\varepsilon f^\varepsilon(\overline{U}_n) - F^{\mathrm{ex}}(\overline{U}_n)] + [F^{\mathrm{ex}}(\overline{U}_n) - F^{\mathrm{ex}}(U)]$. □

3.4. **Gradient folding and two-scale convergence of Sobolev functions.** Even for smooth functions $U:\Omega\times\mathcal{Y}\to\mathbb{R}$ the folded function $\mathcal{F}_\varepsilon U$ is only piecewise constant in x, hence $\nabla(\mathcal{F}_\varepsilon U)$ cannot be determined in the classical sense. Therefore we now define a so-called *gradient folding operator* $\mathcal{G}_\varepsilon$, which assigns to each differentiable two-scale function $U \in H^1(\Omega\times\mathcal{Y})$ a one-scale function $u^\varepsilon \in H^1(\Omega)$. The definition of the above mentioned gradient folding operator $\mathcal{G}_\varepsilon$ follows [22] for $\gamma = 1$. There, the operator $\mathcal{G}_\varepsilon$ is constructed via $\mathcal{T}_\varepsilon$ and various projections, but then it is shown that $\mathcal{G}_\varepsilon$ is uniquely characterized by solving a linear elliptic PDE, see [22, Prop. 2.11] based on [53, Thm. 6.1] and [35, Prop. 2.10].

Definition 3.7 (Gradient folding)**.** The gradient folding operator $\mathcal{G}_\varepsilon : L^2(\Omega; H^1(\mathcal{Y})) \to H^1(\Omega)$ maps a two-scale function $U \in L^2(\Omega; H^1(\mathcal{Y}))$ to $u^\varepsilon := \mathcal{G}_\varepsilon U$, where $u^\varepsilon \in H^1(\Omega)$ is the unique weak solution of the elliptic problem

$$\int_\Omega (u^\varepsilon - \mathcal{F}_\varepsilon U^{\mathrm{ex}})\cdot\varphi + (\varepsilon\nabla u^\varepsilon - \mathcal{F}_\varepsilon[\nabla_y U]^{\mathrm{ex}}) : \varepsilon\nabla\varphi \,\mathrm{d}x = 0 \quad \text{for all } \varphi \in H^1(\Omega). \tag{3.6}$$

While $\mathcal{F}_\varepsilon : L^2(\mathbb{R}^d \times \mathcal{Y}) \to L^2(\Omega)$, we have $\mathcal{G}_\varepsilon : L^2(\Omega; H^1(\mathcal{Y})) \to H^1(\Omega)$. Thus the domains of the two operators differ not only with respect to the regularity of the admissible functions, but also with respect to the underlying domains for the space variable x, i.e. $x \in \mathbb{R}^d$ versus $x \in \Omega$. However, since both operators require L^2–regularity in x only, extending $U \in L^2(\Omega; H^1(\mathcal{Y}))$ by 0 outside of Ω yields $U^{\mathrm{ex}} \in L^2(\mathbb{R}^d; H^1(\mathcal{Y}))$. Thus, $\mathcal{F}_\varepsilon U^{\mathrm{ex}}$ indeed is well-defined in (3.6). In particular, $\mathcal{F}_\varepsilon U^{\mathrm{ex}}$ and $\mathcal{F}_\varepsilon[\nabla_y U]^{\mathrm{ex}} \in L^2(\Omega)$ can be understood as linear operators acting on U and moved, as inhomogeneities for the determination of u^ε, to the right-hand side of (3.6). Thus for $\varepsilon > 0$ fixed, the Lax-Milgram lemma yields the existence of a unique weak solution $u^\varepsilon \in H^1(\Omega)$, so that the gradient folding operator $\mathcal{G}_\varepsilon$ is indeed well-defined. Since (3.6) implies $\|\mathcal{G}_\varepsilon U\|_{L^2(\Omega)} + \varepsilon\|\nabla(\mathcal{G}_\varepsilon U)\|_{L^2(\Omega)} \le C$, Theorem 3.5(b) supplies the existence of a weakly two-scale convergent subsequence. However, for given $U \in L^2(\Omega; H^1(\mathcal{Y}))$ the gradient folding operator guarantees even *strong* two-scale convergence. Since $(\mathcal{G}_\varepsilon U)_\varepsilon \subset H^1(\Omega)$ *recovers* any function $U \in L^2(\Omega; H^1(\mathcal{Y}))$ via strong two-scale convergence, $\mathcal{G}_\varepsilon$ is also called *recovery operator* in [22, pp. 10-12].

Proposition 3.8 (Recovery property of $\mathcal{G}_\varepsilon$, [22, Prop. 2.11])**.** *For all two-scale functions $U \in L^2(\Omega; H^1(\mathcal{Y}))$, we have $\mathcal{G}_\varepsilon U \xrightarrow{2\mathrm{s}} U$ & $\varepsilon\nabla[\mathcal{G}_\varepsilon U] \xrightarrow{2\mathrm{s}} \nabla_y U$ in $L^2(\Omega \times \mathcal{Y})$.*

Later on, in the proof of Theorem 4.1, it will be essential to interchange differentiation and folding of two-scale functions $U \in L^2(\Omega; H^1(\mathcal{Y}))$. However, convenient commutation relations, such as $\mathcal{F}_\varepsilon(\nabla_y U^{\mathrm{ex}}) = \varepsilon\nabla(\mathcal{F}_\varepsilon U^{\mathrm{ex}})$ or $\mathcal{G}_\varepsilon(\nabla_y U) = \varepsilon\nabla(\mathcal{G}_\varepsilon U)$, cannot be expected, since $\mathcal{F}_\varepsilon U \notin H^1(\Omega)$ and $\nabla_y U \notin L^2(\Omega; H^1(\mathcal{Y}))$. Instead, we establish a kind of commutation between $\mathcal{F}_\varepsilon(\nabla_y U^{\mathrm{ex}})$ and $\varepsilon\nabla(\mathcal{G}_\varepsilon U)$. More precisely, the following result shows that $\mathcal{F}_\varepsilon U^{\mathrm{ex}}$ and $\mathcal{G}_\varepsilon U$ are comparable in the sense that their difference vanishes.

Theorem 3.9 (Comparison of $\mathcal{F}_\varepsilon$ and $\mathcal{G}_\varepsilon$)**.** *For all $U \in L^2(\Omega; H^1(\mathcal{Y}))$ we have*

$$\|\mathcal{F}_\varepsilon U^{\mathrm{ex}} - \mathcal{G}_\varepsilon U\|_{L^2(\Omega)} + \|\mathcal{F}_\varepsilon(\nabla_y U)^{\mathrm{ex}} - \varepsilon\nabla(\mathcal{G}_\varepsilon U)\|_{L^2(\Omega)} \to 0 \text{ for } \varepsilon \to 0.$$

Proof. Recalling the abbreviation $H = L^2(\Omega)$, we have

$$\begin{aligned}&\|\mathcal{F}_\varepsilon U^{\mathrm{ex}} - \mathcal{G}_\varepsilon U\|_H + \|\mathcal{F}_\varepsilon(\nabla_y U)^{\mathrm{ex}} - \varepsilon\nabla(\mathcal{G}_\varepsilon U)\|_H \\ &\le \|\mathcal{F}_\varepsilon U^{\mathrm{ex}} - U\|_H + \|U - \mathcal{G}_\varepsilon U\|_H + \|\mathcal{F}_\varepsilon(\nabla_y U)^{\mathrm{ex}} - \nabla_y U\|_H + \|\nabla_y U - \varepsilon\nabla(\mathcal{G}_\varepsilon U)\|_H,\end{aligned}$$

which converges to 0 by Proposition 3.4(d) for the terms involving $\mathcal{F}_\varepsilon$ and by Proposition 3.8 for the terms involving $\mathcal{G}_\varepsilon$. □

4. **Homogenization of the degenerating equation.** In this section we consider reaction-diffusion systems with nonlinear reactions and a diffusion term degenerating for $\varepsilon \to 0$,

$$\begin{aligned} v_t^\varepsilon &= \operatorname{div}(\varepsilon^2\mathbb{D}^\varepsilon\nabla v^\varepsilon) + f^\varepsilon(v^\varepsilon) && \text{in } \Omega_T, \\ 0 &= (\varepsilon^2\mathbb{D}^\varepsilon\nabla v^\varepsilon)\cdot\vec{n} && \text{on } \Gamma_T, \\ v^\varepsilon(0) &= v_0^\varepsilon && \text{in } \Omega. \end{aligned} \tag{4.1P$_\varepsilon$}$$

Relying on the fruits of Section 2.2, we rigorously derive a homogenization result for $\varepsilon \to 0$ in (4.1P$_\varepsilon$), stating the existence of a uniquely determined effective equation given by (4.4P$_0$) below. In Section 4.3, we prove that the weak solutions of (4.1P$_\varepsilon$) converge in the two-scale sense to the weak solution of (4.4P$_0$).

We will assume that $\mathbb{D}^\varepsilon \rightsquigarrow \mathbb{D}$ and $f^\varepsilon \rightsquigarrow F$ in a suitable manner, specified in assumption (4.9A$_{\varepsilon\to0}$) in Section 4.1. In view of the convergences $v^\varepsilon \xrightarrow{2\text{w}} V$ and $\varepsilon\nabla v^\varepsilon \xrightarrow{2\text{w}} \nabla_y V$ in $L^2(\Omega\times\mathcal{Y})$, see Theorem 3.5(b), we formally expect a result of the following type:

$$\begin{array}{ccc} \int_\Omega v_t^\varepsilon \cdot \varphi \,\mathrm{d}x & = & \int_\Omega -\mathbb{D}^\varepsilon \varepsilon\nabla v^\varepsilon : \varepsilon\nabla\varphi + f^\varepsilon(v^\varepsilon)\cdot\varphi\,\mathrm{d}x \quad \text{for all } \varphi \in X_\varepsilon \\ \downarrow & & \downarrow \text{ for } \varepsilon \to 0 \\ \int_{\Omega\times\mathcal{Y}} V_t \cdot \Phi \,\mathrm{d}x\,\mathrm{d}y & = & \int_{\Omega\times\mathcal{Y}} -\mathbb{D}\nabla_y V : \nabla_y\Phi + F(V)\cdot\Phi\,\mathrm{d}x\,\mathrm{d}y \quad \text{for all } \Phi \in \mathbb{X}. \end{array} \tag{4.2}$$

To deduce the convergence of the weak forms (4.2), we have to cope with the fact that $(v^\varepsilon)_\varepsilon$ converges a priori only weakly in the two-scale sense and therefore the passage $f^\varepsilon(v^\varepsilon) \rightsquigarrow F(V)$ is not straight forward, because f^ε and F are in general nonlinear. If we had the *strong* two-scale convergence of the sequence of solutions $(v^\varepsilon)_\varepsilon$, then $f^\varepsilon(v^\varepsilon) \rightsquigarrow F(V)$ would follow easily. For the special case of f^ε being the gradient of a λ-convex potential ϕ, a rigorous convergence result of the type (4.2) was deduced in [24, Prop. 12] via methods of convex analysis. In contrast to this, our approach to verify convergence (4.2) (in Theorem 4.1), indeed is to show that the sequence of solutions $(v^\varepsilon)_\varepsilon \subset H^1(0,T;H^1(\Omega))$ converges even *strongly* in the two-scale sense to some limit $V \in H^1(0,T;L^2(\Omega;H^1(\mathcal{Y})))$, more precisely:

$$\begin{array}{l} \text{uniformly for all } t\in[0,T]:\ v^\varepsilon(t) \xrightarrow{2\text{s}} V(t) \text{ in } L^2(\Omega\times\mathcal{Y}), \\ \text{i.e. } \max_{0\le t\le T} \| \mathcal{T}_\varepsilon\, v^\varepsilon(t) - V^{\text{ex}}(t)\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} \to 0, \end{array} \tag{4.3a}$$

$$\begin{array}{l} \text{pointwise for all } t\in[0,T]:\ \varepsilon\nabla v^\varepsilon(t) \xrightarrow{2\text{s}} \nabla_y V(t) \text{ in } L^2(\Omega\times\mathcal{Y}) \\ \text{and } \varepsilon\nabla v^\varepsilon \xrightarrow{2\text{s}} \nabla_y V \text{ in } L^2(0,T;L^2(\Omega\times\mathcal{Y})), \end{array} \tag{4.3b}$$

$$v_t^\varepsilon \xrightarrow{2\text{w}} V_t \text{ in } L^2(0,T;L^2(\Omega\times\mathcal{Y})), \tag{4.3c}$$

where $V \in H^1(0,T;L^2(\Omega;H^1(\mathcal{Y})))$ is the unique weak solution of the effective equation

$$\begin{aligned} V_t &= \operatorname{div}_y(\mathbb{D}\nabla_y V) + F(V) && \text{in } \Omega_T\times\mathcal{Y}, \\ V(0,x,y) &= V_0(x,y) && \text{in } \Omega\times\mathcal{Y}. \end{aligned} \tag{4.4P$_0$}$$

The proof of convergence (4.2), in particular of the strong two-scale convergence results in (4.3a)–(4.3b), relies on a clever choice of test functions, suitable for the weak formulations of the ε- and the limit problem, i.e. (4.1P$_\varepsilon$) and (4.4P$_0$), respectively. For the latter, suitable test functions must belong to $L^2(\Omega;H^1(\mathcal{Y}))$, in particular they have to be Y-periodic. The most direct candidate $(\mathcal{T}_\varepsilon\, v^\varepsilon(t))_\varepsilon$ for $t\in[0,T]$ fixed, supplies the required convergence but is *incompatible* with Y-periodicity, since $\mathcal{T}_\varepsilon\, v^\varepsilon(t) \in L^2(\mathbb{R}^d, H^1(Y))$, only, and $H^1(\mathcal{Y}) = H^1_{\text{per}}(Y) \subsetneqq H^1(Y)$. But the $\mathcal{T}_\varepsilon$-property (1.7) guarantees the *recovery of Y-periodicity for the limit*, which thus is *compatible* with the space of test functions of the limit problem. This is an essential observation for the proof of the strong two-scale convergence (4.3a).

