GLOBAL MARTINGALE SOLUTIONS FOR STOCHASTIC SHIGESADA–KAWASAKI–TERAMOTO POPULATION MODELS

MARCEL BRAUKHOFF, FLORIAN HUBER, AND ANSGAR JÜNGEL

ABSTRACT. The existence of global nonnegative martingale solutions to cross-diffusion systems of Shigesada–Kawasaki–Teramoto type with multiplicative noise is proven. The model describes the stochastic segregation dynamics of an arbitrary number of population species in a bounded domain with no-flux boundary conditions. The diffusion matrix is generally neither symmetric nor positive semidefinite, which excludes standard methods for evolution equations. Instead, the existence proof is based on the entropy structure of the model, a novel regularization of the entropy variable, higher-order moment estimates, and fractional time regularity. The regularization technique is generic and is applied to the population system with self-diffusion in any space dimension and without self-diffusion in two space dimensions.

1. Introduction

Shigesada, Kawasaki, and Teramoto (SKT) suggested in their seminal paper [37] a deterministic cross-diffusion system for two competing species, which is able to describe the segregation of the populations. A random influence of the environment or the lack of knowledge of certain biological parameters motivate the introduction of noise terms, leading to the stochastic system for n species with the population density u_i of the ith species:

$$(1)\qquad \mathrm{d}u_i - \operatorname{div}\bigg(\sum_{j=1}^n A_{ij}(u)\nabla u_j\bigg)\mathrm{d}t = \sum_{j=1}^n \sigma_{ij}(u)\mathrm{d}W_j(t) \quad \text{in } \mathcal{O},\ t>0,\ i=1,\ldots,n,$$

with initial and no-flux boundary conditions

$$(2)\qquad u_i(0)=u_i^0 \quad \text{in } \mathcal{O}, \quad \sum_{j=1}^n A_{ij}(u)\nabla u_j\cdot\nu = 0 \quad \text{on } \partial\mathcal{O},\ t>0,\ i=1,\ldots,n,$$

Date: February 25, 2022.

2000 *Mathematics Subject Classification.* 60H15, 35R60, 35Q92.

Key words and phrases. Population dynamics, cross diffusion, martingale solutions, multiplicative noise, entropy method, tightness of laws.

The authors thank Prof. Alexandra Neamţu (Konstanz, Germany) for her constructive help in the preparation of this manuscript. The second and third authors acknowledge partial support from the Austrian Science Fund (FWF), grants I3401, P33010, W1245, and F65. The second author has been supported by the FWF, grant Y1235 of the START program. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, ERC Advanced Grant no. 101018153.

and diffusion coefficients

$$A_{ij}(u) = \delta_{ij}\bigg(a_{i0} + \sum_{k=1}^n a_{ik}u_k\bigg) + a_{ij}u_i, \quad i,j=1,\ldots,n, \tag{3}$$

where $\mathcal{O} \subset \mathbb{R}^d$ ($d \geq 1$) is a bounded domain, ν is the exterior unit normal vector to $\partial\mathcal{O}$, $(W_1,\ldots,W_n)$ is an n-dimensional cylindrical Wiener process, and $a_{ij} \geq 0$ for $i=1,\ldots,n$, $j=0,\ldots,n$ are parameters. The stochastic framework is detailed in Section 2.

The deterministic analog of (1)–(3) generalizes the two-species model of [37] to an arbitrary number of species. The deterministic model can be derived rigorously from nonlocal population systems [19, 35], stochastic interacting particle systems [8], and finite-state jump Markov models [2, 13]. The original system in [37] also contains a deterministic environmental potential and Lotka–Volterra terms, which are neglected here for simplicity.

We call a_{i0} the diffusion coefficients, a_{ii} the self-diffusion coefficients, and a_{ij} for $i \neq j$ the cross-diffusion coefficients. We say that system (1)-(3) is *with self-diffusion* if $a_{i0} \geq 0$, $a_{ii} > 0$ for all $i=1,\ldots,n$, and *without self-diffusion* if $a_{i0} > 0$, $a_{ii} = 0$ for all $i=1,\ldots,n$.

The aim of this work is to prove the existence of global nonnegative martingale solutions to system (1)–(3) allowing for large cross-diffusion coefficients. The existence of a local pathwise mild solution to (1)–(3) with $n=2$ was shown in [30, Theorem 4.3] under the assumption that the diffusion matrix is positive definite. Global martingale solutions to a SKT model with quadratic instead of linear coefficients $A_{ij}(u)$ were found in [18]. Besides detailed balance, this result needs a moderate smallness condition on the cross-diffusion coefficients. We prove the existence of global martingale solutions to the SKT model for general coefficients satisfying detailed balance. This result seems to be new.

There are two major difficulties in the analysis of system (1). The first difficulty is the fact that the diffusion matrix associated to (1) is generally neither symmetric nor positive semidefinite. In particular, standard semigroup theory is not applicable. These issues have been overcome in [9, 10] in the deterministic case by revealing a formal gradient-flow or entropy structure. The task is to extend this idea to the stochastic setting.

In the deterministic case, usually an implicit Euler time discretization is used [24]. In the stochastic case, we need an explicit Euler scheme because of the stochastic Itô integral, but this excludes entropy estimates. An alternative is the Galerkin scheme, which reduces the infinite-dimensional stochastic system to a finite-dimensional one; see, e.g., the proof of [32, Theorem 4.2.4]. This is possible only if energy-type (L^2) estimates are available, i.e. if u_i can be used as a test function. In the present case, however, only entropy estimates are available with the test function $\log u_i$, which is not an element of the Galerkin space.

In the following, we describe our strategy to overcome these difficulties. We say that system (1) has an entropy structure if there exists a function $h : [0,\infty)^n \to [0,\infty)$, called an entropy density, such that the deterministic analog of (1) can be written in terms of the entropy variables (or chemical potentials) $w_i = \partial h/\partial u_i$ as

$$\partial_t u_i(w) - \operatorname{div}\bigg(\sum_{j=1}^n B_{ij}(w)\nabla w_j\bigg) = 0, \quad i = 1\ldots,n, \tag{4}$$

where $w = (w_1, \ldots, w_n)$, u_i is interpreted as a function of w, and $B(w) = A(u(w))h''(u(w))^{-1}$ with $B = (B_{ij})$ is positive semidefinite. For the deterministic analog of (1), it was shown in [11] that the entropy density is given by

$$h(u) = \sum_{i=1}^n \pi_i\big(u_i(\log u_i - 1) + 1\big), \quad u \in [0,\infty)^n, \tag{5}$$

where the numbers $\pi_i > 0$ are assumed to satisfy $\pi_i a_{ij} = \pi_j a_{ji}$ for all $i, j = 1, \ldots, n$. This condition is the detailed-balance condition for the Markov chain associated to (a_{ij}), and $(\pi_1, \ldots, \pi_n)$ is the corresponding reversible stationary measure [11]. Using $w_i = \pi_i \log u_i$ in (4) as a test function and summing over $i = 1, \ldots, n$, a formal computation shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{O}} h(u)\mathrm{d}x + 2\int_{\mathcal{O}}\sum_{i=1}^n \pi_i\bigg(2a_{i0}|\nabla\sqrt{u_i}|^2 + 2a_{ii}|\nabla u_i|^2 + \sum_{j\neq i} a_{ij}|\nabla\sqrt{u_i u_j}|^2\bigg)\mathrm{d}x = 0. \tag{6}$$

A similar expression holds in the stochastic setting; see (29). It provides L^2 estimates for $\nabla\sqrt{u_i}$ if $a_{i0} > 0$ and for ∇u_i if $a_{ii} > 0$. Moreover, having proved the existence of a solution w to an approximate version of (1) leads to the positivity of $u_i(w) = \exp(w_i/\pi_i)$ (and nonnegativity after passing to the de-regularization limit).

To define the approximate scheme, our idea is to "regularize" the entropy variable w. Indeed, instead of the algebraic mapping $w \mapsto u(w)$, we introduce the mapping $Q_\varepsilon(w) = u(w) + \varepsilon L^*Lw$, where $L : D(L) \to H$ with domain $D(L) \subset H$ is a suitable operator and L^* its dual; see Section 3 for details. The operator L is chosen in such a way that all elements of $D(L)$ are bounded functions, implying that $u(w)$ is well defined. Introducing the regularization operator $R_\varepsilon : D(L)' \to D(L)$ as the inverse of $Q_\varepsilon : D(L) \to D(L)'$, the approximate scheme to (1) is defined, written in compact form, as

$$\mathrm{d}v(t) = \operatorname{div}\big(B(R_\varepsilon(v))\nabla R_\varepsilon(v)\big)\mathrm{d}t + \sigma\big(u(R_\varepsilon(v))\big)\mathrm{d}W(t), \quad t > 0. \tag{7}$$

The existence of a local weak solution v^ε to (7) with suitable initial and boundary conditions is proved by applying the abstract result of [32, Theorem 4.2.4]; see Theorem 12. The entropy inequality for $w^\varepsilon := R_\varepsilon(v^\varepsilon)$ and $u^\varepsilon := u(w^\varepsilon)$,

$$\begin{aligned}\mathbb{E}\sup_{0<t<T\wedge\tau_R}\int_{\mathcal{O}} h(u^\varepsilon(t))\mathrm{d}x + \frac{\varepsilon}{2}\mathbb{E}\sup_{0<t<T\wedge\tau_R}\|Lw^\varepsilon(t)\|^2_{L^2(\mathcal{O})} \\ + \mathbb{E}\sup_{0<t<T\wedge\tau_R}\int_0^t\int_{\mathcal{O}}\nabla w^\varepsilon(s) : B(w^\varepsilon(s))\nabla w^\varepsilon(s)\mathrm{d}x\mathrm{d}s \le C(u^0, T),\end{aligned}$$

up to some stopping time $\tau_R > 0$ allows us to extend the local solution to a global one (Proposition 15).

For the de-regularization limit $\varepsilon \to 0$, we need suitable uniform bounds. The entropy inequality provides gradient bounds for u_i^ε in the case with self-diffusion and for $(u_i^\varepsilon)^{1/2}$ in the case without self-diffusion. Based on these estimates, we use the Gagliardo–Nirenberg inequality to prove uniform bounds for u_i^ε in $L^q(0,T;L^q(\mathcal{O}))$ with $q \ge 2$. Such an estimate is crucial to define, for instance, the product $u_i^\varepsilon u_j^\varepsilon$. Furthermore, we show a uniform estimate for u_i^ε in the Sobolev–Slobodeckij space $W^{\alpha,p}(0,T;D(L)')$ for some $\alpha < 1/2$ and

$p > 2$ such that $\alpha p > 1$. These estimates are needed to prove the tightness of the laws of (u^ε) in some sub-Polish space and to conclude strong convergence in L^2 thanks to the Skorokhod–Jakubowski theorem.

For the uniform estimates, we need to distinguish the cases with and without self-diffusion. In the former case, we obtain an $L^2(0,T;H^1(\mathcal{O}))$ estimate for u_i^ε, such that the product $u_i^\varepsilon \nabla u_j^\varepsilon$ is integrable, and we can pass to the limit in the coefficients $A_{ij}(u_i^\varepsilon)$. Without self-diffusion, we can only conclude that (u_i^ε) is bounded in $L^2(0,T;W^{1,1}(\mathcal{O}))$, and products like $u_i^\varepsilon \nabla u_j^\varepsilon$ may be not integrable. To overcome this issue, we use the fact that

$$\operatorname{div}\bigg(\sum_{j=1}^n A_{ij}(u^\varepsilon)\nabla u_j^\varepsilon\bigg) = \Delta\bigg(u_i^\varepsilon\bigg(a_{i0} + \sum_{j=1}^n a_{ij}u_j^\varepsilon\bigg)\bigg) \tag{8}$$

and write (1) in a "very weak" formulation by applying the Laplace operator to the test function. Since the bound in $L^2(0,T;W^{1,1}(\mathcal{O}))$ implies a bound in $L^2(0,T;L^2(\mathcal{O}))$ bound in two space dimensions, products like $u_i^\varepsilon u_j^\varepsilon$ are integrable. In the deterministic case, we can exploit the L^2 bound for $\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}$ to find a bound for $u_i^\varepsilon u_j^\varepsilon$ in $L^1(0,T;L^1(\mathcal{O}))$ in any space dimension, but the limit involves an identification that we could not extend to the martingale solution concept.

On an informal level, we may state our main result as follows. We refer to Section 2 for the precise formulation.

Theorem 1 (Informal statement)**.** *Let $a_{ij} \geq 0$ satisfy the detailed-balance condition, let the stochastic diffusion σ_{ij} be Lipschitz continuous on the space of Hilbert–Schmidt operators, and let a certain interaction condition between the entropy and stochastic diffusion hold (see Assumption (A5) below). Then there exists a global nonnegative martingale solution to* (1)–(3) *in the case with self-diffusion in any space dimension and in the case without self-diffusion in at most two space dimensions.*

We discuss examples for $\sigma_{ij}(u)$ in Section 7. Here, we only remark that an admissible diffusion term is

$$\sigma_{ij}(u) = \delta_{ij}u_i^\alpha \sum_{k=1}^\infty a_k(e_k,\cdot)_U, \quad i.j = 1,\ldots,n, \tag{9}$$

where $1/2 \leq \alpha \leq 1$, δ_{ij} is the Kronecker symbol, $a_k \geq 0$ decays sufficiently fast, (e_k) is a basis of the Hilbert space U with inner product $(\cdot,\cdot)_U$.

We end this section by giving a brief overview of the state of the art for the deterministic SKT model. First existence results for the two-species model were proven under restrictive conditions on the parameters, for instance in one space dimension [26], for the triangular system with $a_{21} = 0$ [33], or for small cross-diffusion parameters, since in the latter situation the diffusion matrix becomes positive definite [17]. Amann [1] proved that a priori estimates in the $W^{1,p}(\mathcal{O})$ norm with $p > d$ are sufficient to conclude the global existence of solutions to quasilinear parabolic systems, and he applied this result to the triangular SKT system. The first global existence proof without any restriction on the parameters a_{ij} (except nonnegativity) was achieved in [22] in one space dimension. This result was generalized to several space dimensions in [9, 10] and to the whole space problem in [21]. SKT-type systems with nonlinear coefficients $A_{ij}(u)$, but still for two species, were analyzed in [15, 16]. Global existence results for SKT-type models with an arbitrary number of species and under a detailed-balance condition were first proved in [11] and later generalized in [31].

This paper is organized as follows. We present our notation and the main results in Section 2. The operators needed to define the approximative scheme are introduced in Section 3. In Section 4, the existence of solutions to a general approximative scheme is proved and the corresponding entropy inequality is derived. Theorems 3 and 4 are shown in Sections 5 and 6, respectively. Section 7 is concerned with examples for $\sigma_{ij}(u)$ satisfying our assumptions. Finally, the proofs of some auxiliary lemmas are presented in Appendix A, and Appendix B states a tightness criterion that (slightly) extends [5, Corollary 2.6].

2. Notation and main result

2.1. **Notation and stochastic framework.** Let $\mathcal{O}\subset\mathbb{R}^d$ ($d\geq 1$) be a bounded domain. The Lebesgue and Sobolev spaces are denoted by $L^p(\mathcal{O})$ and $W^{k,p}(\mathcal{O})$, respectively, where $p\in[1,\infty]$, $k\in\mathbb{N}$, and $H^k(\mathcal{O})=W^{k,2}(\mathcal{O})$. For notational simplicity, we generally do not distinguish between $W^{k,p}(\mathcal{O})$ and $W^{k,p}(\mathcal{O};\mathbb{R}^n)$. We set $H^m_N(\mathcal{O})=\{v\in H^m(\mathcal{O}):\nabla v\cdot\nu=0$ on $\partial\mathcal{O}\}$ for $m\geq 2$. If $u=(u_1,\ldots,u_n)\in X$ is some vector-valued function in the normed space X, we write $\|u\|_X^2=\sum_{i=1}^n\|u_i\|_X^2$. The inner product of a Hilbert space H is denoted by $(\cdot,\cdot)_H$, and $\langle\cdot,\cdot\rangle_{V',V}$ is the dual product between the Banach space V and its dual V'. If $F:U\to V$ is a Fréchet differentiable function between Banach spaces U and V, we write $\mathrm{D}F[v]:U\to V$ for its Fréchet derivative, for any $v\in U$.

Given two quadratic matrices $A=(A_{ij})$, $B=(B_{ij})\in\mathbb{R}^{n\times n}$, $A:B=\sum_{i,j=1}^n A_{ij}B_{ij}$ is the Frobenius matrix product, $\|A\|_F=(A:A)^{1/2}$ the Frobenius norm of A, and $\operatorname{tr}A=\sum_{i=1}^n A_{ii}$ the trace of A. The constants $C>0$ in this paper are generic and their values change from line to line.