In Section 4.1, we state the assumptions on the given data, which are needed to rigorously carry out the limit passage in (4.2). Based on these assumptions, we expound the existence of unique weak solutions v^ε to (4.1P$_\varepsilon$) and V to (4.4P$_0$), independently of the limit passage. From the uniform boundedness of the solutions $(v^\varepsilon)_\varepsilon$, by Theorem 3.5(b), we conclude that there exists a two-scale function $\widetilde{V}$ such that $v^\varepsilon \xrightarrow{2\mathrm{w}} \widetilde{V}$ and $\varepsilon\nabla v^\varepsilon \xrightarrow{2\mathrm{w}} \nabla_y\widetilde{V}$, up to subsequences, and it is not known a-priori that $\widetilde{V}$ solves (4.4P$_0$). In Theorem 4.1, Section 4.3, it is shown that $v^\varepsilon \xrightarrow{2\mathrm{s}} V$, where V solves (4.4P$_0$) and hence that $\widetilde{V} = V$, which makes the passage to the limit in (4.2) rigorous. The abstract strategy of the proof is presented in Section 4.2.

4.1. **Assumptions and a priori bounds.** Adjusted to the structure of problem (4.1P$_\varepsilon$), respectively (4.4P$_0$), we introduce the following function spaces so that the definitions and results from Section 2 are immediately applicable. Recalling (2.2), we set for $\varepsilon > 0$

$$X_\varepsilon := X \text{ equipped with the norm } \|v\|_{X_\varepsilon} := \|v\|_H + \varepsilon\|\nabla v\|_H. \tag{4.5}$$

Since for $\varepsilon > 0$ fixed, both norms, $\|\cdot\|_X$ and $\|\cdot\|_{X_\varepsilon}$, are equivalent, we have that $X = X_\varepsilon \subset H$. Thus $X_\varepsilon \subset H$ is dense and continuously embedded and we obtain that $X_\varepsilon \subset H \subset X_\varepsilon^*$ is an evolution triple. Moreover we define

$$\mathbb{X} := L^2(\Omega; H^1(\mathcal{Y})) \quad \text{and} \quad \mathbb{H} := L^2(\Omega\times\mathcal{Y}), \tag{4.6}$$

which again yields an evolution triple $\mathbb{X} \subset \mathbb{H} \subset \mathbb{X}^*$. Throughout Section 4, we impose the following assumptions on the given data:

Data specification for the ε-problem (4.1P$_\varepsilon$):

$$\begin{aligned} \forall\,\varepsilon > 0: \quad & \mathbb{D}^\varepsilon \in \mathbb{M}(\Omega,\mu,D_\infty),\ f^\varepsilon \in \mathbb{F}(\Omega,L,C_0),\ \text{and} \\ & \exists\, C \geq 0: \|\operatorname{div}(\varepsilon^2\mathbb{D}^\varepsilon\nabla v_0^\varepsilon)\|_H + \|v_0^\varepsilon\|_H \leq C. \end{aligned} \tag{4.7A$_\varepsilon$}$$

Data specification for the limit problem (4.4P$_0$):

$$\begin{aligned} & \mathbb{D} \in \mathbb{M}(\Omega\times\mathcal{Y},\mu,D_\infty),\ F \in \mathbb{F}(\Omega\times\mathcal{Y},L,C_0),\ \text{and} \\ & \operatorname{div}_y(\mathbb{D}\nabla_y V_0),\ V_0 \in \mathbb{H}. \end{aligned} \tag{4.8A$_0$}$$

Convergence of data:

$$\begin{aligned} \mathcal{T}_\varepsilon\,\mathbb{D}^\varepsilon(x,y) &\to \mathbb{D}^{\mathrm{ex}}(x,y) && \text{for a.a. } (x,y) \in \mathbb{R}^d\times\mathcal{Y}, \\ f^\varepsilon(t,\cdot,A) &\xrightarrow{2\mathrm{s}} F(t,\cdot,\cdot,A) && \text{in } \mathbb{H},\ \text{for all } (t,A) \in [0,T]\times\mathbb{R}^m, \\ v_0^\varepsilon &\xrightarrow{2\mathrm{s}} V_0 && \text{in } \mathbb{H}. \end{aligned} \tag{4.9A$_{\varepsilon\to0}$}$$

By (4.7A$_\varepsilon$), relying on Theorem 2.1 and Proposition 2.2, we have the existence of a unique weak solution $v^\varepsilon \in H^1(0,T;X_\varepsilon)\cap H^2(0,T;X_\varepsilon^*)$ of (4.1P$_\varepsilon$). The additional regularity assumptions in (4.7A$_\varepsilon$), i.e. $\operatorname{div}(\varepsilon^2\mathbb{D}^\varepsilon\nabla v_0^\varepsilon)$ uniformly bounded in H, as well as their analogies in (4.8A$_0$), ensure improved time-regularity for the solutions via Proposition 2.2, see (4.10), respectively (4.11), below. On the one hand, this improved time-regularity helps to overcome the technical difficulty that the operators $\mathcal{T}_\varepsilon$ and $\mathcal{F}_\varepsilon$ from Section 3.2 are defined for integrable functions only. On the other hand, it allows us to find uniform upper bounds pointwise and uniform in time, which will simplify the proof of Theorem 4.1:

$$\|v^\varepsilon\|_{C^1([0,T];H)} + \|v^\varepsilon\|_{H^1(0,T;X_\varepsilon)} + \|v^\varepsilon\|_{H^2(0,T;X_\varepsilon^*)} \leq C_b. \tag{4.10}$$

To determine the constant C_b we have used that, by (4.7A$_\varepsilon$), the diffusion tensor $\mathbb{D}^\varepsilon$, resp. reaction term f^ε, belong to the same class for all $\varepsilon > 0$, namely $\mathbb{M}(\Omega,\mu,D_\infty)$, resp. $\mathbb{F}(\Omega,L,C_0)$. Therefore $\frac{\varepsilon^2 D_\infty}{\sqrt{\varepsilon^2\mu}} \leq \frac{D_\infty}{\sqrt{\mu}}$ provides a uniform bound on v^ε in (2.8) and (2.10) for all $\varepsilon \in (0,1)$. In view of the definition of X_ε, we obtain the existence of a constant $C_b \geq 0$, independent of ε, such that (4.10) holds for all $\varepsilon \in (0,1)$.

Analogously, for the limit problem (4.4P$_0$), we may apply Theorem 2.1 and Proposition 2.2 to obtain the existence of a unique weak solution $V \in H^1(0,T;\mathbb{X}) \cap H^2(0,T;\mathbb{X}^*)$ of (4.4P$_0$). Again, estimates (2.8) and (2.10) imply

$$\|V\|_{C^1([0,T];\mathbb{H})} + \|V\|_{H^1(0,T;\mathbb{X})} + \|V\|_{H^2(0,T;\mathbb{X}^*)} \leq C_b, \tag{4.11}$$

with C_b from (4.10), since the given data belong to the same classes of diffusion tensors and reaction terms, in particular with the same parameters $(\mu, D_\infty, L, C_\infty)$.

4.2. **Abstract strategy for proving strong two-scale convergence.** To highlight the general approach to the proof of the strong two-scale convergence result (4.3a), we consider the two abstract systems

$$v^\varepsilon_t = \mathcal{A}^\varepsilon v^\varepsilon + f^\varepsilon(v^\varepsilon) \quad \text{and} \quad V_t = \mathbb{A}V + \mathbb{F}(V) \tag{4.12}$$

in the Hilbert spaces $\mathcal{X} \subset \mathcal{H}$ and $\mathbb{X} \subset \mathbb{H}$, respectively. The operators $\mathcal{A}^\varepsilon$ and $\mathbb{A}$ are given in terms of uniformly bounded and uniformly elliptic quadratic forms, namely

$$\mathcal{B}_\varepsilon(v,w) = \langle -\mathcal{A}^\varepsilon v, w\rangle \quad \text{and} \quad \mathbb{B}(V,W) = \langle\!\langle -\mathbb{A}V, W\rangle\!\rangle.$$

We consider an unfolding operator $\mathcal{T}_\varepsilon : \mathcal{H} \to \mathbb{H}$ which also satisfies $\mathcal{T}_\varepsilon : \mathcal{X} \to \widetilde{\mathbb{X}}$, where $\mathbb{X} \subsetneqq \widetilde{\mathbb{X}}$ is a closed subspace. For the corresponding folding operators $\mathcal{F}_\varepsilon : \mathbb{H} \to \mathcal{H}$ and $\mathcal{G}_\varepsilon : \mathbb{X} \to \mathcal{X}$, we assume that $\mathcal{T}_\varepsilon{}' = \mathcal{F}_\varepsilon$ and that $\mathcal{F}_\varepsilon$ and $\mathcal{G}_\varepsilon$ are comparable in the sense of Theorem 3.9.

We want to show that the solution v^ε converges to V, i.e. $W^\varepsilon \to 0$ in $\mathbb{H}$ or $w^\varepsilon \to 0$ in $\mathcal{H}$, where

$$W^\varepsilon := \mathcal{T}_\varepsilon v^\varepsilon - V \quad \text{and} \quad w^\varepsilon := v^\varepsilon - \mathcal{G}_\varepsilon V.$$

For the proof we resort to working with W^ε instead of w^ε, since this gives the desired two-scale convergence more directly. In particular, to establish this convergence for $(W^\varepsilon)_\varepsilon$, we derive a Gronwall estimate

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|W^\varepsilon\|^2 \leq L\|W^\varepsilon\|^2 + \Delta^\varepsilon, \tag{4.13}$$

where $\|\cdot\|$ stands for the norm in the Hilbert space $\mathbb{H}$. From

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|W^\varepsilon\|^2 = \langle\!\langle W^\varepsilon_t, W^\varepsilon\rangle\!\rangle = \langle\!\langle (\mathcal{T}_\varepsilon v^\varepsilon)_t, W^\varepsilon\rangle\!\rangle - \langle\!\langle V_t, W^\varepsilon\rangle\!\rangle \tag{4.14}$$

we see that it is desirable to test the equations (4.12) with $\mathcal{F}_\varepsilon W^\varepsilon$ and W^ε, respectively. However this is not possible as we do neither have $\mathcal{F}_\varepsilon W^\varepsilon \in \mathcal{X}$ nor $W^\varepsilon \in \mathbb{X}$. Indeed $w^\varepsilon \in \mathcal{X}$ is an admissible test function for $(4.12)_1$, but $\mathcal{T}_\varepsilon w^\varepsilon \notin \mathbb{X}$, due to $\mathcal{T}_\varepsilon v^\varepsilon \in \widetilde{\mathbb{X}} \supsetneqq \mathbb{X}$.

Observe that $\mathcal{B}_\varepsilon : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, whereas $\mathbb{B} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$. To overcome this discrepancy in the underlying spaces $\mathcal{X}$ and $\mathbb{X}$, we replace $\mathcal{B}_\varepsilon$ with a quadratic form $\mathbb{B}_\varepsilon : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ with the same properties as $\mathcal{B}_\varepsilon$ and compensate their mismatch by an additional error term. Thus, we obtain four different types of errors, namely

1. Δ^ε_1 for the *folding mismatch* between $\mathcal{F}_\varepsilon V$ and $\mathcal{G}_\varepsilon V$,
2. Δ^ε_2 for the *incompatibility* of $\mathcal{T}_\varepsilon v^\varepsilon \in \widetilde{\mathbb{X}} \supsetneqq \mathbb{X}$,
3. Δ^ε_3 for the *approximation error* between $\mathbb{B}$ and $\mathbb{B}_\varepsilon$, and
4. Δ^ε_4 for the *approximation error* between f^ε and $\mathbb{F}$.

More precisely, we test $(4.12)_1$ with $w^\varepsilon = (v^\varepsilon - \mathcal{F}_\varepsilon V) + (\mathcal{F}_\varepsilon V - \mathcal{G}_\varepsilon V)$, transform the equation from $\mathcal{X}$ to $\mathbb{X}$ using $\mathcal{T}_\varepsilon$ and $\mathbb{B}_\varepsilon$ so that we obtain

$$\begin{aligned} &\langle\!\langle (\mathcal{T}_\varepsilon v^\varepsilon)_t, W^\varepsilon \rangle\!\rangle = -\mathbb{B}_\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon, W^\varepsilon) + \langle\!\langle \mathcal{T}_\varepsilon f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon), W^\varepsilon \rangle\!\rangle + \Delta_1^\varepsilon, \qquad (4.15)\\ &\text{where } \Delta_1^\varepsilon := \langle\!\langle (\mathcal{T}_\varepsilon v^\varepsilon)_t, W^\varepsilon \rangle\!\rangle + \mathbb{B}_\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon, W^\varepsilon) - \langle\!\langle \mathcal{T}_\varepsilon f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon), W^\varepsilon \rangle\!\rangle \\ &\qquad - \langle v_t^\varepsilon, w^\varepsilon\rangle \quad - \quad \mathcal{B}_\varepsilon(v^\varepsilon, w^\varepsilon) \quad + \quad \langle f^\varepsilon(v^\varepsilon), w^\varepsilon\rangle. \end{aligned}$$

We may additionally assume that $\mathbb{B}$ is well-defined on $\widetilde{\mathbb{X}}$ as well. However, testing $(4.12)_2$ with $W^\varepsilon \in \widetilde{\mathbb{X}}$ is not allowed, since equation $(4.12)_2$ is valid in the subspace $\mathbb{X}$, only. Nevertheless, each of the expressions $\langle\!\langle V_t, W^\varepsilon \rangle\!\rangle$, $\mathbb{B}(V, W^\varepsilon)$, $\langle\!\langle \mathbb{F}(V), W^\varepsilon \rangle\!\rangle$ is well-defined. Therefore we test $(4.12)_2$ with V only, include the missing terms containing $\mathcal{T}_\varepsilon v^\varepsilon$ and compensate them by the *incompatibility error term* Δ_2^ε via

$$\begin{aligned} \langle\!\langle V_t, W^\varepsilon \rangle\!\rangle &= -\langle\!\langle V_t, V \rangle\!\rangle + \langle\!\langle V_t, \mathcal{T}_\varepsilon v^\varepsilon \rangle\!\rangle = \mathbb{B}(V,V) - \langle\!\langle \mathbb{F}(V), V \rangle\!\rangle + \langle\!\langle V_t, \mathcal{T}_\varepsilon v^\varepsilon \rangle\!\rangle \\ &= -\mathbb{B}(V, W^\varepsilon) + \langle\!\langle \mathbb{F}(V), W^\varepsilon \rangle\!\rangle - \Delta_2^\varepsilon, \qquad (4.16)\\ \text{where } \Delta_2^\varepsilon &:= -\langle\!\langle V_t, \mathcal{T}_\varepsilon v^\varepsilon \rangle\!\rangle - \mathbb{B}(V, \mathcal{T}_\varepsilon v^\varepsilon) + \langle\!\langle \mathbb{F}(V), \mathcal{T}_\varepsilon v^\varepsilon \rangle\!\rangle. \end{aligned}$$

Since V is a weak solution in $\mathbb{X}$, the error Δ_2^ε would vanish, if $\mathcal{T}_\varepsilon v^\varepsilon \in \mathbb{X}$, i.e. $\mathcal{T}_\varepsilon v^\varepsilon$ would be an admissible test function. In general this is not the case, but in analogy to the $\mathcal{T}_\varepsilon$-property of recovered periodicity (1.7), we may assume that $V = \text{w-}\lim_{\varepsilon\to 0} \mathcal{T}_\varepsilon v^\varepsilon$ is compatible with the space $\mathbb{X}$, despite the fact that $\mathcal{T}_\varepsilon v^\varepsilon \notin \mathbb{X}$. Thus, we have $\lim_{\varepsilon\to 0} \Delta_2^\varepsilon = 0$.

Inserting (4.15)&(4.16) into (4.14), we obtain

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|W^\varepsilon\|^2 &= -\mathbb{B}_\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon, W^\varepsilon) + \mathbb{B}(V, W^\varepsilon) + \langle\!\langle \mathcal{T}_\varepsilon f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon), W^\varepsilon \rangle\!\rangle - \langle\!\langle \mathbb{F}(V), W^\varepsilon \rangle\!\rangle \\ &\quad + \Delta_1^\varepsilon + \Delta_2^\varepsilon \\ &= -\mathbb{B}_\varepsilon(W^\varepsilon, W^\varepsilon) + \langle\!\langle \mathcal{T}_\varepsilon f^\varepsilon(\mathcal{T}_\varepsilon f^\varepsilon) - \mathcal{T}_\varepsilon f^\varepsilon(V), W^\varepsilon \rangle\!\rangle + \Delta^\varepsilon, \qquad (4.17) \end{aligned}$$

where $\Delta^\varepsilon := \sum_{i=1}^4 \Delta_i^\varepsilon$ collects also the *approximation errors* of the given data, viz.

$$\Delta_3^\varepsilon := \mathbb{B}(V, W^\varepsilon) - \mathbb{B}_\varepsilon(V, W^\varepsilon) \quad \text{and} \quad \Delta_4^\varepsilon := \langle\!\langle \mathcal{T}_\varepsilon f^\varepsilon(V), W^\varepsilon \rangle\!\rangle - \langle\!\langle \mathbb{F}(V), W^\varepsilon \rangle\!\rangle.$$

Exploiting the uniform ellipticity of $\mathcal{B}_\varepsilon$ and the global Lipschitz continuity of f^ε, equation (4.17) yields the Gronwall estimate (4.13). It is then left to show that the error Δ^ε vanishes for $\varepsilon \to 0$. Together with the assumption $W^\varepsilon(0) \to 0$, one then obtains the desired result $W^\varepsilon(t) \to 0$ for all $t > 0$.

4.3. **Main theorem on the strong two-scale convergence.** In this section we prove the strong two-scale convergence of the slow diffusive variable v^ε, which is the most critical ingredient for the homogenization of the degenerating equation (4.1P_ε) and the system of coupled equations (1.1$\mathrm{P}_\varepsilon^{\mathrm{cp}}$).

Theorem 4.1 (Strong two-scale convergence of the slow diffusive variable)**.** *Let the assumptions* (4.7A_ε)*,* (4.8A_0)*, and* (4.9$\mathrm{A}_{\varepsilon\to 0}$) *hold true; then the weak solutions* $(v^\varepsilon)_\varepsilon$ *of* (4.1P_ε) *two-scale converge to the weak solution* V *of* (4.4P_0) *as stated in* (4.3)*.*

The proof goes along the abstract strategy explained in Section 4.2, where we set

$$\begin{aligned}&\mathcal{H}=H=L^2(\Omega),\quad \mathcal{X}=X=H^1(\Omega),\quad \mathbb{H}=L^2(\Omega\times\mathcal{Y}),\\&\mathbb{X}=L^2(\Omega;H^1(\mathcal{Y}))\subset\widetilde{\mathbb{X}}=L^2(\Omega;H^1(Y)),\\&\mathcal{A}^\varepsilon v=\operatorname{div}(\varepsilon^2\mathbb{D}^\varepsilon\nabla v)\quad\text{and}\quad \mathbb{A}V=\operatorname{div}_y(\mathbb{D}\nabla_y V)\quad\text{for } v\in X \text{ and } V\in\mathbb{X}.\end{aligned}\tag{4.18}$$

Integrating by parts and using the no-flux resp. periodic boundary conditions, the quadratic forms read

$$\mathcal{B}_\varepsilon(v,w)=\langle-\operatorname{div}(\varepsilon^2\mathbb{D}^\varepsilon\nabla v),w\rangle_{X^*,X}=\int_\Omega \mathbb{D}^\varepsilon\varepsilon\nabla v:\varepsilon\nabla w\,\mathrm{d}x,$$

$$\mathbb{B}(V,W)=\langle\!\langle-\operatorname{div}_y(\mathbb{D}\nabla_y V),W\rangle\!\rangle_{\mathbb{X}^*,\mathbb{X}}=\int_{\Omega\times\mathcal{Y}}\mathbb{D}\nabla_y V:\nabla_y W\,\mathrm{d}x\,\mathrm{d}y.$$

For technical reasons, we prove the convergence $W^\varepsilon(t)\to 0$ in $L^2(\mathbb{R}^d\times\mathcal{Y})$ for all $t\in[0,T]$ and not in $\mathbb{H}=L^2(\Omega\times\mathcal{Y})$. Here,

$$W^\varepsilon(t):=\mathcal{T}_\varepsilon v^\varepsilon(t)-V^{\mathrm{ex}}(t)\in\widetilde{\mathbb{X}},\tag{4.19}$$

by Theorem 3.2(a). To this end, we extend all functions $V\in\mathbb{H}$, resp. $V\in\mathbb{X}$, to the whole $\mathbb{R}^d$ with 0 outside of Ω so that $V^{\mathrm{ex}}\in L^2(\mathbb{R}^d\times\mathcal{Y})$, resp. $V^{\mathrm{ex}}\in L^2(\mathbb{R}^d;H^1(Y))$. Thereby no regularity is lost, since in any case V is only square-integrable with respect to $x\in\Omega$.

Before entering the details of the proof, let us now sketch its main steps:

Step 1. Extraction of weakly convergent subsequences: The a priori bound (4.10) allows us to extract a weakly two-scale converging subsequence of $(v^\varepsilon)_\varepsilon$ with a limit $\widetilde{V}$. By improving the convergence from weak to strong in the subsequent steps, we are able to show that $\widetilde{V}$ equals the unique solution V of (4.4P$_0$) and to conclude the convergence of the whole sequence.

Step 2. Reformulation of (4.1P$_\varepsilon$) *and specification of the folding mismatch* Δ_1^ε*:* The underlying domains of the ε-problem (4.1P$_\varepsilon$) and the effective one (4.4P$_0$) are Ω and $\Omega\times\mathcal{Y}$. To subtract their weak formulations, as in (4.14)–(4.16), we unfold the ε-problem to the common domain of integration $\mathbb{R}^d\times\mathcal{Y}$ by using the folding and unfolding operators from Sections 3.2 and 3.4. Inserting a suitable test function, we arrive at the definition of the folding mismatch Δ_1^ε as specified in (4.15).

Step 3. Specification of the incompatibility error Δ_2^ε*:* We derive equation (4.16) and the exact form of the error term Δ_2^ε induced by the incompatibility of $\mathcal{T}_\varepsilon v^\varepsilon$, see (1.7). The error terms Δ_1^ε and Δ_2^ε look in principle as in Section 4.2, but are a little more involved owing to the precise definition of the folding and unfolding operators.

Step 4. Preparation of the Gronwall estimate and the approximation errors Δ_3^ε *and* Δ_4^ε*:* As in (4.13) & (4.17), we subtract the reformulated weak formulations of (4.1P$_\varepsilon$) and (4.4P$_0$), derived in Step 2–3, and we precise the error terms Δ_3^ε and Δ_4^ε, which contain the approximation errors $\mathbb{D}^\varepsilon\rightsquigarrow\mathbb{D}$ and $f^\varepsilon\rightsquigarrow F$.

Step 5. Estimation of W^ε *via Gronwall's lemma as in* (4.13).

Step 6. Control of the error terms Δ_i^ε and the strong two-scale convergence (4.3a).

Step 7. Derivation of the remaining convergences (4.3b)–(4.3c).

Proof of Theorem 4.1. **Step 1. Extraction of weakly convergent subsequences.** The uniform bound (4.10) implies $\|v^\varepsilon\|_{C^1([0,T];H)}+\|\varepsilon\nabla v^\varepsilon\|_{H^1(0,T;H)}\le C$.

Using the Arzelà-Ascoli theorem together with Theorem 3.5(b), we find a subsequence ε' of ε and a limit $\widetilde{V} \in C^1([0,T];\mathbb{H}) \cap H^1(0,T;\mathbb{X})$ such that

$$\forall\, t \in [0,T]: \quad v^{\varepsilon'}(t) \xrightarrow{2\text{w}} \widetilde{V}(t) \;\text{ and }\; \varepsilon'\nabla v^{\varepsilon'}(t) \xrightarrow{2\text{w}} \nabla_y \widetilde{V}(t) \quad \text{in } \mathbb{H}. \tag{4.20}$$

For the subsequent steps we resort to working with the above extracted subsequence, labeling it by ε again for notational simplicity.

Step 2. Reformulation of (4.1P$_\varepsilon$) and specification of the folding mismatch Δ_1^ε. Let $t \in [0,T]$ be arbitrary but fixed and let all upcoming equations hold for all $t \in [0,T]$, if not stated otherwise. The weak formulation of (4.1P$_\varepsilon$) reads

$$\int_\Omega v_t^\varepsilon \cdot \varphi \,\mathrm{d}x = \int_\Omega -\mathbb{D}^\varepsilon \varepsilon\nabla v^\varepsilon : \varepsilon\nabla\varphi + f^\varepsilon(v^\varepsilon)\cdot\varphi\,\mathrm{d}x \quad \text{for all } \varphi \in X_\varepsilon. \tag{4.21}$$

Let $V \in H^1(0,T;\mathbb{X}) \cap H^2(0,T;\mathbb{X}^*)$ be the unique weak solution of (4.4P$_0$) and we choose the test function $\varphi^\varepsilon = v^\varepsilon - \mathcal{G}_\varepsilon V \in X_\varepsilon$, according to Definition 3.7. Using the identity $\mathcal{F}_\varepsilon \mathcal{T}_\varepsilon = \mathrm{id}|_{L^2(\Omega)}$ and adding $\pm\mathcal{F}_\varepsilon V$ and $\pm\mathcal{F}_\varepsilon(\nabla_y V)$, we obtain

$$\begin{aligned}\int_\Omega v_t^\varepsilon \cdot \mathcal{F}_\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon - V)\,\mathrm{d}x = \int_\Omega &-\mathbb{D}^\varepsilon\varepsilon\nabla v^\varepsilon : \mathcal{F}_\varepsilon\left[\mathcal{T}_\varepsilon(\varepsilon\nabla v^\varepsilon) - \nabla_y V\right]\\ &+ f^\varepsilon(v^\varepsilon)\cdot\mathcal{F}_\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon - V)\,\mathrm{d}x + \Delta_1^\varepsilon\end{aligned} \tag{4.22}$$

with the folding mismatch error

$$\Delta_1^\varepsilon := \int_\Omega -\varepsilon\mathbb{D}^\varepsilon\nabla v^\varepsilon : [\mathcal{F}_\varepsilon(\nabla_y V) - \varepsilon\nabla(\mathcal{G}_\varepsilon V)] + [f^\varepsilon(v^\varepsilon) - v_t^\varepsilon]\cdot(\mathcal{F}_\varepsilon V - \mathcal{G}_\varepsilon V)\,\mathrm{d}x.$$