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space endowed with a complete right-continuous filtration $\mathbb{F}=(\mathcal{F}_t)_{t\geq 0}$ and let H be a Hilbert space. Then $L^0(\Omega;H)$ consists of all measurable functions from Ω to H, and $L^2(\Omega;H)$ consists of all H-valued random variables v such that $\mathbb{E}\|v\|_H^2=\int_\Omega\|v(\omega)\|_H^2\mathbb{P}(\mathrm{d}\omega)<\infty$. Let U be a separable Hilbert space and $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of U. The space of Hilbert–Schmidt operators from U to $L^2(\mathcal{O})$ is defined by

$$\mathcal{L}_2(U;L^2(\mathcal{O}))=\bigg\{F:U\to L^2(\mathcal{O})\text{ linear, continuous}:\sum_{k=1}^\infty\|Fe_k\|_{L^2(\mathcal{O})}^2<\infty\bigg\},$$

and it is endowed with the norm $\|F\|_{\mathcal{L}_2(U;L^2(\mathcal{O}))}=(\sum_{k=1}^\infty\|Fe_k\|_{L^2(\mathcal{O})}^2)^{1/2}$.

Let $W=(W_1,\ldots,W_n)$ be an n-dimensional U-cylindrical Wiener process, taking values in the separable Hilbert space $U_0\supset U$ and adapted to the filtration $\mathbb{F}$. We can write $W_j=\sum_{k=1}^\infty e_kW_j^k$, where (W_j^k) is a sequence of independent standard one-dimensional

Brownian motions [12, Section 4.1.2]. Then $W_j(\omega)\in C^0([0,\infty);U_0)$ for a.e. ω [32, Section 2.5.1].

2.2. **Assumptions.** We impose the following assumptions:

- (A1) Domain: $\mathcal{O}\subset\mathbb{R}^d$ ($d\ge 1$) is a bounded domain with Lipschitz boundary. Let $T>0$ and set $Q_T=\mathcal{O}\times(0,T)$.
- (A2) Initial datum: $u^0=(u_1^0,\ldots,u_n^0)\in L^\infty(\Omega;L^2(\mathcal{O};\mathbb{R}^n))$ is a $\mathcal{F}_0$-measurable random variable satisfying $u^0(x)\ge 0$ for a.e. $x\in\mathcal{O}$ $\mathbb{P}$-a.s.
- (A3) Diffusion matrix: $a_{ij}\ge 0$ for $i=1,\ldots,n$, $j=0,\ldots,n$ and there exist $\pi_1,\ldots,\pi_n>0$ such that $\pi_i a_{ij}=\pi_j a_{ji}$ for all $i,j=1,\ldots,n$ (detailed-balance condition).
- (A4) Multiplicative noise: $\sigma=(\sigma_{ij})$ is an $n\times n$ matrix, where $\sigma_{ij}:L^2(\mathcal{O};\mathbb{R}^n)\to\mathcal{L}_2(U;L^2(\mathcal{O}))$ is $\mathcal{B}(L^2(\mathcal{O};\mathbb{R}^n))/\ \mathcal{B}(\mathcal{L}_2(U;L^2(\mathcal{O})))$-measurable and $\mathbb{F}$-adapted. Furthermore, there exists $C_\sigma>0$ such that for all u, $v\in L^2(\mathcal{O};\mathbb{R}^n)$,
$$\begin{aligned}\|\sigma(u)-\sigma(v)\|_{\mathcal{L}_2(U;L^2(\mathcal{O}))}&\le C_\sigma\|u-v\|_{L^2(\mathcal{O})},\\ \|\sigma(v)\|_{\mathcal{L}_2(U;L^2(\mathcal{O}))}&\le C_\sigma(1+\|v\|_{L^2(\mathcal{O})}).\end{aligned}$$
- (A5) Interaction between entropy and noise: There exists $C_h>0$ such that for all $u\in L^\infty(\mathcal{O}\times(0,T))$,
$$\bigg\{\int_0^t\sum_{k=1}^\infty\sum_{i,j=1}^n\bigg(\int_{\mathcal{O}}\frac{\partial h}{\partial u_i}(u(s))\sigma_{ij}(u(s))e_k\mathrm{d}x\bigg)^2\mathrm{d}s\bigg\}^{1/2}\le C_h\bigg(1+\int_0^t\int_{\mathcal{O}}h(u(s))\mathrm{d}x\mathrm{d}s\bigg),$$
$$\int_0^t\sum_{k=1}^\infty\int_{\mathcal{O}}\operatorname{tr}\big[(\sigma(u)e_k)^T h''(u)\sigma(u)e_k\big](s)\mathrm{d}x\mathrm{d}s\le C_h\bigg(1+\int_0^t\int_{\mathcal{O}}h(u(s))\mathrm{d}x\mathrm{d}s\bigg),$$
where h is the entropy density defined in (5).

Remark 2 (Discussion of the assumptions)**.** (A1) The Lipschitz regularity of the boundary $\partial\mathcal{O}$ is needed to apply the Sobolev and Gagliardo–Nirenberg inequalities.

- (A2) The regularity condition on u^0 can be weakened to $u^0\in L^p(\Omega;L^2(\mathcal{O};\mathbb{R}^n))$ for sufficiently large $p\ge 2$ (only depending on the space dimension); it is used to derive the higher-order moment estimates.
- (A3) The detailed-balance condition is also needed in the deterministic case to reveal the entropy structure of the system; see [11].
- (A4) The Lipschitz continuity of the stochastic diffusion $\sigma(u)$ is a standard condition for stochastic PDEs; see, e.g., [36].
- (A5) This is the most restrictive assumption. It compensates for the singularity of $(\partial h/\partial u_i)(u)=\pi_i\log u_i$ at $u_i=0$. We show in Lemma 33 that
$$\sigma_{ij}(u)(\cdot)=\frac{u_i\delta_{ij}}{1+u_i^{1/2+\eta}}\sum_{k=1}^\infty a_k(e_k,\cdot)_U$$
satisfies Assumption (A5), where $\eta>0$ and $(a_k)\in\ell^2(\mathbb{R})$. Taking into account the gradient estimate from the entropy inequality (see (6)), we can allow for more general stochastic diffusion terms like (9); see Lemma 34.

2.3. **Main results.** Let $T>0$, $m\in\mathbb{N}$ with $m>d/2+1$, and $D(L)=H^m_N(\mathcal{O})$.

Definition 1 (Martingale solution)**.** *A* martingale solution *to* (1)–(3) *is the triple* $(\widetilde{U},\widetilde{W},\widetilde{u})$ *such that* $\widetilde{U}=(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}},\widetilde{\mathbb{F}})$ *is a stochastic basis with filtration* $\widetilde{\mathbb{F}}=(\widetilde{\mathcal{F}}_t)_{t\geq 0}$*,* $\widetilde{W}$ *is an* n*-dimensional cylindrical Wiener process, and* $\widetilde{u}=(\widetilde{u}_1,\ldots,\widetilde{u}_n)$ *is a continuous* $D(L)'$*-valued* $\widetilde{\mathbb{F}}$*-adapted process such that* $\widetilde{u}_i\geq 0$ *a.e. in* $\mathcal{O}\times(0,T)$ $\widetilde{\mathbb{P}}$*-a.s.,*

$$\widetilde{u}_i\in L^0(\widetilde{\Omega};C^0([0,T];D(L)'))\cap L^0(\widetilde{\Omega};L^2(0,T;H^1(\mathcal{O}))), \tag{10}$$

the law of $\widetilde{u}_i(0)$ *is the same as for* u_i^0*, and for all* $\phi\in D(L)$*,* $t\in(0,T)$*,* $i=1,\ldots,n$*,* $\widetilde{\mathbb{P}}$*-a.s.,*

$$\begin{aligned}\langle\widetilde{u}_i(t),\phi\rangle_{D(L)',D(L)} &= \langle\widetilde{u}_i(0),\phi\rangle_{D(L)',D(L)} - \sum_{j=1}^n\int_0^t\int_{\mathcal{O}}A_{ij}(\widetilde{u}(s))\nabla\widetilde{u}_j(s)\cdot\nabla\phi\,\mathrm{d}x\mathrm{d}s \\ &\quad + \sum_{j=1}^n\int_{\mathcal{O}}\bigg(\int_0^t\sigma_{ij}(\widetilde{u}(s))\mathrm{d}\widetilde{W}_j(s)\bigg)\phi\,\mathrm{d}x.\end{aligned} \tag{11}$$

Our main results read as follows.