Since $\mathcal{T}_\varepsilon$ is a linear and bounded operator, it commutes with differentiation, i.e. $\mathcal{T}_\varepsilon(v_t^\varepsilon) = (\mathcal{T}_\varepsilon v^\varepsilon)_t$. Exploiting $\mathcal{F}_\varepsilon{}' = \mathcal{T}_\varepsilon$, as well as $\mathcal{T}_\varepsilon[\mathbb{D}^\varepsilon\varepsilon\nabla v^\varepsilon] = \mathcal{T}_\varepsilon\,\mathbb{D}^\varepsilon\,\mathcal{T}_\varepsilon(\varepsilon\nabla v^\varepsilon)$, $\mathcal{T}_\varepsilon[f^\varepsilon(v^\varepsilon)] = \mathcal{T}_\varepsilon\, f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon)$ and $\mathcal{T}_\varepsilon(\varepsilon\nabla v^\varepsilon) = \nabla_y(\mathcal{T}_\varepsilon v^\varepsilon)$ (Theorem 3.2 and (3.3)), we arrive at

$$\begin{aligned}&\int_{\mathbb{R}^d\times\mathcal{Y}} (\mathcal{T}_\varepsilon v^\varepsilon)_t \cdot W^\varepsilon\,\mathrm{d}x\,\mathrm{d}y = \\ &\qquad\int_{\mathbb{R}^d\times\mathcal{Y}} \mathcal{T}_\varepsilon\, f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon)\cdot W^\varepsilon - \mathcal{T}_\varepsilon\,\mathbb{D}^\varepsilon\nabla_y(\mathcal{T}_\varepsilon v^\varepsilon) : \nabla_y W^\varepsilon\,\mathrm{d}x\,\mathrm{d}y + \Delta_1^\varepsilon.\end{aligned} \tag{4.23}$$

Hence, the reformulation of (4.1P$_\varepsilon$) is completed, and (4.15) is established with

$$\mathbb{B}_\varepsilon(V,W) = \int_{\mathbb{R}^d\times\mathcal{Y}} \mathcal{T}_\varepsilon\,\mathbb{D}^\varepsilon\nabla_y V : \nabla_y W\,\mathrm{d}x\,\mathrm{d}y.$$

Step 3. Specification of the incompatibility error Δ_2^ε. Next we consider the weak formulation of the effective equation (4.4P$_0$)

$$\int_{\Omega\times\mathcal{Y}} V_t\cdot\Phi\,\mathrm{d}x\,\mathrm{d}y = \int_{\Omega\times\mathcal{Y}} -\mathbb{D}\nabla_y V : \nabla_y\Phi + F(V)\cdot\Phi\,\mathrm{d}x\,\mathrm{d}y \quad \text{for all } \Phi\in\mathbb{X}. \tag{4.24}$$

We aim to derive (4.16), but we observe a discrepancy in the domains of integration in (4.23) and (4.24). Therefore we reformulate (4.24) by extending all the functions by 0 outside of Ω, i.e.

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathcal{Y}} V_t^{\mathrm{ex}}\cdot\Phi\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^d\times\mathcal{Y}} -\mathbb{D}^{\mathrm{ex}}\nabla_y V^{\mathrm{ex}} : \nabla_y\Phi + F^{\mathrm{ex}}(V^{\mathrm{ex}})\cdot\Phi\,\mathrm{d}x\,\mathrm{d}y&\\ \text{for all } \Phi\in L^2(\mathbb{R}^d;H^1(\mathcal{Y})).&\end{aligned} \tag{4.25}$$

Although $\Phi = \mathcal{T}_\varepsilon v^\varepsilon$ is not admissible in (4.25) because of the $\mathcal{T}_\varepsilon$-property (1.7), each integral expression in (4.25), considered on its own, is well-defined for $\Phi = \mathcal{T}_\varepsilon v^\varepsilon$. Because of this, we test (4.25) with $\Phi = V^{\mathrm{ex}}$ only and then add and subtract the missing terms $-V_t^{\mathrm{ex}} \cdot \mathcal{T}_\varepsilon v^\varepsilon + \mathbb{D}\nabla_y V^{\mathrm{ex}} : \nabla_y(\mathcal{T}_\varepsilon v^\varepsilon) - F^{\mathrm{ex}}(V^{\mathrm{ex}}) \cdot v^\varepsilon$ at the cost of creating the error Δ_2^ε:

$$\begin{aligned}
&-\int_{\mathbb{R}^d\times\mathcal{Y}} V_t^{\mathrm{ex}} \cdot W^\varepsilon \,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^d\times\mathcal{Y}} V_t^{\mathrm{ex}} \cdot V^{\mathrm{ex}} \,\mathrm{d}x\,\mathrm{d}y - \int_{\mathbb{R}^d\times\mathcal{Y}} V_t^{\mathrm{ex}} \cdot \mathcal{T}_\varepsilon v^\varepsilon \,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{\mathbb{R}^d\times\mathcal{Y}} -\mathbb{D}^{\mathrm{ex}}\nabla_y V^{\mathrm{ex}} : \nabla_y V^{\mathrm{ex}} + F^{\mathrm{ex}}(V^{\mathrm{ex}}) \cdot V^{\mathrm{ex}} \,\mathrm{d}x\,\mathrm{d}y - \int_{\mathbb{R}^d\times\mathcal{Y}} V_t^{\mathrm{ex}} \cdot \mathcal{T}_\varepsilon v^\varepsilon \,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{\mathbb{R}^d\times\mathcal{Y}} \mathbb{D}^{\mathrm{ex}}\nabla_y V^{\mathrm{ex}} : \nabla_y W^\varepsilon - F^{\mathrm{ex}}(V^{\mathrm{ex}}) \cdot W^\varepsilon \,\mathrm{d}x\,\mathrm{d}y + \Delta_2^\varepsilon, \qquad (4.26)
\end{aligned}$$

where the incompatibility error is given by

$$\Delta_2^\varepsilon := \int_{\mathbb{R}^d\times\mathcal{Y}} -\mathbb{D}^{\mathrm{ex}}\nabla_y V^{\mathrm{ex}} : \nabla_y(\mathcal{T}_\varepsilon v^\varepsilon) + [F^{\mathrm{ex}}(V^{\mathrm{ex}}) - V_t^{\mathrm{ex}}] \cdot \mathcal{T}_\varepsilon v^\varepsilon \,\mathrm{d}x\,\mathrm{d}y.$$

Thus (4.16) is established.

Step 4. Preparation of the Gronwall estimate and the approximation errors Δ_3^ε and Δ_4^ε. For applying Gronwall's lemma at the end of Step 5, we now prepare the estimate (4.13). We begin by adding (4.23) and (4.26), as suggested in (4.14) & (4.17),

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|W^\varepsilon\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} = \int_{\mathbb{R}^d\times\mathcal{Y}} W_t^\varepsilon \cdot W^\varepsilon \,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^d\times\mathcal{Y}} (\mathcal{T}_\varepsilon v^\varepsilon)_t \cdot W^\varepsilon - V_t^{\mathrm{ex}} \cdot W^\varepsilon \,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{\mathbb{R}^d\times\mathcal{Y}} -\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon \nabla_y(\mathcal{T}_\varepsilon v^\varepsilon) : \nabla_y W^\varepsilon + \mathcal{T}_\varepsilon f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon) \cdot W^\varepsilon \,\mathrm{d}x\,\mathrm{d}y + \Delta_1^\varepsilon \\
&\quad + \int_{\mathbb{R}^d\times\mathcal{Y}} \mathbb{D}^{\mathrm{ex}}\nabla_y V^{\mathrm{ex}} : \nabla_y W^\varepsilon - F^{\mathrm{ex}}(V^{\mathrm{ex}}) \cdot W^\varepsilon \,\mathrm{d}x\,\mathrm{d}y + \Delta_2^\varepsilon. \qquad (4.27)
\end{aligned}$$

Rewriting the gradient terms via

$$\begin{aligned}
&-\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon \nabla_y(\mathcal{T}_\varepsilon v^\varepsilon) : \nabla_y W^\varepsilon + \mathbb{D}^{\mathrm{ex}}\nabla_y V^{\mathrm{ex}} : \nabla_y W^\varepsilon \\
&= -\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon \nabla_y W^\varepsilon : \nabla_y W^\varepsilon + (\mathbb{D}^{\mathrm{ex}} - \mathcal{T}_\varepsilon \mathbb{D}^\varepsilon)\nabla_y V^{\mathrm{ex}} : \nabla_y W^\varepsilon,
\end{aligned}$$

equation (4.27) takes the form

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|W^\varepsilon\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} \\
&= \int_{\mathbb{R}^d\times\mathcal{Y}} -\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon \nabla_y W^\varepsilon : \nabla_y W^\varepsilon + \mathcal{T}_\varepsilon f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon) \cdot W^\varepsilon - F^{\mathrm{ex}}(V^{\mathrm{ex}}) \cdot W^\varepsilon \,\mathrm{d}x\,\mathrm{d}y \\
&\quad + \Delta_1^\varepsilon + \Delta_2^\varepsilon + \Delta_3^\varepsilon, \qquad (4.28) \\
&\quad \text{where} \quad \Delta_3^\varepsilon := \int_{\mathbb{R}^d\times\mathcal{Y}} (\mathbb{D}^{\mathrm{ex}} - \mathcal{T}_\varepsilon \mathbb{D}^\varepsilon)\nabla_y V^{\mathrm{ex}} : \nabla_y W^\varepsilon \,\mathrm{d}x\,\mathrm{d}y.
\end{aligned}$$

Analogously we rearrange the reaction terms in (4.28) via

$$\begin{aligned}
&\mathcal{T}_\varepsilon f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon) \cdot W^\varepsilon - F^{\mathrm{ex}}(V^{\mathrm{ex}}) \cdot W^\varepsilon \\
&= [\mathcal{T}_\varepsilon f^\varepsilon(\mathcal{T}_\varepsilon v^\varepsilon) - \mathcal{T}_\varepsilon f^\varepsilon(V^{\mathrm{ex}})] \cdot W^\varepsilon + [\mathcal{T}_\varepsilon f^\varepsilon(V^{\mathrm{ex}}) - F^{\mathrm{ex}}(V^{\mathrm{ex}})] \cdot W^\varepsilon. \qquad (4.29)
\end{aligned}$$

Moreover we define the approximation error Δ_4^ε for f^ε via

$$\Delta_4^\varepsilon := \int_{\mathbb{R}^d\times\mathcal{Y}} [\mathcal{T}_\varepsilon f^\varepsilon(V^{\rm ex}) - F^{\rm ex}(V^{\rm ex})]\cdot W^\varepsilon \,{\rm d}x\,{\rm d}y \quad \text{and} \quad \Delta^\varepsilon := \sum_{i=1}^4 \Delta_i^\varepsilon.$$

Hence, $\Delta^\varepsilon(t)$ depends implicitly on t via the functions f^ε, F and the solutions v^ε, V.

Step 5. Estimating W^ε in terms of Δ^ε via Gronwall's lemma. Using (4.29), the ellipticity of $\mathbb{D}^\varepsilon$ and the global Lipschitz continuity of f^ε in (4.28) yield for all $t\in[0,T]$

$$\begin{aligned}
&\frac{1}{2}\frac{\rm d}{{\rm d}t}\|W^\varepsilon(t)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})}\\
&= \int_{\mathbb{R}^d\times\mathcal{Y}} -\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon\nabla_y W^\varepsilon(t):\nabla_y W^\varepsilon(t) + [\mathcal{T}_\varepsilon f^\varepsilon(t,\mathcal{T}_\varepsilon v^\varepsilon(t)) - \mathcal{T}_\varepsilon f^\varepsilon(t,V^{\rm ex}(t))]\cdot W^\varepsilon(t)\,{\rm d}x\,{\rm d}y\\
&\quad + \Delta^\varepsilon(t)\\
&\le -\mu\|\nabla_y W^\varepsilon(t)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} + L\|W^\varepsilon(t)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} + \Delta^\varepsilon(t)\\
&\le L\|W^\varepsilon(t)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} + \Delta^\varepsilon(t).
\end{aligned} \tag{4.30}$$

Applying Gronwall's lemma we obtain, for all $t\in[0,T]$, the estimate

$$\begin{aligned}
\|W^\varepsilon(t)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} &\le {\rm e}^{2Lt}\Big(\|W^\varepsilon(0)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} + 2\int_0^t \Delta^\varepsilon(s)\,{\rm d}s\Big)\\
&\le {\rm e}^{2LT}\Big(\|W^\varepsilon(0)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} + 2\int_0^T \Delta^\varepsilon(s)\,{\rm d}s\Big).
\end{aligned} \tag{4.31}$$

Step 6. Controlling the error terms Δ_j^ε and the strong convergence (4.3a). To show that the right-hand side in (4.31) vanishes as $\varepsilon\to 0$, we provide an ε-independent and integrable majorant for each $\Delta_i^\varepsilon:[0,T]\to[0,\infty)$ and show further that $\Delta_i^\varepsilon(t)\to 0$ for all $t\in[0,T]$ as $\varepsilon\to 0$.

For the folding mismatch error Δ_1^ε we apply Hölder's inequality, the assumptions on the given data (4.7A$_\varepsilon$), and the uniform bound (4.10) so that we arrive at

$$\begin{aligned}
|\Delta_1^\varepsilon(t)| &= \Big|\int_\Omega -v_t^\varepsilon(t)\cdot(\mathcal{F}_\varepsilon V(t) - \mathcal{G}_\varepsilon V(t)) - \mathbb{D}^\varepsilon\varepsilon\nabla v^\varepsilon(t) : (\mathcal{F}_\varepsilon[\nabla_y V(t)] - \varepsilon\nabla[\mathcal{G}_\varepsilon V(t)])\\
&\qquad + f^\varepsilon(t,v^\varepsilon(t))\cdot(\mathcal{F}_\varepsilon V(t) - \mathcal{G}_\varepsilon V(t))\,{\rm d}x\,{\rm d}y\Big|\\
&\le C(C_b)\left(\|\mathcal{F}_\varepsilon V(t) - \mathcal{G}_\varepsilon V(t)\|_H + \|\mathcal{F}_\varepsilon[\nabla_y V(t)] - \varepsilon\nabla[\mathcal{G}_\varepsilon V(t)]\|_H\right).
\end{aligned}$$

By the a priori estimate (4.11) we find the uniform L^∞–majorant $|\Delta_1^\varepsilon(t)|\le C(C_b)$. Moreover, Theorem 3.9 guarantees the pointwise convergence $\Delta_1^\varepsilon(t)\xrightarrow{\varepsilon\to 0} 0$ for all $t\in[0,T]$.

For the incompatibility error $\Delta_2^\varepsilon(t)$ the uniform bounds (4.10)-(4.11) provide an L^∞–majorant, whereas the pointwise convergence follows form the $\mathcal{T}_\varepsilon$-property of recovered periodicity (1.7), since $\widetilde{V}(t)\in\mathbb{X}$ obtained in (4.20) is an admissible test function for the weak solution V. Indeed, we have

$$
\begin{aligned}
&\Delta_2^\varepsilon(t) = \\
&\int_{\mathbb{R}^d\times\mathcal{Y}} -\mathbb{D}^{\rm ex}\nabla_y V^{\rm ex}(t) : \nabla_y(\mathcal{T}_\varepsilon\, v^\varepsilon(t)) + [F(t,V^{\rm ex}(t)) - V_t^{\rm ex}(t)]\cdot \mathcal{T}_\varepsilon\, v^\varepsilon(t)\,\mathrm{d}x\,\mathrm{d}y \\
&\xrightarrow{\varepsilon\to0} \int_{\mathbb{R}^d\times\mathcal{Y}} -\mathbb{D}^{\rm ex}\nabla_y V^{\rm ex}(t) : \nabla_y\widetilde{V}^{\rm ex}(t) + [F^{\rm ex}(t,V^{\rm ex}(t)) - V_t^{\rm ex}(t)]\cdot \widetilde{V}^{\rm ex}(t)\,\mathrm{d}x\,\mathrm{d}y \\
&\quad = 0,
\end{aligned}
$$

since V solves (4.24).

With the same arguments we obtain L^∞–majorants for the approximation errors Δ_3^ε and Δ_4^ε. Moreover, Lemma 3.6(a&b) and $(4.9\mathrm{A}_{\varepsilon\to0})_{1,2}$ yield

$$
\begin{aligned}
|\Delta_3^\varepsilon(t)| &= \left|\int_{\mathbb{R}^d\times\mathcal{Y}} (\mathbb{D}^{\rm ex} - \mathcal{T}_\varepsilon\,\mathbb{D}^\varepsilon)\nabla_y V^{\rm ex}(t) : \nabla_y W^\varepsilon(t)\,\mathrm{d}x\,\mathrm{d}y\right| \\
&\le 2C_b\|(\mathbb{D}^{\rm ex} - \mathcal{T}_\varepsilon\,\mathbb{D}^\varepsilon)\nabla_y V^{\rm ex}(t)\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} \xrightarrow{\varepsilon\to0} 0 \quad \text{and} \\
|\Delta_4^\varepsilon(t)| &= \left|\int_{\mathbb{R}^d\times\mathcal{Y}} [\mathcal{T}_\varepsilon\, f^\varepsilon(t,V^{\rm ex}(t)) - F^{\rm ex}(t,V^{\rm ex}(t))]\cdot W^\varepsilon(t)\,\mathrm{d}x\,\mathrm{d}y\right| \\
&\le 2C_b\|\,\mathcal{T}_\varepsilon\, f^\varepsilon(t,V^{\rm ex}(t)) - F^{\rm ex}(t,V^{\rm ex}(t))\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} \xrightarrow{\varepsilon\to0} 0 \quad \text{for all } t\in[0,T].
\end{aligned}
$$

Since the dominated convergence yields $\int_0^T|\Delta^\varepsilon(t)|\,\mathrm{d}t \xrightarrow{\varepsilon\to0} 0$ and $(4.9\mathrm{A}_{\varepsilon\to0})_3$ implies $\|W^\varepsilon(0)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} = \|\,\mathcal{T}_\varepsilon\, v_0^\varepsilon - V_0^{\rm ex}\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} \xrightarrow{\varepsilon\to0} 0$, we conclude $\max_{t\in[0,T]}\|W^\varepsilon(t)\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} \xrightarrow{\varepsilon\to0} 0$ and the strong convergence of the subsequence extracted in (4.20) is established.

With the usual arguments, by considering another, different subsequence and since $V\in C^0([0,T];\mathbb{H})$ is the unique weak solution of $(4.4\mathrm{P}_0)$, we conclude that $W^\varepsilon(t) \xrightarrow{2\mathrm{s}} 0$ in $\mathbb{H}$ uniformly for all $t\in[0,T]$, even for the *whole* sequence. Hence, (4.3a) is proved.

Step 7. Proof of the remaining convergence results (4.3b)–(4.3c)**.** Let $t\in[0,T]$ be arbitrary but fixed. From (4.30) we obtain

$$
\begin{aligned}
&\mu\|\nabla_y W^\varepsilon(t)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} \le \int_{\mathbb{R}^d\times\mathcal{Y}} \mathcal{T}_\varepsilon\,\mathbb{D}^\varepsilon\nabla_y W^\varepsilon(t) : \nabla_y W^\varepsilon(t)\,\mathrm{d}x\,\mathrm{d}y \\
&\le \int_{\mathbb{R}^d\times\mathcal{Y}} -W_t^\varepsilon(t)\cdot W^\varepsilon(t) + [\mathcal{T}_\varepsilon\, f^\varepsilon(t,\mathcal{T}_\varepsilon\, v^\varepsilon(t)) - \mathcal{T}_\varepsilon\, f^\varepsilon(t,V^{\rm ex}(t))]\cdot W^\varepsilon(t)\,\mathrm{d}x\,\mathrm{d}y \\
&\quad + \Delta^\varepsilon(t) \\
&\le 2C_b\|W^\varepsilon(t)\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} + L\|W^\varepsilon(t)\|^2_{L^2(\mathbb{R}^d\times\mathcal{Y})} + \Delta^\varepsilon(t) \to 0, \qquad (4.32)
\end{aligned}
$$

where we have used $\|W_t^\varepsilon(t)\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} \le \|v_t^\varepsilon(t)\|_H + \|V_t(t)\|_{\mathbb{H}} \le 2C_b$, (4.3a) and $\Delta^\varepsilon(t)\to0$ pointwise for all $t\in[0,T]$. Integrating (4.32) over $(0,T)$ yields $\nabla_y W^\varepsilon \to 0$ strongly in $L^2(0,T;L^2(\mathbb{R}^d\times\mathcal{Y}))$ and hence (4.3b).

It remains to prove $v_t^\varepsilon \xrightarrow{2\mathrm{w}} V_t$ in $\mathbf{H} = L^2(0,T;\mathbb{H})$. By the a priori bound (4.10) we know that v_t^ε is bounded in the Hilbert space $\mathbf{H}$, and hence has a weakly convergent subsequence with limit U. Choosing $\Phi\in C_{\rm c}^\infty([0,T];\mathbb{H})$, the definition of the weak time derivative gives

$$(U,\Phi)_{\mathbf{H}} \xleftarrow{\varepsilon\to 0} (\mathcal{T}_\varepsilon(v_t^\varepsilon),\Phi)_{\mathbf{H}} = ((\mathcal{T}_\varepsilon\, v^\varepsilon)_t,\Phi)_{\mathbf{H}}$$
$$= -(\mathcal{T}_\varepsilon\, v^\varepsilon,\Phi_t)_{\mathbf{H}} \xrightarrow{\varepsilon\to 0} -(V,\Phi_t)_{\mathbf{H}} = (V_t,\Phi)_{\mathbf{H}}.$$

Since Φ was arbitrary we conclude $U = V_t$ and (4.3c) is established. Thus, the proof of Theorem 4.1 is complete. □

5. **Homogenization of the coupled system.** In this section we consider the system of two coupled reaction-diffusion systems ($5.1\mathrm{P}_\varepsilon^{\mathrm{cp}}$), where the coupling arises via the reaction term $f^\varepsilon = (f_1^\varepsilon, f_2^\varepsilon)$, whereas the diffusion tensor $\mathbb{D}^\varepsilon$ has block structure. We write ($1.1\mathrm{P}_\varepsilon^{\mathrm{cp}}$) shortly in the form

$$\begin{aligned} w_t^\varepsilon &= (\mathbb{D}^\varepsilon\nabla w^\varepsilon) + f^\varepsilon(w^\varepsilon) && \text{in } \Omega_T,\\ 0 &= (\mathbb{D}^\varepsilon\nabla w^\varepsilon)\cdot\vec{n} && \text{on } \Gamma_T,\\ w^\varepsilon(0) &= w_0^\varepsilon && \text{in } \Omega. \end{aligned} \qquad (5.1\mathrm{P}_\varepsilon^{\mathrm{cp}})$$

We are looking for solutions $w^\varepsilon := (u^\varepsilon, v^\varepsilon) : \Omega_T \to \mathbb{R}^{m_1+m_2}$ of ($5.1\mathrm{P}_\varepsilon^{\mathrm{cp}}$), where the diffusion tensor $\mathbb{D}^\varepsilon : \Omega \to \mathbb{R}^{([m_1+m_2]\times d)\times([m_1+m_2]\times d)}$ is assumed to have the form

$$\mathbb{D}^\varepsilon(x) = \begin{pmatrix} \mathbb{D}_1^\varepsilon(x) & 0\\ 0 & \varepsilon^2\mathbb{D}_2^\varepsilon(x)\end{pmatrix} \quad \text{with } \mathbb{D}_1^\varepsilon, \mathbb{D}_2^\varepsilon \in \mathbb{M}(\Omega,\mu,D_\infty). \qquad (5.2)$$

Hence the component v^ε diffuses much slower than the "classically" diffusing one u^ε.

The main result of this section is Theorem 5.1, which states that w^ε converges (for $\varepsilon \to 0$) to a limit W that decomposes into a one-scale function u and a two-scale function V, i.e. $W(t,x,y) = (u(t,x), V(t,x,y))$. We prove that W solves the coupled effective system

$$\begin{aligned} W_t &= \begin{pmatrix}\mathrm{div}(\mathbb{D}_{\mathrm{eff}}\nabla u)\\ \mathrm{div}_y(\mathbb{D}_2\nabla_y V)\end{pmatrix} + \begin{pmatrix} f_{\mathrm{eff}}(W)\\ F_2(W)\end{pmatrix} && \text{in } \Omega_T\times\mathcal{Y},\\ 0 &= (\mathbb{D}_{\mathrm{eff}}\nabla u)\cdot\vec{n} && \text{on } \Gamma_T,\\ W(0) &= W_0 && \text{in } \Omega, \end{aligned} \qquad (5.3\mathrm{P}_0^{\mathrm{cp}})$$

where the effective diffusion tensor $\mathbb{D}_{\mathrm{eff}}$ and the effective u-reaction f_{eff} only depend on the macroscopic variable $x \in \Omega$, while the diffusion tensor $\mathbb{D}_2$ and the V-reaction F_2 depend on the two-scale variables $(x,y) \in \Omega\times\mathcal{Y}$, see ($5.5\mathrm{A}_0$) and (5.6)-(5.8) below.