Theorem 3 (Existence for the SKT model with self-diffusion)**.** *Let Assumptions (A1)–(A5) be satisfied and let* $a_{ii}>0$ *for* $i=1,\ldots,n$*. Then* (1)–(3) *has a global nonnegative martingale solution in the sense of Definition 1.*

Theorem 4 (Existence for the SKT model without self-diffusion)**.** *Let Assumptions (A1)–(A5) be satisfied, let* $d\leq 2$*, and let* $a_{0i}>0$ *for* $i=1,\ldots,n$*. We strengthen Assumption (A4) slightly by assuming that for all* $v\in L^2(\mathcal{O};\mathbb{R}^n)$*,*

$$\|\sigma(v)\|_{\mathcal{L}_2(U;L^2(\mathcal{O}))}\leq C_\sigma(1+\|v\|^\gamma_{L^2(\mathcal{O})}),$$

where $\gamma<1$ *if* $d=2$ *and* $\gamma=1$ *if* $d=1$*. Then* (1)–(3) *has a global nonnegative martingale solution in the sense of Definition 1 with the exception that* (10) *and* (11) *are replaced by*

$$\widetilde{u}_i\in L^0(\widetilde{\Omega};C^0([0,T];D(L)'))\cap L^0(\widetilde{\Omega};L^2(0,T;W^{1,1}(\mathcal{O})))$$

and, for all $\phi\in D(L)\cap W^{2,\infty}(\mathcal{O})$*,*

$$\begin{aligned}\langle\widetilde{u}_i(t),\phi\rangle_{D(L)',D(L)} &= \langle\widetilde{u}_i(0),\phi\rangle_{D(L)',D(L)} - \int_0^t\int_{\mathcal{O}}\widetilde{u}_i(s)\bigg(a_{i0}+\sum_{j=1}^n a_{ij}\widetilde{u}_j(s)\bigg)\Delta\phi\,\mathrm{d}x\mathrm{d}s \\ &\quad + \sum_{j=1}^n\int_{\mathcal{O}}\bigg(\int_0^t\sigma_{ij}(\widetilde{u}(s))\mathrm{d}\widetilde{W}_j(s)\bigg)\phi\,\mathrm{d}x.\end{aligned}$$

The weak formulation for the SKT system without self-diffusion is weaker than that one with self-diffusion, since we have only the gradient regularity $\nabla\widetilde{u}_i\in L^1(\mathcal{O})$, and $A_{ij}(\widetilde{u})$ may be nonintegrable. However, system (1) can be written in Laplacian form according to (8), which allows for the "very weak" formulation stated in Theorem 4. The condition on γ if $d=2$ is needed to prove the fractional time regularity for the approximative solutions.

Remark 5 (Nonnegativity of the solution)**.** The a.s. nonnegativity of the population densities is a consequence of the entropy structure, since the approximate densities u_i^ε satisfy $u_i^\varepsilon = u_i(R_\varepsilon(v^\varepsilon)) = \exp(R_\varepsilon(v^\varepsilon)/\pi_i) > 0$ a.e. in Q_T. This may be surprising since we do not assume that the noise vanishes at zero, i.e. $\sigma_{ij}(u) = 0$ if $u_i = 0$. This condition is replaced by the weaker integrability condition for $\sigma_{ij}(u)\log u_i$ in Assumption (A5). A similar, but pointwise condition was imposed in the deterministic case; see Hypothesis (H3) in [25, Section 4.4]. The examples in Section 7 satisfy $\sigma_{ij}(u) = 0$ if $u_i = 0$. □

3. Operator setup

In this section, we introduce the operators needed to define the approximate scheme.

3.1. **Definition of the connection operator L.**

We define an operator L that "connects" two Hilbert spaces V and H satisfying $V \subset H$. This abstract operator allows us to define a regularization operator that "lifts" the dual space V' to V.

Proposition 6 (Operator L)**.** *Let V and H be separable Hilbert spaces such that the embedding $V \hookrightarrow H$ is continuous and dense. Then there exists a bounded, self-adjoint, positive operator $L : D(L) \to H$ with domain $D(L) = V$. Moreover, it holds for L and its dual operator $L^* : H \to V'$ (we identify H and its dual H') that, for some $0 < c < 1$,*

$$c\|v\|_V \le \|L(v)\|_H = \|v\|_V, \quad \|L^*(w)\|_{V'} \le \|w\|_H, \quad v \in V,\ w \in H. \tag{12}$$

We abuse slightly the notation by denoting both dual and adjoint operators by A^*. The proof is similar to [27, Theorem 1.12]. For the convenience of the reader, we present the full proof.

Proof. We first construct some auxiliary operator by means of the Riesz representation theorem. Let $w \in H$. The mapping $V \to \mathbb{R}$, $v \mapsto (v,w)_H$, is linear and bounded. Hence, there exists a unique element $\widetilde{w} \in V$ such that $(v,\widetilde{w})_V = (v,w)_H$ for all $v \in V$. This defines the linear operator $G : H \to V$, $G(w) := \widetilde{w}$, such that

$$(v,w)_H = (v,G(w))_V \quad \text{for all } v \in V,\ w \in H.$$

The operator G is bounded and symmetric, since $\|G(w)\|_V = \|\widetilde{w}\|_V = \|w\|_H$ and

$$(G(w),v)_H = (G(w),G(v))_V = (w,G(v))_H \quad \text{for all } v,w \in H. \tag{13}$$

This means that G is self-adjoint as an operator on H. Choosing $v = w \in H$ in (13) gives $(G(v),v)_H = \|G(v)\|_V^2 \ge 0$, i.e., G is positive. We claim that G is also one-to-one. Indeed, let $G(w) = 0$ for some $w \in H$. Then $0 = (v,G(w))_V = (v,w)_H$ for all $v \in V$ and, by the density of the embedding $V \hookrightarrow H$, for all $v \in H$. This implies that $w = 0$ and shows the claim.

The properties on G allow us to define $\Lambda := G^{-1} : D(\Lambda) \to H$, where $D(\Lambda) = \operatorname{ran}(G) \subset V$ and $D(\Lambda)$ denotes the domain of Λ. By definition, this operator satisfies

$$(v,\Lambda(w))_H = (v,w)_V \quad \text{for all } v \in V,\ w \in D(\Lambda).$$

Hence, for all $v,w \in D(\Lambda)$, we have $(v,\Lambda(w))_H = (v,w)_V = (\Lambda(v),w)_H$, i.e., Λ is symmetric. Since $G = G^*$, we have $D(\Lambda^*) = \operatorname{ran}(G^*) = \operatorname{ran}(G) = D(\Lambda)$ and consequently, Λ is self-adjoint. Moreover, Λ is densely defined (since $V \hookrightarrow H$ is dense). As a densely defined, self-adjoint operator, it is also closed. Finally, Λ is one-to-one and positive:

$$C\|\Lambda(v)\|_H\|v\|_V \geq \|\Lambda(v)\|_H\|v\|_H \geq (\Lambda(v),v)_H = (v,v)_V = \|v\|_V^2 \geq 0$$

for all $v \in D(\Lambda)$ and some $C>0$ and consequently, $\|\Lambda(v)\|_H \geq C^{-1}\|v\|_V$.

Therefore, we can define the square root of Λ, $\Lambda^{1/2} : D(\Lambda^{1/2}) \to H$, which is densely defined and closed. Its domain can be obtained by closing $D(\Lambda)$ with respect to

$$\|\Lambda^{1/2}(v)\|_H = (\Lambda^{1/2}(v),\Lambda^{1/2}(v))_H^{1/2} = (\Lambda(v),v)_H^{1/2} = (v,v)_V^{1/2} = \|v\|_V \tag{14}$$

for $v \in D(\Lambda^{1/2})$. In particular, the graph norm $\|\cdot\|_H + \|\Lambda^{1/2}(\cdot)\|_H$ is equivalent to the norm in V. We claim that $D(\Lambda^{1/2}) = V$. To prove this, let $w \in V$ be orthogonal to $D(\Lambda^{1/2})$. Then $(w,v)_V = 0$ for all $v \in D(\Lambda^{1/2})$ and, since $D(\Lambda) \subset D(\Lambda^{1/2})$, in particular for all $v \in D(\Lambda)$. It follows that $0 = (w,v)_V = (w,\Lambda(v))_H$ for $v \in D(\Lambda)$. Since Λ is the inverse of $G : H \to V$, we have $\operatorname{ran}(\Lambda) = H$, and it holds that $(w,\xi)_H = 0$ for all $\xi \in H$, implying that $w = 0$. This shows the claim.