5.1. **Notation and existence for the coupled and effective systems.** Following the notations (2.2) and (4.5)–(4.6), we define the function spaces

$$\overline{X}_\varepsilon := X\times X_\varepsilon,\quad \overline{\mathbb{X}} := X\times\mathbb{X},\quad \overline{H} := H\times H,\quad \text{and}\quad \overline{\mathbb{H}} := H\times\mathbb{H} \qquad (5.4)$$

and we obtain the two evolution triples $\overline{X}_\varepsilon \subset \overline{H} \subset \overline{X}_\varepsilon^*$ and $\overline{\mathbb{X}} \subset \overline{\mathbb{H}} \subset \overline{\mathbb{X}}^*$. In the spirit of Section 4.1, we impose the following assumptions on the data of ($5.1\mathrm{P}_\varepsilon^{\mathrm{cp}}$)-($5.3\mathrm{P}_0^{\mathrm{cp}}$):

($4.7\mathrm{A}_\varepsilon$) holds for $\mathbb{D}_i^\varepsilon$ in (5.2), $f^\varepsilon := (f_1^\varepsilon, f_2^\varepsilon)$, and $w_0^\varepsilon = (u_0^\varepsilon, v_0^\varepsilon)$ satisfies $\|\,\mathrm{div}(\mathbb{D}_1^\varepsilon\nabla u_0^\varepsilon)\|_H + \|\,\mathrm{div}(\varepsilon^2\mathbb{D}_2^\varepsilon\nabla v_0^\varepsilon)\|_H + \|w_0^\varepsilon\|_{\overline{H}} \le C$; ($5.5\mathrm{A}_\varepsilon$)

($4.8\mathrm{A}_0$) holds for $\mathbb{D}_i : \Omega\times\mathcal{Y} \to \mathbb{R}^{(m_i\times d)\times(m_i\times d)}$, $i = 1,2$, $F := (F_1,F_2) : [0,T]\times\Omega\times\mathcal{Y}\times\mathbb{R}^{m_1+m_2} \to \mathbb{R}^{m_1+m_2}$ and W_0; ($5.5\mathrm{A}_0$)

($4.9\mathrm{A}_{\varepsilon\to 0}$) holds for $\mathbb{D}_i^\varepsilon \rightsquigarrow \mathbb{D}_i$, $f_i^\varepsilon \rightsquigarrow F_i$, $i = 1,2$, and $w_0^\varepsilon \rightsquigarrow W_0$. ($5.5\mathrm{A}_{\varepsilon\to 0}$)

The effective diffusion tensor $\mathbb{D}_{\text{eff}}$ and the effective reaction term f_{eff} of the classical equation $(5.3\text{P}_0^{\text{cp}})_2$ are given as follows. The function-to-function map $f_{\text{eff}} : [0,T] \times \Omega \times \mathbb{R}^{m_1} \times L^2(\mathcal{Y};\mathbb{R}^{m_2}) \to \mathbb{R}^{m_1}$ is defined as

$$f_{\text{eff}}(t,x,u,Z) := \int_{\mathcal{Y}} F_1(t,x,y,u,Z(y))\,\mathrm{d}y. \tag{5.6}$$

The effective diffusion tensor $\mathbb{D}_{\text{eff}} : \Omega \to \mathbb{R}^{(m_1\times d)\times(m_1\times d)}$ is given componentwise via the classical homogenization formula, see e.g. [2, Eq. (2.20)], [1, Eq. (2.6)-(2.7)], [26, Eq. (46)-(47)],

$$\mathbb{D}_{\text{eff}}(x)_{ijkl} := \int_{\mathcal{Y}} \mathbb{D}_1(x,y)_{ijkl} + \sum_{r=1}^{d} \mathbb{D}_1(x,y)_{ijkr} \cdot \partial_{y_r} z(y)_{kl}\,\mathrm{d}y, \tag{5.7}$$

for $i,k = 1,...,m_1,\ j,l = 1,...,d$, where the so-called correctors $z_{ij} \in H^1_{\text{av}}(\mathcal{Y})$ solve the local problem in the weak sense:

$$\mathrm{div}_y \left(\mathbb{D}_1(x,y)_{ijkl} + \sum_{r=1}^{d} \mathbb{D}_1(x,y)_{ijkr} \cdot \partial_{y_r} z(y)_{kl} \right) = 0 \quad \text{in } \mathcal{Y} \text{ for a.a. } x \in \Omega. \tag{5.8}$$

It is easy to check that $f_{\text{eff}} \in \mathbb{F}(\Omega, L, C_\infty)$ and $\mathbb{D}_{\text{eff}}$ again satisfies $\mathbb{D}_{\text{eff}} \in \mathbb{M}(\Omega,\mu,\frac{D_\infty^2}{\mu})$, see e.g. [8, Thm. 13.4] or [37, Thm. 2], since $\mathbb{D}_1^\varepsilon$ in particular H-converges to $\mathbb{D}_{\text{eff}}$.

For given $T > 0$ and all $\varepsilon \in (0,1)$, Theorem 2.1 and Proposition 2.2 yield the existence of unique weak solutions $w^\varepsilon \in H^1(0,T;\overline{X}_\varepsilon) \cap H^2(0,T;\overline{X}_\varepsilon^*)$ of $(5.1\text{P}_\varepsilon^{\text{cp}})$ and $W \in H^1(0,T;\overline{\mathbb{X}}) \cap H^2(0,T;\overline{\mathbb{X}}^*)$ of $(5.3\text{P}_0^{\text{cp}})$ with

$$\exists\, C_b \geq 0 : \quad \|w^\varepsilon\|_{H^1(0,T;\overline{X}_\varepsilon)\cap H^2(0,T;\overline{X}_\varepsilon^*)} + \|W\|_{H^1(0,T;\overline{\mathbb{X}})\cap H^2(0,T;\overline{\mathbb{X}}^*)} \leq C_b. \tag{5.9}$$

Note that this improved time-regularity is valid in the weighted space $\overline{X}_\varepsilon$, i.e. $\|w^\varepsilon\|^2_{\overline{X}_\varepsilon} = \|w^\varepsilon\|^2_{L^2(\Omega)} + \|\nabla u^\varepsilon\|^2_{L^2(\Omega)} + \varepsilon^2\|\nabla v^\varepsilon\|^2_{L^2(\Omega)}$.

5.2. **Convergence of the coupled system.** Finally we present the convergence result for the coupled system involving the concentrations $u(t,x) \in \mathbb{R}^{m_1}$ for the classical diffusive species and the concentrations $v(t,x) \in \mathbb{R}^{m_2}$ for the slow diffusive species.

Theorem 5.1. *Let the assumptions* $(5.5\text{A}_\varepsilon)$, (5.5A_0) *and* $(5.5\text{A}_{\varepsilon\to 0})$ *be satisfied. The sequence of weak solutions* $w^\varepsilon = (u^\varepsilon, v^\varepsilon) \in H^1(0,T;\overline{X}_\varepsilon) \cap H^2(0,T;\overline{X}_\varepsilon^*)$ *of* $(5.1\text{P}_\varepsilon^{\text{cp}})$ *converges to the weak solution* $(u,V) \in H^1(0,T;\overline{\mathbb{X}}) \cap H^2(0,T;\overline{\mathbb{X}}^*)$ *of* $(5.3\text{P}_0^{\text{cp}})$ *in the following sense: There exists a function* $U \in L^2(0,T;L^2(\Omega;H^1_{av}(\mathcal{Y})))$ *such that*

$$\begin{aligned}
&u^\varepsilon \rightharpoonup u \text{ in } H^1(0,T;X), && v^\varepsilon(t) \xrightarrow{2\text{s}} V(t) \text{ in } \mathbb{H} \text{ for all } t \in [0,T],\\
&\nabla u^\varepsilon \xrightharpoonup{2\text{w}} \nabla u + \nabla_y U \text{ in } L^2(0,T;\mathbb{H}), && \varepsilon\nabla v^\varepsilon(t) \xrightarrow{2\text{s}} \nabla_y V(t) \text{ in } L^2(0,T;\mathbb{H}),\\
&u_t^\varepsilon \rightharpoonup u_t \text{ in } H^1(0,T;X^*), && v_t^\varepsilon \xrightharpoonup{2\text{w}} V_t \text{ in } L^2(0,T;\mathbb{H}).
\end{aligned} \tag{5.10}$$

Moreover, one can prove $\nabla u^\varepsilon \xrightarrow{2\text{s}} \nabla u + \nabla_y U$ in $L^2(0,T;\mathbb{H})$ in a similar manner as $\varepsilon\nabla v^\varepsilon(t) \xrightarrow{2\text{s}} \nabla_y V(t)$ in $L^2(0,T;\mathbb{H})$ is proved, by defining another gradient folding operator $\widetilde{\mathcal{G}}_\varepsilon : X \times \mathbb{X} \to X$, cf. [22, 35].

The proof relies heavily on Theorem 4.1 and the well-known techniques from the classical theory of non-degenerating two-scale homogenization. In order to give a rigorous derivation of $(5.3\text{P}_0^{\text{cp}})$, it is essential to know that v^ε converges strongly in the two-scale sense, otherwise one cannot pass to the limit with the term $f_1^\varepsilon(u^\varepsilon, v^\varepsilon)$. Hence, we want to use the improved time-regularity (Proposition 2.2) for v^ε in order to apply Theorem 4.1. Therefore we need in particular $u_t^\varepsilon \in X_\varepsilon^* \subset X^*$, which is satisfied for $u_t^\varepsilon \in H \subset X_\varepsilon^*$ as stated in (5.9).

Proof of Theorem 5.1. Let w^ε be the weak solution of (5.1$\mathrm{P}_\varepsilon^{\mathrm{cp}}$) satisfying (5.9). Applying Banach's selection principle yields the existence of a function $u \in H^1(0,T;X) \cap H^2(0,T;X^*)$ such that $u^\varepsilon \rightharpoonup u$ in $H^1(0,T;X)$ and $u_t^\varepsilon \rightharpoonup u_t$ in $H^1(0,T;X^*)$ hold true, up to a subsequence. Moreover, Theorem 3.5(c) yields $U \in L^2(0,T;L^2(\Omega;H^1_{\mathrm{av}}(\mathcal{Y})))$ such that $\nabla u^\varepsilon \xrightarrow{2\mathrm{w}} \nabla u + \nabla_y U$ in $L^2(0,T;\mathbb{H})$, up to a subsequence. In particular, the compact embeddings $X \subset H$ and $H^1(0,T;X) \subset C^\alpha([0,T];X)$, for $\alpha \in (0,\frac{1}{2})$, together with Proposition 3.4(d) yield

$$\forall\, t \in [0,T]: \quad u^\varepsilon(t) \to u(t) \quad \text{in } H \quad \Longrightarrow \quad u^\varepsilon(t) \xrightarrow{2\mathrm{s}} u(t) \quad \text{in } \mathbb{H}. \tag{5.11}$$

Testing (5.1$\mathrm{P}_\varepsilon^{\mathrm{cp}}$) with functions of the kind $(\varphi,0) \in \overline{X}_\varepsilon$ or $(0,\varphi) \in \overline{X}_\varepsilon$, we can consider the first, or "classical", equation $(5.1\mathrm{P}_\varepsilon^{\mathrm{cp}})_1$ separated from the degenerating second one $(5.1\mathrm{P}_\varepsilon^{\mathrm{cp}})_2$.

Step 1. Convergence of the slow diffusive variable v^ε. In view of the notation from Section 4, we set $g^\varepsilon(t,x,A) := f_2^\varepsilon(t,x,u^\varepsilon(t,x),A)$, $F(t,x,y,A) := F_2(t,x,y,u(t,x),A)$ for $A \in \mathbb{R}^{m_2}$ and $\widetilde{\mathbb{H}} := L^2(\mathbb{R}^d \times \mathcal{Y})$ for brevity. Then Lemma 3.6(b) with $U(x,y) = u(x)$ and the strong convergence (5.11) give

$$\begin{aligned} &\| \mathcal{T}_\varepsilon[g^\varepsilon(A)] - F^{\mathrm{ex}}(A)\|_{C^0([0,T];\widetilde{\mathbb{H}})} \\ &\le \| \mathcal{T}_\varepsilon\, f_2^\varepsilon(\mathcal{T}_\varepsilon\, u^\varepsilon, A) - \mathcal{T}_\varepsilon\, f_2^\varepsilon(u,A)\|_{C^0([0,T];\widetilde{\mathbb{H}})} + \| \mathcal{T}_\varepsilon\, f_2^\varepsilon(u,A) - F_2^{\mathrm{ex}}(u,A)\|_{C^0([0,T];\widetilde{\mathbb{H}})} \\ &\le L\| \mathcal{T}_\varepsilon\, u^\varepsilon - u^{\mathrm{ex}}\|_{C^0([0,T];\widetilde{\mathbb{H}})} + \| \mathcal{T}_\varepsilon\, f_2^\varepsilon(u,A) - F_2^{\mathrm{ex}}(u,A)\|_{C^0([0,T];\widetilde{\mathbb{H}})} \longrightarrow 0, \end{aligned} \tag{5.12}$$

which implies $g^\varepsilon(t,\cdot,A) \xrightarrow{2\mathrm{s}} F^{\mathrm{ex}}(t,\cdot,\cdot,A)$ in $\mathbb{H}$ for all $(t,A) \in [0,T] \times \mathbb{R}^{m_2}$. Thus we can apply Theorem 4.1 to the degenerating equation $v_t^\varepsilon = \mathrm{div}(\varepsilon^2 \mathbb{D}_2^\varepsilon \nabla v^\varepsilon) + g^\varepsilon(v^\varepsilon)$ in Ω_T and we obtain $v^\varepsilon(t) \xrightarrow{2\mathrm{s}} V(t)$ in $\mathbb{H}$ for all $t \in [0,T]$, where V solves

$$V_t = \mathrm{div}_y(\mathbb{D}_2 \nabla_y V) + F(V) \quad \text{in } \Omega_T \times \mathcal{Y} \tag{5.13}$$

with $F(V) = F_2(u,V)$. The convergences $\varepsilon \nabla v^\varepsilon \xrightarrow{2\mathrm{s}} \nabla_y V$ and $v_t^\varepsilon \xrightarrow{2\mathrm{w}} V_t$ in $L^2(0,T;\mathbb{H})$ follow from (4.3). Hence (5.10) is shown, up to a subsequence, and it is left to prove that the limit $W = (u,V)$, in particular u, solves (5.3$\mathrm{P}_0^{\mathrm{cp}}$).