Finally, we define $L := \Lambda^{1/2} : D(L) = V \to H$, which is a positive and self-adjoint operator. Estimate (14) shows that $\|L(v)\|_H = \|v\|_V$ for $v \in V$. We deduce from the equivalence between the norm in V and the graph norm of L that, for some $C>0$ and all $v \in V$,

$$\|v\|_V \leq C(\|L(v)\|_H + \|v\|_H) = C(\|L(v)\|_V + \|L^{-1}L(v)\|_H) \leq C(1+\|L^{-1}\|)\|L(v)\|_H,$$

which proves the lower bound in (12). The dual operator $L^* : H \to V'$ is bounded too, since it holds for all $w \in H$ that

$$\|L^*(w)\|_{V'} = \sup_{\|v\|_V=1} |(w,L(v))_H| \leq \sup_{\|v\|_V=1} \|w\|_H\|v\|_V = \|w\|_H.$$

This ends the proof. ☐

We apply Proposition 6 to $V = H_N^m(\mathcal{O})$ and $H = L^2(\mathcal{O})$, recalling that $H_N^m(\mathcal{O}) = \{v \in H^m(\mathcal{O}) : \nabla v \cdot \nu = 0 \text{ on } \partial\mathcal{O}\}$ and $m > d/2+1$. Then, by Sobolev's embedding, $D(L) \hookrightarrow W^{1,\infty}(\mathcal{O})$. Observe the following two properties that are used later:

$$\|L^*L(v)\|_{V'} \leq \|v\|_V, \quad \|L^*(w)\|_{V'} \leq \|w\|_H \quad \text{for all } v \in V,\ w \in H. \tag{15}$$

The following lemma is used in the proof of Proposition 15 to apply Itô's lemma.

Lemma 7 (Operator L^{-1}). *Let $L^{-1} : \operatorname{ran}(L) \to D(L)$ be the inverse of L and let $D(L^{-1}) := \overline{D(\Lambda)}$ be the closure of $D(\Lambda)$ with respect to $\|L^{-1}(\cdot)\|_H$. Then $D(L)'$ is isometric to $D(L^{-1})$. In particular, it holds that $(L^{-1}(v),L^{-1}(w))_H = (v,w)_{D(L)'}$ for all v, $w \in D(L)'$.*

Proof. The proof is essentially contained in [27, p. 136ff] and we only sketch it. Let $F \in D(L^{-1})'$. Then $|F(v)| \leq C\|L^{-1}(v)\|_H$ for all $v \in D(\Lambda)$ and, as a consequence, $|F(Lu)| \leq C\|u\|_H$ for $u = L^{-1}(v) \in D(L)$. The density of $L^{-1}(D(\Lambda))$ in H guarantees the unique representation $F(Lu) = (u,w)_H$ for some $w \in H$, and we can represent F in the form $F(v) = (L^{-1}v,w)_H = (v,L^{-1}w)_H$, where $L^{-1}w \in D(L)$. This shows that every element of $D(L^{-1})'$ can be identified with an element of $D(L)$.

Conversely, if $w \in D(L)$, we consider functionals of the type $v \mapsto (v,w)_H$ for $v \in D(\Lambda)$, which are bounded in $\|L^{-1}(\cdot)\|_H$. These functionals can be extended by continuity to functionals F belonging to $D(L^{-1})'$. The proof in [27, p. 137] shows that $\|F\|_{D(L^{-1})'} = \|w\|_{D(L)}$. We conclude that $D(L^{-1})'$ is isometric to $D(L)$. Since Hilbert spaces are reflexive, $D(L^{-1})$ is isometric to $D(L)'$. $\square$

Lemma 8 (Operator u)**.** *The mapping $u := (h')^{-1}$ from $D(L)$ to $L^\infty(\mathcal{O})$ is Fréchet differentiable and, as a mapping from $D(L)$ to $D(L)'$, monotone.*

Proof. Let $w \in D(L) \hookrightarrow L^\infty(\mathcal{O})$ (here we use $m > d/2$). Then $u(w) = (x \mapsto u(w(x))) \in L^\infty(\mathcal{O})$, showing that $u : D(L) \to L^\infty(\mathcal{O}) = (L^1(\mathcal{O}))' \hookrightarrow D(L)'$ is well defined. It follows from the mean-value theorem that for all w, $\xi \in D(L)$,

$$\|u(w+\xi) - u(w) - u'(w)\xi\|_{L^\infty(\mathcal{O})} \le C\|\xi\|_{D(L)}^2 \bigg\| \int_0^1 (1-s)u''(w+s\xi)\mathrm{d}s \bigg\|_{L^\infty(\mathcal{O})}.$$

Since u'' maps bounded sets to bounded sets, the integral is bounded. Thus, $u : D(L) \to L^\infty(\mathcal{O})$ is Fréchet differentiable. For the monotonicity, we use the convexity of h and hence the monotonicity of h':

$$\begin{aligned}\langle u(v) - u(w), v - w\rangle_{D(L)',D(L)} &= (u(v) - u(w), v - w)_{L^2(\mathcal{O})} \\ &= (u(v) - u(w), h'(u(v)) - h'(u(w)))_{L^2(\mathcal{O})} \ge 0\end{aligned}$$

for all v, $w \in D(L)$. This proves the lemma. $\square$

3.2. **Definition of the regularization operator R_ε.** First, we define another operator that maps $D(L)$ to $D(L)'$. Its inverse is the desired regularization operator.

Lemma 9 (Operator Q_ε)**.** *Let $\varepsilon > 0$ and define $Q_\varepsilon : D(L) \to D(L)'$ by $Q_\varepsilon(w) = u(w) + \varepsilon L^*Lw$, where $w \in D(L)$. Then Q_ε is Fréchet differentiable, strongly monotone, coercive, and invertible. Its Fréchet derivative $\mathrm{D}Q_\varepsilon[w](\xi) = u'(w)\xi + \varepsilon L^*L\xi$ for w, $\xi \in D(L)$ is continuous, strongly monotone, coercive, and invertible.*

Proof. The mapping Q_ε is well defined since $w \in D(L) \hookrightarrow L^\infty(\mathcal{O})$ implies that $u(w) \in L^\infty(\mathcal{O})$ and hence, $\|u(w)\|_{D(L)'} \le C\|u(w)\|_{L^1(\mathcal{O})}$ is finite. We show that Q_ε is strongly monotone. For this, let v, $w \in D(L)$ and compute

$$\begin{aligned}(16)\qquad &\langle Q_\varepsilon(v) - Q_\varepsilon(w), v - w\rangle_{D(L)',D(L)} \\ &= (u(v) - u(w), v - w)_H + \varepsilon\langle L^*L(v-w), v - w\rangle_{D(L)',D(L)} \\ &\ge \varepsilon\langle L^*L(v-w), v-w\rangle_{D(L)',D(L)} = \varepsilon\|L(v-w)\|_H^2 \ge \varepsilon c\|v-w\|_{D(L)}^2\end{aligned}$$

where we used the monotonicity of $w \mapsto u(w)$ and the lower bound in (12). The coercivity of Q_ε is a consequence of the strong monotonicity:

$$\begin{aligned}\langle Q_\varepsilon(v), v\rangle_{D(L)',D(L)} &= \langle Q_\varepsilon(v) - Q_\varepsilon(0), v - 0\rangle_{D(L)',D(L)} + \langle Q_\varepsilon(0), v\rangle_{D(L)',D(L)} \\ &\ge \varepsilon c\|v\|_{D(L)}^2 + (u(0), v)_H \ge \varepsilon c\|v\|_{D(L)}^2 - C|u(0)|\,\|v\|_{D(L)}\end{aligned}$$

for $v \in D(L)$. Based on these properties, the invertibility of Q_ε now follows from Browder's theorem [20, Theorem 6.1.21].

Next, we show the properties for $\mathrm{D}Q_\varepsilon$. The operator $\mathrm{D}Q_\varepsilon[w] : D(L) \to D(L)'$ is well defined for all $w \in D(L)$, since

$$\|u'(w)\xi\|_{D(L)'} \le C\|u'(w)\xi\|_{L^2(\mathcal{O})} \le C\|u'(w)\|_{L^2(\mathcal{O})}\|\xi\|_{L^\infty(\mathcal{O})} \le C\|u'(w)\|_{L^2(\mathcal{O})}\|\xi\|_{D(L)}$$

for all $\xi \in D(L) \hookrightarrow L^\infty(\mathcal{O})$. The strong monotonicity of $\mathrm{D}Q_\varepsilon[w]$ for $w \in D(L)$ follows from the positive semidefiniteness of $u'(w) = (h'')^{-1}(u(w))$ and the lower bound in (12):

$$\begin{aligned}&\langle \mathrm{D}Q_\varepsilon[w](\xi) - \mathrm{D}Q_\varepsilon[w](\eta), \xi - \eta\rangle_{D(L)',D(L)} \\ &\quad = (u'(w)(\xi-\eta), \xi-\eta)_H + \varepsilon\langle L^*L(\xi-\eta), \xi-\eta\rangle_{D(L)',D(L)} \\ &\quad \ge \varepsilon\|L(\xi-\eta)\|_H^2 \ge \varepsilon c\|\xi-\eta\|_{D(L)}^2\end{aligned}$$

for ξ, $\eta \in D(L)$. The choice $\eta = 0$ yields immediately the coercivity of $\mathrm{D}Q_\varepsilon[w]$. The invertibility of $\mathrm{D}Q_\varepsilon[w]$ follows again from Browder's theorem. □

Lemma 9 shows that the inverse of Q_ε exists. We set $R_\varepsilon := Q_\varepsilon^{-1} : D(L)' \to D(L)$, which is the desired regularization operator. It has the following properties.

Lemma 10 (Operator R_ε)**.** *The operator $R_\varepsilon : D(L)' \to D(L)$ is Fréchet differentiable and strictly monotone. In particular, it is Lipschitz continuous with Lipschitz constant C/ε, where $C > 0$ does not depend on ε. The Fréchet derivative equals*

$$\mathrm{D}R_\varepsilon[v] = (\mathrm{D}Q_\varepsilon[R_\varepsilon(v)])^{-1} = (u'(R_\varepsilon(v)) + \varepsilon L^*L)^{-1} \quad \text{for } v \in D(L)',$$

and it is Lipschitz continuous with constant C/ε, satisfying $\|\mathrm{D}R_\varepsilon[v](\xi)\|_{D(L)} \le \varepsilon^{-1}C\|\xi\|_{D(L)'}$ for v, $\xi \in D(L)'$.

Proof. We show first the Lipschitz continuity of R_ε. Let v_1, $v_2 \in D(L)'$. Then there exist w_1, $w_2 \in D(L)$ such that $v_1 = Q_\varepsilon(w_1)$, $v_2 = Q_\varepsilon(w_2)$. Hence, using (12) and (16),

$$\begin{aligned}\|R_\varepsilon(v_1) - R_\varepsilon(v_2)\|_{D(L)}^2 &= \|w_1 - w_2\|_{D(L)}^2 \le C\|L(w_1-w_2)\|_H^2 \\ &\le \varepsilon^{-1}C\langle Q_\varepsilon(w_1) - Q_\varepsilon(w_2), w_1 - w_2\rangle_{D(L)',D(L)} \\ &\le \varepsilon^{-1}C\|Q_\varepsilon(w_1) - Q_\varepsilon(w_2)\|_{D(L)'}\|w_1 - w_2\|_{D(L)} \\ &= \varepsilon^{-1}C\|v_1 - v_2\|_{D(L)'}\|R_\varepsilon(v_1) - R_\varepsilon(v_2)\|_{D(L)},\end{aligned}$$

proving that R_ε is Lipschitz continuous with Lipschitz constant C/ε. The Fréchet differentiability is a consequence of the inverse function theorem and $\mathrm{D}R_\varepsilon[v] = (\mathrm{D}Q_\varepsilon[R_\varepsilon(v)])^{-1}$ for $v \in D(L)'$.

We verify the strict monotonicity of R_ε. Let v, $w \in D(L)'$ with $v \neq w$. Because of the strong monotonicity of Q_ε, we have

$$\begin{aligned}\langle v - w, R_\varepsilon(v) - R_\varepsilon(w)\rangle_{D(L)',D(L)} &= \langle Q_\varepsilon(R_\varepsilon(v)) - Q_\varepsilon(R_\varepsilon(w)), R_\varepsilon(v) - R_\varepsilon(w)\rangle_{D(L)',D(L)} \\ &\ge \varepsilon^{-1}c\|R_\varepsilon(v) - R_\varepsilon(w)\|_{D(L)}^2 > 0,\end{aligned}$$

and the right-hand side vanishes only if $v = w$, since R_ε is one-to-one.

Next, we show that $\mathrm{D}R_\varepsilon[v]$ is Lipschitz continuous. Let w_1, $w_2 \in D(L)$. By Lemma 9, $\mathrm{D}Q_\varepsilon[w]$ is strongly monotone. Thus, for any $w \in D(L)$,

$$\varepsilon c\|w_1 - w_2\|_{D(L)}^2 \le \langle \mathrm{D}Q_\varepsilon[w](w_1) - \mathrm{D}Q_\varepsilon[w](w_2), w_1 - w_2\rangle_{D(L)',D(L)}$$

$$\le \|\mathrm{D}Q_\varepsilon[w](w_1) - \mathrm{D}Q_\varepsilon[w](w_2)\|_{D(L)'}\|w_1 - w_2\|_{D(L)}.$$

Let $v_1 = \mathrm{D}Q_\varepsilon[w](w_1)$ and $v_2 = \mathrm{D}Q_\varepsilon[w](w_2)$. We infer that

$$\begin{aligned}\|(\mathrm{D}Q_\varepsilon[w])^{-1}(v_1) - (\mathrm{D}Q_\varepsilon[w])^{-1}(v_2)\|_{D(L)} &= \|w_1 - w_2\|_{D(L)} \\ &\le \varepsilon^{-1}C\|\mathrm{D}Q_\varepsilon[w](w_1) - \mathrm{D}Q_\varepsilon[w](w_2)\|_{D(L)'} = \varepsilon^{-1}C\|v_1 - v_2\|_{D(L)'},\end{aligned}$$

showing the Lipschitz continuity of $(\mathrm{D}Q_\varepsilon[w])^{-1}$ and $\mathrm{D}R_\varepsilon[v] = (\mathrm{D}Q_\varepsilon[R_\varepsilon(v)])^{-1}$. Finally, choosing $w = R_\varepsilon[v]$ and $v_2 = 0$, $\|\mathrm{D}R_\varepsilon[v](v_1)\|_{D(L)} \le \varepsilon^{-1}C\|v_1\|_{D(L)'}$. □

4. Existence of approximate solutions

In the previous section, we have introduced the regularization operator $R_\varepsilon : D(L)' \to D(L)$. The entropy variable w is replaced by the regularized variable $R_\varepsilon(v)$ for $v \in D(L)'$. Setting $v = u(R_\varepsilon(v)) + \varepsilon L^*LR_\varepsilon(v)$, we consider the regularized problem

$$\mathrm{d}v = \operatorname{div}\big(B(R_\varepsilon(v))\nabla R_\varepsilon(v)\big)\mathrm{d}t + \sigma\big(u(R_\varepsilon(v))\big)\mathrm{d}W(t) \quad \text{in } \mathcal{O},\ t \in [0, T\wedge\tau), \tag{17}$$

$$v(0) = u^0 \quad \text{in } \mathcal{O}, \quad \nabla R_\varepsilon(v)\cdot\nu = 0 \quad \text{on } \partial\mathcal{O},\ t > 0, \tag{18}$$

recalling that $B(w) = A(u(w))h''(u(w))^{-1}$ for $w \in \mathbb{R}^n$.

We clarify the notion of solution to problem (17)–(18). Let $T > 0$, let τ be an $\mathbb{F}$-adapted stopping time, and let v be a continuous, $D(L)'$-valued, $\mathbb{F}$-adapted process. We call (τ, v) a *local weak solution* to (17) if

$$v(\omega,\cdot,\cdot) \in L^2([0, T\wedge\tau(\omega)); D(L)') \cap C^0([0, T\wedge\tau(\omega)); D(L)')$$

for a.e. $\omega \in \Omega$ and for all $t \in [0, T\wedge\tau)$,

$$v(t) = v(0) + \int_0^t \operatorname{div}\big(B(R_\varepsilon(v(s)))\nabla R_\varepsilon(v(s))\big)\mathrm{d}s + \int_0^t \sigma\big(u(R_\varepsilon(v(s)))\big)\mathrm{d}W(s), \tag{19}$$

$$\nabla R_\varepsilon(v)\cdot\nu = 0 \quad \text{on } \partial\mathcal{O} \quad \mathbb{P}\text{-a.s.} \tag{20}$$

It can be verified that R_ε is strongly measurable and, if v is progressively measurable, also progressively measurable. Furthermore, if w is progressively measurable then so does $u(w)$, and if $v \in C^0([0,T]; D(L)')$, we have $R_\varepsilon(v) \in C^0([0,T]; D(L))$ and $u(R_\varepsilon(v)) \in L^\infty(Q_T)$. Finally, if $v \in L^0(\Omega; L^p(0,T; D(L)'))$ for $1 \le p \le \infty$, then $\operatorname{div}(B(u(R_\varepsilon(v)))\nabla R_\varepsilon(v)) \in L^0(\Omega; L^p(0,T; D(L)')))$. Therefore, the integrals in (19) are well defined. The local weak solution is called a *global weak solution* if $\mathbb{P}(\tau = \infty) = 1$. Given $t > 0$ and a process $v \in L^2(\Omega; C^0([0,t]; D(L)'))$, we introduce the stopping time

$$\tau_R := \inf\{s \in [0,t] : \|v(s)\|_{D(L)'} > R\} \quad \text{for } R > 0.$$

This time is positive. Indeed, by Chebychev's inequality, it holds for $\delta > 0$ that

$$\mathbb{P}(\tau_R > \delta) \ge \mathbb{P}\Big(\sup_{0<t<\delta}\|v(t\wedge\tau_R)\|_{D(L)'} \le R\Big) \ge 1 - \frac{1}{R^2}\mathbb{E}\sup_{0<t<\delta}\|v(t\wedge\tau_R)\|_{D(L)'}^2.$$

Then, inserting (19) and using the properties of the operators introduced in Section 3, we can show that $\mathbb{P}(\tau_R > \delta) \ge 1 - C(\delta)$, where $C(\delta) \to 0$ as $\delta \to 0$, which proves the claim.