Step 2. "Classical" homogenization. We begin with considering the convergence of the reaction term f_1^ε. Arguing as in (5.12) gives

$$\begin{aligned} &\| \mathcal{T}_\varepsilon[f_1^\varepsilon(u^\varepsilon, v^\varepsilon)] - F_1^{\mathrm{ex}}(u,V)\|_{C^0([0,T];\widetilde{\mathbb{H}})} \\ &\le L\|(\mathcal{T}_\varepsilon\, u^\varepsilon, \mathcal{T}_\varepsilon\, v^\varepsilon) - (u,V)^{\mathrm{ex}}\|_{C^0([0,T];\widetilde{\mathbb{H}})} + \| \mathcal{T}_\varepsilon\, f_1^\varepsilon(u,V) - F_1^{\mathrm{ex}}(u,V)\|_{C^0([0,T];\widetilde{\mathbb{H}})} \\ &\to 0. \end{aligned} \tag{5.14}$$

From (5.14) and Proposition 3.4(e), we infer $f_1^\varepsilon(u^\varepsilon, v^\varepsilon) \rightharpoonup f_{\mathrm{eff}}(u,V)$ in $L^2(0,T;H)$. We derive $(5.3\mathrm{P}_0^{\mathrm{cp}})_1$ by exploiting the convergences in (5.10), up to subsequences so far, and by choosing two different test functions, one after another.

(1) For $\varphi \in X$, we clearly have $\nabla\varphi \xrightarrow{2s} \nabla\varphi$ in $\mathbb{H}$. Testing (5.1$\mathrm{P}_\varepsilon^{\mathrm{cp}}$) with $(\varphi, 0) \in \overline{X}_\varepsilon$ and using formula (3.4) gives

$$\begin{aligned}
&\int_\Omega u_t^\varepsilon \cdot \varphi \,\mathrm{d}x = \int_\Omega -\mathbb{D}_1^\varepsilon \nabla u^\varepsilon : \nabla\varphi + f_1^\varepsilon(u^\varepsilon, v^\varepsilon) \cdot \varphi \,\mathrm{d}x \\
\Longleftrightarrow &\int_\Omega u_t^\varepsilon \cdot \varphi \,\mathrm{d}x = \int_{\mathbb{R}^d \times \mathcal{Y}} -\mathcal{T}_\varepsilon \,\mathbb{D}_1^\varepsilon \,\mathcal{T}_\varepsilon(\nabla u^\varepsilon) : \mathcal{T}_\varepsilon(\nabla\varphi) \,\mathrm{d}x \,\mathrm{d}y + \int_\Omega f_1^\varepsilon(u^\varepsilon, v^\varepsilon) \cdot \varphi \,\mathrm{d}x \\
\xrightarrow{\varepsilon\to 0} &\int_\Omega u_t \cdot \varphi \,\mathrm{d}x = \int_\Omega -\left(\int_{\mathcal{Y}} \mathbb{D}_1[\nabla u + \nabla_y U] \,\mathrm{d}y\right) : \nabla\varphi + f_{\text{eff}}(u, V) \cdot \varphi \,\mathrm{d}x \qquad (5.15)
\end{aligned}$$

a.e. in $(0, T)$. For the passage $\xrightarrow{\varepsilon\to 0}$, we have applied Lemma 3.6(a) to each component of $(\mathcal{T}_\varepsilon \,\mathbb{D}_1^\varepsilon)\,\mathcal{T}_\varepsilon(\nabla u^\varepsilon) : \mathcal{T}_\varepsilon(\nabla\varphi)$. The strong formulation of (5.15) then reads

$$u_t = \operatorname{div}\left(\int_{\mathcal{Y}} \mathbb{D}_1[\nabla u + \nabla_y U] \,\mathrm{d}y\right) + f_{\text{eff}}(u, V) \quad \text{in } \Omega_T, \qquad (5.16)$$

where $\int_{\mathcal{Y}} \mathbb{D}_1[\nabla u + \nabla_y U] \,\mathrm{d}y = \mathbb{D}_{\text{eff}} \nabla u$ with $\mathbb{D}_{\text{eff}}$ from (5.7) will be deduced in part (3) below. Note that (5.16) already resembles the structure of $(5.3\mathrm{P}_0^{\mathrm{cp}})_1$.

(2) Secondly we test (5.1$\mathrm{P}_\varepsilon^{\mathrm{cp}}$) with $(\varphi^\varepsilon, 0) \in \overline{X}_\varepsilon$ for $\varphi^\varepsilon(x) := \varepsilon\varphi_1(x)\varphi_2(\frac{x}{\varepsilon})$, where $\varphi_1 \in C^\infty(\overline{\Omega})$ and $\varphi_2 \in C^\infty(\mathcal{Y})$. Notice that $\nabla\varphi^\varepsilon(x) = \varepsilon\nabla\varphi_1(x)\varphi_2(\frac{x}{\varepsilon}) + \varphi_1(x)\nabla_y\varphi_2(\frac{x}{\varepsilon})$ as well as $(\varphi_2(\frac{\cdot}{\varepsilon}), \nabla\varphi_2(\frac{\cdot}{\varepsilon})) \xrightarrow{2s} (\varphi_2, \nabla_y\varphi_2)$ in $\mathbb{X}$ since $\mathcal{T}_\varepsilon[\varphi(\frac{x}{\varepsilon})] = \varphi(y)$. Thus $(\varphi^\varepsilon, \nabla\varphi^\varepsilon) \xrightarrow{2s} (0, \varphi_1\nabla_y\varphi_2)$ in $\mathbb{X}$ and we proceed as in (5.15), which gives

$$0 = \int_{\Omega\times\mathcal{Y}} \mathbb{D}_1[\nabla u + \nabla_y U] : \varphi_1 \nabla_y \varphi_2 \,\mathrm{d}x \,\mathrm{d}y.$$

Applying the fundamental lemma in the calculus of variations twice and using integration by parts once, we arrive at the local problem

$$\operatorname{div}_y(\mathbb{D}_1[\nabla u + \nabla_y U]) = 0 \quad \text{in } \Omega_T \times \mathcal{Y}, \qquad (5.17)$$

which is an elliptic PDE on $\mathcal{Y}$ and Ω_T may be considered as a set of parameters. With the three equations (5.13), (5.16), and (5.17), the functions u, U, and V are uniquely determined (in the sense of weak solutions).

(3) In order to derive (5.7)–(5.8), we use the ansatz of separation of variables $U_i(x, y) = \nabla u_i(x) \cdot z_i(y) = \sum_{k=1}^d \partial_{x_k} u_i(x) \cdot z_{ik}(y)$ for $z_{ik} \in H^1_{\text{av}}(\mathcal{Y})$, $i = 1, ..., m_1$, $k = 1, ..., d$ (cf. [26, Eq. (48)]). Let us assume we can write $\mathbb{D}_{\text{eff}}\nabla u$ in the form

$$[\mathbb{D}_{\text{eff}}\nabla u]\,(x) = \int_{\mathcal{Y}} \mathbb{D}(x, y)\,[\nabla u(x) + \nabla_y U(x, y)] \,\mathrm{d}y. \qquad (5.18)$$

Inserting $\nabla_{y_j} U_i(x, y) = \nabla u_i(x) \cdot \partial_{y_j} z_i(y)$, $j = 1, ..., d$, in (5.17)-(5.18) yields (5.7)-(5.8).

Since the limit W solves (5.3$\mathrm{P}_0^{\mathrm{cp}}$) uniquely, the *whole* sequence converges in (5.10) and the identity in (5.18) is justified. □

5.3. **Discussion of the assumptions.** For given $\mathbb{D} \in \mathbb{M}(\Omega \times \mathcal{Y}, \mu, D_\infty)$ and $F \in \mathbb{F}(\Omega \times \mathcal{Y}, L, C_\infty)$, the natural choice for data satisfying (4.7A_ε)-(4.9$\mathrm{A}_{\varepsilon\to 0}$) is to set,

$$\mathbb{D}^\varepsilon := \mathcal{F}_\varepsilon \,\mathbb{D}^{\text{ex}} \text{ and } f^\varepsilon(t, \cdot, A) := \mathcal{F}_\varepsilon \,F^{\text{ex}}(t, \cdot, \cdot, A) \text{ for all } (t, A) \in [0, T] \times \mathbb{R}^m. \quad (5.19)$$

Clearly $\mathbb{D}^\varepsilon \in \mathbb{M}(\Omega, \mu, D_\infty)$ and $f^\varepsilon \in \mathbb{F}(\Omega, L, C_\infty)$. It remains to verify (4.9$\mathrm{A}_{\varepsilon\to 0}$) for the choice (5.19). Proposition 3.4(c) implies directly $f^\varepsilon(t, \cdot, A) \xrightarrow{2s} F(t, \cdot, \cdot, A)$ in $\mathbb{H}$ for all $(t, A) \in [0, T] \times \mathbb{R}^m$. The pointwise convergence $\mathcal{T}_\varepsilon \,\mathbb{D}^\varepsilon \to \mathbb{D}^{\text{ex}}$ a.e. in $\mathbb{R}^d \times \mathcal{Y}$ is proved in the following result.

Proposition 5.2. *For $\mathbb{D} \in L^\infty(\Omega \times \mathcal{Y})$, we have $\mathcal{T}_\varepsilon \mathcal{F}_\varepsilon \mathbb{D}^{\rm ex}(x,y) \to \mathbb{D}^{\rm ex}(x,y)$ for a.a. $(x,y) \in \mathbb{R}^d \times \mathcal{Y}$.*

Proof. Recalling the notations introduced in Section 3.1, we set $A_\varepsilon := \hat{\Omega}_\varepsilon \times \mathcal{Y}$ and $B_\varepsilon := \{x \in \mathbb{R}^d \,|\, (\mathcal{N}_\varepsilon(x) + \varepsilon Y) \cap \Omega = \emptyset\} \times \mathcal{Y}$. Let $N_{2\varepsilon \mathrm{diam}(Y)}(\Gamma)$ denote the $2\varepsilon\mathrm{diam}(Y)$-neighborhood of the boundary Γ, then we set $N_\varepsilon := N_{2\varepsilon\mathrm{diam}(Y)}(\Gamma) \times \mathcal{Y}$ and it holds $A_\varepsilon \cap B_\varepsilon = \emptyset$ and $A_\varepsilon \cup N_\varepsilon \cup B_\varepsilon = \mathbb{R}^d \times \mathcal{Y}$.

Let an arbitrary point $(x,y) \in \mathbb{R}^d \times \mathcal{Y}$ be given. Then there exists $\varepsilon_0 > 0$ such that $(x,y) \notin N_\varepsilon$ for all $\varepsilon \le \varepsilon_0$, see (3.1). Therefore it holds either (1) $(x,y) \in A_\varepsilon$ or (2) $(x,y) \in B_\varepsilon$.

Ad (1). For $(x,y) \in A_\varepsilon$, the Lebesgue-Besicovitch differentiation theorem, cf. [15, Thm. 1 p. 43], yields

$$(\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon)(x,y) = (\mathcal{T}_\varepsilon \mathcal{F}_\varepsilon \mathbb{D}^{\rm ex})(x,y) = \frac{1}{\varepsilon^d} \int_{\mathcal{N}_\varepsilon(x)+\varepsilon Y} \mathbb{D}(\xi, y)\,\mathrm{d}\xi \xrightarrow{\varepsilon\to 0} \mathbb{D}(x,y).$$

Ad (2). If $(x,y) \in B_\varepsilon$, then $(\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon)(x,y) = 0 = \mathbb{D}^{\rm ex}(x,y)$ and the proof is finished. □

The choice of the initial values v_0^ε and V_0 is more involved, cf. (2.12), and is elaborated in the following proposition.

Proposition 5.3. *For arbitrary $G \in \mathbb{H}$ given, let V_0 be the unique weak solution of*

$$\mathrm{div}_y(\mathbb{D}\nabla_y V_0) - V_0 = G \quad \textit{in } \Omega \times \mathcal{Y}. \tag{5.20}$$

Then there exists a sequence of functions $(v_0^\varepsilon)_\varepsilon$ bounded in X_ε such that $v_0^\varepsilon \xrightarrow{2\mathrm{s}} V_0$ in $\mathbb{X}$ and

$$\mathrm{div}(\varepsilon^2 \mathbb{D}^\varepsilon \nabla v_0^\varepsilon) - v_0^\varepsilon = \mathcal{F}_\varepsilon G^{\rm ex} \quad \textit{in } \Omega. \tag{5.21}$$

Recalling (4.18), we have in particular that $\mathbb{A}V_0$ and $\mathcal{A}^\varepsilon v_0^\varepsilon$ are uniformly bounded in $\mathbb{H}$ and H, respectively. Hence choosing V_0 and v_0^ε as in (5.20) and (5.21), respectively, the assumption (4.9A$_{\varepsilon\to 0}$) is satisfied. An obvious choice for $G \in \mathbb{H}$ is $G = F(V_0)$.