We impose the following general assumptions.

(H1) Entropy density: Let $\mathcal{D}\subset\mathbb{R}^n$ be a domain and let $h\in C^2(\mathcal{D};[0,\infty))$ be such that $h':\mathcal{D}\to\mathbb{R}^n$ and $h''(u)\in\mathbb{R}^{n\times n}$ for $u\in\mathcal{D}$ are invertible and there exists $C>0$ such that $|u|\le C(1+h(u))$ for all $u\in\mathcal{D}$.

(H2) Initial datum: $u^0=(u_1^0,\ldots,u_n^0)\in L^\infty(\Omega;L^2(\mathcal{O};\mathbb{R}^n))$ is $\mathcal{F}_0$-measurable satisfying $u^0(x)\in\mathcal{D}$ for a.e. $x\in\mathcal{O}$ $\mathbb{P}$-a.s.

(H3) Diffusion matrix: $A=(A_{ij})\in C^1(\overline{\mathcal{O}};\mathbb{R}^{n\times n})$ grows at most linearly and the matrix $h''(u)A(u)$ is positive semidefinite for all $u\in\mathcal{D}$.

Remark 11 (Discussion of the assumptions)**.** Hypothesis (H1) and the positive semidefiniteness condition of $h''(u)A(u)h''(u)$ in (H3) are necessary for the entropy structure of the general cross-diffusion system. The entropy density (5) with $\mathcal{D}=(0,\infty)^n$ satisfies Hypothesis (H1), and the diffusion matrix (3) fulfills (H3). The differentiability of A is needed to apply [32, Prop. 4.1.4] (stating that the assumptions of the abstract existence Theorem 4.2.2 are satisfied) and can be weakened to continuity, weak monotonicity, and coercivity conditions. The growth condition for A is technical; it guarantees that the integral formulation associated to (1) is well defined. Hypothesis (H2) guarantees that $h(u^0)$ is well defined. □

We consider general approximate stochastic cross-diffusion systems, since the existence result for (17) may be useful also for other stochastic cross-diffusion systems.

Theorem 12 (Existence of approximate solutions)**.** *Let Assumptions (A1)–(A2), (A4)–(A5), (H1)–(H3) be satisfied and let $\varepsilon>0$, $R>0$. Then problem* (17)–(18) *has a unique local solution (τ_R,v^ε).*

Proof. We want to apply Theorem 4.2.4 and Proposition 4.1.4 of [32]. To this end, we need to verify that the operator $M:D(L)'\to D(L)'$, $M(v):=\operatorname{div}(B(R_\varepsilon(v))\nabla R_\varepsilon(v))$, is Fréchet differentiable and has at most linear growth, $\mathrm{D}M[v]-cI$ is negative semidefinite for all $v\in D(L)'$ and some $c>0$, and σ is Lipschitz continuous.

By the regularity of the matrix A and the entropy density h, the operator $D(L)\to D(L)'$, $w\mapsto\operatorname{div}(B(w)\nabla w)$, is Fréchet differentiable. Then the Fréchet differentiability of R_ε (see Lemma 10) and the chain rule imply that the operator M is also Fréchet differentiable with derivative

$$\mathrm{D}M[v](\xi)=\operatorname{div}\big(\mathrm{D}B[R_\varepsilon(v)](\mathrm{D}R_\varepsilon[v](\xi))\nabla R_\varepsilon(v)\big)+\operatorname{div}\big(B(R_\varepsilon(v))\nabla\mathrm{D}R_\varepsilon[v](\xi)\big),$$

where v, $\xi\in D(L)'$. We claim that this derivative is locally bounded, i.e. if $\|v\|_{D(L)'}\le K$ then $\|\mathrm{D}M[v](\xi)\|_{D(L)'}\le C(K)\|\xi\|_{D(L)'}$. For this, we deduce from the Lipschitz continuity of R_ε (Lemma 10) and the property $u(R_\varepsilon(v))\in L^\infty(\mathcal{O})$ for $v\in D(L)'$ that

$$\|B(R_\varepsilon(v))\|_{L^\infty(\mathcal{O})}+\|\mathrm{D}B[R_\varepsilon(v)]\|_{L^\infty(\mathcal{O})}\le C(1+\|R_\varepsilon(v)\|_{D(L)})\le C(\varepsilon)(1+\|v\|_{D(L)'}),$$

where $\mathrm{D}B[R_\varepsilon(v)]$ is interpreted as a matrix. Recalling from Lemma 10 that

$$\|\mathrm{D}R_\varepsilon[v](\xi)\|_{D(L)}\le C(\varepsilon)\|\xi\|_{D(L)'}\quad\text{for all }\xi\in D(L)',$$

we obtain for $\|v\|_{D(L)'}\le K$ and $\xi\in D(L)'$:

$$\|\mathrm{D}M[v](\xi)\|_{D(L)'}\le C\big\|\mathrm{D}B[R_\varepsilon(v)](\mathrm{D}R_\varepsilon[v](\xi))\nabla R_\varepsilon(v)+B(R_\varepsilon(v))\nabla\mathrm{D}R_\varepsilon[v](\xi)\big\|_{L^1(\mathcal{O})}$$

$$\begin{aligned}
&\le C\|\mathrm{D}B[R_\varepsilon(v)](\mathrm{D}R_\varepsilon[v](\xi))\|_{L^\infty(\mathcal{O})}\|\nabla R_\varepsilon(v)\|_{L^1(\mathcal{O})} \\
&\quad + C\|B(R_\varepsilon(v))\|_{L^\infty(\mathcal{O})}\|\nabla \mathrm{D}R_\varepsilon[v](\xi)\|_{L^1(\mathcal{O})} \\
&\le C\|\mathrm{D}B[R_\varepsilon(v)]\|_{L^\infty(\mathcal{O})}\|\mathrm{D}R_\varepsilon[v](\xi)\|_{D(L)}\|R_\varepsilon(v)\|_{D(L)} \\
&\quad + C\|B(R_\varepsilon(v))\|_{L^\infty(\mathcal{O})}\|\mathrm{D}R_\varepsilon[v](\xi)\|_{D(L)} \\
&\le C(\varepsilon)(1+\|v\|_{D(L)'})\|\xi\|_{D(L)'} \le C(\varepsilon,K)\|\xi\|_{D(L)'}.
\end{aligned}$$

This proves the claim. Thus, if $\|v\|_{D(L)'} \le K$, there exists $c > 0$ such that

$$(\xi, \mathrm{D}M[v](\xi) - c\xi)_{D(L)'} \le 0 \quad \text{for } \xi \in D(L)'.$$

Moreover, by Lemma 10 again,

$$\begin{aligned}
\|M(v)\|_{D(L)'} &\le C\|B(R_\varepsilon(v))\nabla R_\varepsilon(v)\|_{L^1(\mathcal{O})} \le C\|\nabla R_\varepsilon(v)\|_{L^1(\mathcal{O})} \\
&\le C\|R_\varepsilon(v)\|_{D(L)} \le \varepsilon^{-1}C(1+\|v\|_{D(L)'}).
\end{aligned}$$

It follows from Assumption (A4) and Lemma 8 that for v, $\bar v \in D(L)'$ with $\|v\|_{D(L)'} \le K$ and $\|\bar v\|_{D(L)'} \le K$,

$$\begin{aligned}
&\|\sigma(u(R_\varepsilon(v))) - \sigma(u(R_\varepsilon(\bar v)))\|_{\mathcal{L}_2(U;D(L)')} \le C\|\sigma(u(R_\varepsilon(v))) - \sigma(u(R_\varepsilon(\bar v)))\|_{\mathcal{L}_2(U;L^2(\mathcal{O}))} \\
&\quad \le C(K)\|u(R_\varepsilon(v))) - u(R_\varepsilon(\bar v))\|_{L^2(\mathcal{O})} \\
&\quad \le C(K)\|R_\varepsilon(v) - R_\varepsilon(\bar v)\|_{D(L)} \le C(\varepsilon,K)\|v-\bar v\|_{D(L)'},
\end{aligned}$$

where $C(K)$ also depends on the $L^\infty(\mathcal{O})$ norms of $u'(R_\varepsilon(v))$ and $u'(R_\varepsilon(\bar v))$.