Proof. It is well-known in the literature, cf. [1, 44, 22, 31], that the sequence $(v_0^\varepsilon)_\varepsilon$ of solutions of (5.21) are uniformly bounded in X_ε and that $v_0^\varepsilon \xrightarrow{2\mathrm{w}} V_0$ in $\mathbb{X}$, where V_0 solves (5.20). The strong two-scale convergence follows from the estimate (cf. [22, Thm. 4.1 & Rem. 5.1])

$$\begin{aligned}
&\min\{\mu, 1\} \| \mathcal{T}_\varepsilon v_0^\varepsilon - V_0^{\rm ex} \|^2_{L^2(\mathbb{R}^d; H^1(Y))} \\
&\le \int_{\mathbb{R}^d \times \mathcal{Y}} \mathcal{T}_\varepsilon \mathbb{D}^\varepsilon \nabla_y (\mathcal{T}_\varepsilon v_0^\varepsilon - V_0^{\rm ex}) : \nabla_y (\mathcal{T}_\varepsilon v_0^\varepsilon - V_0^{\rm ex}) + (\mathcal{T}_\varepsilon v_0^\varepsilon - V_0^{\rm ex})^2 \,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{\mathbb{R}^d \times \mathcal{Y}} \Big\{ \underbrace{\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon \nabla_y (\mathcal{T}_\varepsilon v_0^\varepsilon) : \nabla_y (\mathcal{T}_\varepsilon v_0^\varepsilon) + (\mathcal{T}_\varepsilon v_0^\varepsilon)^2}_{= \mathcal{T}_\varepsilon \mathcal{F}_\varepsilon G^{\rm ex} \cdot \mathcal{T}_\varepsilon v_0^\varepsilon} + \underbrace{\mathbb{D}^{\rm ex} \nabla_y V_0^{\rm ex} : \nabla_y V_0^{\rm ex} + (V_0^{\rm ex})^2}_{= G^{\rm ex} \cdot V_0^{\rm ex}} \\
&\qquad - \mathcal{T}_\varepsilon \mathbb{D}^\varepsilon \nabla_y (\mathcal{T}_\varepsilon v_0^\varepsilon) : \nabla_y V_0^{\rm ex} - \mathcal{T}_\varepsilon \mathbb{D}^\varepsilon \nabla_y V_0^{\rm ex} : \nabla_y (\mathcal{T}_\varepsilon v_0^\varepsilon) - 2\, \mathcal{T}_\varepsilon v_0^\varepsilon \cdot V_0^{\rm ex} \\
&\qquad + (\mathcal{T}_\varepsilon \mathbb{D}^\varepsilon - \mathbb{D}^{\rm ex}) \nabla_y V_0^{\rm ex} : \nabla_y V_0^{\rm ex} \Big\} \,\mathrm{d}x\,\mathrm{d}y \\
&\xrightarrow{\varepsilon\to 0} \int_{\mathbb{R}^d \times \mathcal{Y}} 2G^{\rm ex} \cdot V_0^{\rm ex} - 2\mathbb{D}^{\rm ex} \nabla_y V_0^{\rm ex} : \nabla_y V_0^{\rm ex} - 2(V_0^{\rm ex})^2 + 0 \,\mathrm{d}x\,\mathrm{d}y = 0.
\end{aligned}$$

□

Remark 5.4 (Connection to [11])**.** We assume that the solution of (4.4P$_0$) is smooth, i.e. $V \in C^1(\overline{\Omega} \times \mathcal{Y})$, and we set $[V]^\varepsilon(x) := V(x, \frac{x}{\varepsilon})$. Then one can easily show that $\|v^\varepsilon - [V]^\varepsilon\|_{L^2(\Omega)} = \|v^\varepsilon - \mathcal{F}_\varepsilon V^{\rm ex}\|_{L^2(\Omega)} + O(\sqrt{\varepsilon})$ and hence

$$\|v^\varepsilon - [V]^\varepsilon\|_{L^2(\Omega)} \leq \| \mathcal{T}_\varepsilon v^\varepsilon - V^{\rm ex}\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} + \|V^{\rm ex} - \mathcal{T}_\varepsilon \mathcal{F}_\varepsilon V^{\rm ex}\|_{L^2(\mathbb{R}^d\times\mathcal{Y})} + O(\sqrt{\varepsilon}) \xrightarrow{\varepsilon\to 0} 0,$$

by Theorem 4.1 and Proposition 3.4(c). Under suitable smoothness assumptions, our method reproduces the convergence rates in [11], see [46].

Acknowledgments. The authors acknowledge the (partial) funding by the *Collaborative Research Center 910: Control of self-organizing nonlinear systems* via the project *A5: Pattern formation in systems with multiple scales.* Moreover, S.R. and M.T. thank J.A. Griepentrog and H. Hanke for many fruitful discussions and helpful comments.

REFERENCES

[1] G. Allaire, Homogenization and two-scale convergence, *SIAM J. Math. Anal.*, **23** (1992), 1482–1518.

[2] A. Bensoussan, J.-L. Lions and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, Studies in Mathematics and its Applications, 5, North-Holland Publishing Co., Amsterdam, 1978.

[3] D. Bothe and D. Hilhorst, A reaction-diffusion system with fast reversible reaction, *J. Math. Anal. Appl.*, **286** (2003), 125–135.

[4] V. Chalupecký, T. Fatima and A. Muntean, Multiscale sulfate attack on sewer pipes: Numerical study of a fast micro-macro mass transfer limit, *Journal of Math-for-Industry*, **2B** (2010), 171–181.

[5] V. Chalupecký and A. Muntean, Semi-discrete finite difference multiscale scheme for a concrete corrosion model: A priori estimates and convergence, *Jpn. J. Ind. Appl. Math.*, **29** (2012), 289–316.

[6] D. Cioranescu, A. Damlamian and G. Griso, Periodic unfolding and homogenization, *C. R. Math. Acad. Sci. Paris*, **335** (2002), 99–104.

[7] D. Cioranescu, A. Damlamian and G. Griso, The periodic unfolding method in homogenization, *SIAM J. Math. Anal.*, **40** (2008), 1585–1620.

[8] D. Cioranescu and P. Donato, *An Introduction to Homogenization*, Oxford Lecture Series in Mathematics and its Applications, 17, The Clarendon Press Oxford University Press, New York, 1999.

[9] D. Cioranescu, A. Damlamian and R. De Arcangelis, Homogenization of quasiconvex integrals via the periodic unfolding method, *SIAM J. Math. Anal.*, **37** (2006), 1435–1453 (electronic).

[10] A. Damlamian, An elementary introduction to periodic unfolding, *Math. Sci. Appl.*, **24** (2005), 119–136.

[11] C. Eck, Homogenization of a phase field model for binary mixtures, *Multiscale Model. Simul.*, **3** (2004/05), 1–27 (electronic).

[12] J. Elstrodt, *Maß- und Integrationstheorie*, 3rd edition, Springer, 2002.

[13] E. K. Essel, K. Kuliev, G. Kulieva and L.-E. Persson, Homogenization of quasilinear parabolic problems by the method of Rothe and two scale convergence, *Appl. Math.*, **55** (2010), 305–327.

[14] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, 19, American Mathematical Society, Providence, RI, 1998.

[15] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[16] T. Fatima, A. Muntean and M. Ptashnyk, Unfolding-based corrector estimates for a reaction-diffusion system predicting concrete corrosion, *Appl. Anal.*, **91** (2012), 1129–1154.

[17] B. Fiedler and M. Vishik, Quantitative homogenization of analytic semigroups and reaction-diffusion equations with Diophantine spatial frequencies, *Adv. Differential Equations*, **6** (2001), 1377–1408.

[18] B. Fiedler and M. Vishik, Quantitative homogenization of global attractors for reaction-diffusion systems with rapidly oscillating terms, *Asymptot. Anal.*, **34** (2003), 159–185.

[19] L. Flodén and M. Olsson, Reiterated homogenization of some linear and nonlinear monotone parabolic operators, *Can. Appl. Math. Q.*, **14** (2006), 149–183.

[20] A. Giacomini and A. Musesti, Two-scale homogenization for a model in strain gradient plasticity, *ESAIM Control Optim. Calc. Var.*, **17** (2011), 1035–1065.

[21] A. Glitzky and R. Hünlich, Global estimates and asymptotics for electro-reaction-diffusion systems in heterostructures, *Appl. Anal.*, **66** (1997), 205–226.

[22] H. Hanke, Homogenization in gradient plasticity, *Math. Models Meth. Appl. Sci. (M^3AS)*, **21** (2011), 1651–1684.

[23] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Mathematics, 840, Springer-Verlag, Berlin, 1981.

[24] U. Hornung, W. Jäger and A. Mikelić, Reactive transport through an array of cells with semi-permeable membranes, *RAIRO Modél. Math. Anal. Numér.*, **28** (1994), 59–94.

[25] V. V. Jikov, S. M. Kozlov and O. A. Oleĭnik, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, Berlin, Heidelberg, 1994.

[26] D. Lukkassen, G. Nguetseng and P. Wall, Two-scale convergence, *Int. J. Pure Appl. Math.*, **2** (2002), 35–86.

[27] H. S. Mahato, *Homogenization of a System of Nonlinear Multi-Species Diffusion-Reaction Equations in an $H^{1,p}$ Setting*, Ph.D thesis, Universität Bremen, 2013.

[28] V. A. Marchenko and E. Y. Khruslov, *Homogenization of Partial Differential Equations*, Birkhäuser Boston Inc., Boston, MA, 2006.

[29] P. A. Markowich, C. A. Ringhofer and C. Schmeiser, *Semiconductor Equations*, Springer-Verlag, Vienna, 1990.

[30] A. Matache and C. Schwab, Two-scale FEM for homogenization problems, *Math. Model. Numer. Anal. (M2AN)*, **36** (2002), 537–572.

[31] S. A. Meier and A. Muntean, A two-scale reaction-diffusion system: Homogenization and fast-reaction limits, in *Current Advances in Nonlinear Analysis and Related Topics*, GAKUTO Internat. Ser. Math. Sci. Appl., 32, Gakkōtosho, Tokyo, 2010, 443–461.

[32] A. Mielke, A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems, *Nonlinearity*, **24** (2011), 1329–1346.

[33] A. Mielke, Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions, *Discr. Cont. Dynam. Systems Ser. S*, **6** (2013), 479–499.

[34] A. Mielke and E. Rohan, Homogenization of elastic waves in fluid-saturated porous media using the Biot model, *Math. Models Meth. Appl. Sci. (M^3AS)*, **23** (2013), 873–916.

[35] A. Mielke and A. M. Timofte, Two-scale homogenization for evolutionary variational inequalities via the energetic formulation, *SIAM J. Math. Analysis*, **39** (2007), 642–668.

[36] A. Muntean and M. Neuss-Radu, A multiscale Galerkin approach for a class of nonlinear coupled reaction-diffusion systems in complex media, *J. Math. Anal. Appl.*, **371** (2010), 705–718.

[37] F. Murat and L. Tartar, H-convergence, in *Topics in the mathematical modelling of composite materials*, Progr. Nonlinear Differential Equations Appl., 31, Birkhäuser, Boston, MA, 1997, 21–43.

[38] J. D. Murray, *Mathematical Biology. I. An Introduction*, 3rd edition, Interdisciplinary Applied Mathematics, 17, Springer-Verlag, New York, 2002.

[39] S. Nesenenko, Homogenization in viscoplasticity, *SIAM J. Math. Anal.*, **39** (2007), 236–262.

[40] M. Neuss-Radu and W. Jäger, Effective transmission conditions for reaction-diffusion processes in domains separated by an interface, *SIAM J. Math. Anal.*, **39** (2007), 687–720.

[41] G. Nguetseng, A general convergence result for a functional related to the theory of homogenization, *SIAM J. Math. Anal.*, **20** (1989), 608–623.

[42] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, 44, Springer-Verlag, New York, 1983.

[43] J. Persson, Homogenization of monotone parabolic problems with several temporal scales, *Appl. Math.*, **57** (2012), 191–214.

[44] M. A. Peter and M. Böhm, Different choices of scaling in homogenization of diffusion and interfacial exchange in a porous medium, *Math. Meth. Appl. Sci.*, **31** (2008), 1257–1282.

[45] M. Pierre, Global existence in reaction-diffusion systems with control of mass: A survey, *Milan J. Math.*, **78** (2010), 417–455.

[46] S. Reichelt, *Multi-scale Analysis of Nonlinear Reaction-Diffusion Systems*, in preparation, Ph.D thesis, Humboldt-Universität zu Berlin, 2014.

[47] B. Schweizer, Homogenization of degenerate two-phase flow equations with oil trapping, *SIAM J. Math. Anal.*, **39** (2008), 1740–1763.

[48] B. Schweizer and M. Veneroni, Periodic homogenization of Prandtl-Reuss plasticity with hardening, *J. Multiscale Model.*, **2** (2010), 69–106.

[49] L. Tartar, *The General Theory of Homogenization. A Personalized Introduction*, Lecture Notes of the Unione Matematica Italiana, 7, Springer-Verlag, Berlin, 2009.

[50] R. Temam, *Infinite-dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, New York, 1988.

[51] A. Visintin, Two-scale convergence of some integral functionals, *Calc. Var. Partial Differential Equations*, **29** (2007), 239–265.

[52] A. Visintin, Homogenization of a parabolic model of ferromagnetism, *J. Differential Equations*, **250** (2011), 1521–1552.

[53] A. Visintin, Some properties of two-scale convergence, *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, **15** (2004), 93–107.

[54] A. Visintin, Towards a two-scale calculus, *ESAIM Control Optim. Calc. Var.*, **12** (2006), 371–397 (electronic).

[55] A. Visintin, Homogenization of the nonlinear Maxwell model of viscoelasticity and of the Prandtl-Reuss model of elastoplasticity, *Roy. Soc. Edinb. Proc. A*, **138** (2008), 1363–1401.

[56] J. L. Woukeng, Periodic homogenization of nonlinear non-monotone parabolic operators with three time scales, *Ann. Mat. Pura Appl. (4)*, **189** (2010), 357–379.

Received November 2013; revised April 2014.

E-mail address: Alexander.Mielke@wias-berlin.de
E-mail address: Sina.Reichelt@wias-berlin.de
E-mail address: Marita.Thomas@wias-berlin.de