These estimates show that the assumptions of Theorem 4.2.4 of [32] are satisfied in the ball $\{v \in D(L)' : \|v\|_{D(L)'} \le K\}$. An inspection of the proof of that theorem, which is based on the Galerkin method and Itô's lemma, shows that *local* bounds are sufficient to conclude the existence of a *local* solution v up to the stopping time τ_R. The boundary conditions follow from $R_\varepsilon(v) \in D(L) = H^m_N(\mathcal{O})$ and the definition of the space $H^m_N(\mathcal{O})$. ☐

For the entropy estimate we need two technical lemmas whose proofs are deferred to Appendix A.

Lemma 13. *Let $w \in D(L)$, $a = (a_{ij}) \in L^1(\mathcal{O};\mathbb{R}^{n\times n})$, and $b = (b_{ij}) \in D(L)^{n\times n}$ satisfying $\mathrm{D}R_\varepsilon[w](a) = b$. Then*

$$\int_{\mathcal{O}} a : b\,\mathrm{d}x \le \int_{\mathcal{O}} \mathrm{tr}[a^T u'(w)^{-1} a]\mathrm{d}x.$$

Lemma 14. *Let $v^0 \in L^p(\Omega; L^1(\mathcal{O}))$ for some $p \ge 1$ satisfies $\mathbb{E}\int_{\mathcal{O}} h(v^0)\mathrm{d}x \le C$. Then*

$$\int_{\mathcal{O}} h(u(R_\varepsilon(v^0)))\mathrm{d}x + \frac{\varepsilon}{2}\|LR_\varepsilon(v^0)\|^2_{L^2(\mathcal{O})} \le \int_{\mathcal{O}} h(v^0)\mathrm{d}x.$$

We turn to the entropy estimate.

Proposition 15 (Entropy inequality)**.** *Let (τ_R, v^ε) be a local solution to* (17)–(18) *and set $v^R(t) = v^\varepsilon(\omega, t\wedge\tau_R(\omega))$ for $\omega \in \Omega$, $t \in (0,\tau_R(\omega))$. Then there exists a constant*

$C(u^0,T)>0$, depending on u^0 and T but not on ε and R, such that

$$\mathbb{E}\sup_{0<t<T\wedge\tau_R}\int_{\mathcal{O}}h(u^\varepsilon(t))\mathrm{d}x+\frac{\varepsilon}{2}\mathbb{E}\sup_{0<t<T\wedge\tau_R}\|Lw^\varepsilon(t)\|^2_{L^2(\mathcal{O})}$$
$$+\,\mathbb{E}\sup_{0<t<T\wedge\tau_R}\int_0^t\int_{\mathcal{O}}\nabla w^\varepsilon(s):B(w^\varepsilon(s))\nabla w^\varepsilon(s)\mathrm{d}x\mathrm{d}s\le C(u^0,T),$$

where $u^\varepsilon:=u(R_\varepsilon(v^R))$ and $w^\varepsilon:=R_\varepsilon(v^R)$.

Proof. The result follows from Itô's lemma using a regularized entropy. More precisely, we want to apply the Itô lemma in the version of [29, Theorem 3.1]. To this end, we verify the assumptions of that theorem. Basically, we need a twice differentiable function $\mathcal{H}$ on a Hilbert space H, whose derivatives satisfy some local growth conditions on H and V, where V is another Hilbert space such that the embedding $V\hookrightarrow H$ is dense and continuous. We choose $V=H=D(L)'$ and the regularized entropy

$$\mathcal{H}(v):=\int_{\mathcal{O}}h(u(R_\varepsilon(v)))\mathrm{d}x+\frac{\varepsilon}{2}\|LR_\varepsilon(v)\|^2_{L^2(\mathcal{O})},\quad v\in D(L)'. \tag{21}$$

Recall that $R_\varepsilon(v)=h'(u(R_\varepsilon(v)))$ for $v\in D(L)'$, since $u=u(w)$ is the inverse of h'. Then, in view of the regularity assumptions for h and Lemma 10, $\mathcal{H}$ is Fréchet differentiable with derivative

$$\begin{aligned}\mathrm{D}\mathcal{H}[v](\xi)&=\int_{\mathcal{O}}\big(h'(u(R_\varepsilon(v)))u'(R_\varepsilon(v))\mathrm{D}R_\varepsilon[v](\xi)+\varepsilon L\mathrm{D}R_\varepsilon[v](\xi)\cdot LR_\varepsilon(v)\big)\mathrm{d}x\\&=\big\langle(u'(R_\varepsilon(v))+\varepsilon L^*L)\mathrm{D}R_\varepsilon[v](\xi),R_\varepsilon(v)\big\rangle_{D(L)',D(L)}\\&=\big\langle\mathrm{D}Q_\varepsilon[R_\varepsilon(v)]\mathrm{D}R_\varepsilon[v](\xi),R_\varepsilon(v)\big\rangle_{D(L)',D(L)}=\langle\xi,R_\varepsilon(v)\rangle_{D(L)',D(L)},\end{aligned}$$

where v, $\xi\in D(L)'$. In other words, $\mathrm{D}\mathcal{H}[v]$ can be identified with $R_\varepsilon(v)\in D(L)$. In a similar way, we can prove that $\mathrm{D}\mathcal{H}[v]$ is Fréchet differentiable with

$$\mathrm{D}^2\mathcal{H}[v](\xi,\eta)=\langle\xi,\mathrm{D}R_\varepsilon[v](\eta)\rangle_{D(L)',D(L)}\quad\text{for } v,\ \xi,\ \eta\in D(L)'.$$

We have, thanks to the Lipschitz continuity of R_ε and $\mathrm{D}R_\varepsilon[v]$ (see Lemma 10) for all v, $\xi\in D(L)'$ with $\|v\|_{D(L)'}\le K$ for some $K>0$,

$$\begin{aligned}|\mathrm{D}\mathcal{H}[v](\xi)|&\le\|R_\varepsilon(v)\|_{D(L)}\|\xi\|_{D(L)'}\le C(\varepsilon)(1+\|v\|_{D(L)'})\|\xi\|_{D(L)'}\le C(\varepsilon,K)\|\xi\|_{D(L)'},\\|\mathrm{D}^2\mathcal{H}[v](\xi,\xi)|&\le\|\mathrm{D}R_\varepsilon[v](\xi)\|_{D(L)}\|\xi\|_{D(L)'}\le C(\varepsilon)\|\xi\|^2_{D(L)'}.\end{aligned}$$

Finally, for any $\eta\in D(L)'$, we need an estimate for the mapping $D(L)'\to\mathbb{R}$, $v\mapsto\mathrm{D}\mathcal{H}[v](\eta)$. We have identified $\mathrm{D}\mathcal{H}[v]$ with $R_\varepsilon(v)\in D(L)$, but we need an identification in $D(L)'$. As in Lemma 7, the operator L can be constructed in such a way that the Riesz representative in $D(L)'$ of a functional acting on $D(L)'$ can be expressed via the application of L^*L to an element of $D(L)$. Indeed, for $F\in D(L)$ and $\xi\in D(L)'$, we infer from Lemma 7 that

$$\begin{aligned}\langle\xi,F\rangle_{D(L)',D(L)}&=(L^{-1}\xi,LF\rangle_{D(L)',D(L)}=((LL^{-1})L^{-1}\xi,LF)_{L^2(\mathcal{O})}\\&=(L^{-1}\xi,L^{-1}L^*LF)_{L^2(\mathcal{O})}=(L^*LF,\xi)_{D(L)'}.\end{aligned}$$

Hence, we can associate $\mathrm{D}\mathcal{H}[v]$ with $L^*LR_\varepsilon(v)\in D(L)'$. Then, by the first estimate in (15) and the Lipschitz continuity of R_ε,

$$\begin{aligned}\|L^*LR_\varepsilon(v)\|_{D(L)'} &\le C\|R_\varepsilon(v)\|_{D(L)} \le C\|R_\varepsilon(v)-R_\varepsilon(0)\|_{D(L)} + C\|R_\varepsilon(0)\|_{D(L)} \\ &\le C(\varepsilon)(1+\|v\|_{D(L)'}) \quad \text{for all } v\in D(L)',\end{aligned}$$

giving the desired estimate for $\mathrm{D}\mathcal{H}[v]$ in $D(L)'$. Thus, the assumptions of the Itô lemma, as stated in [29], are satisfied.

To simplify the notation, we set $u^\varepsilon := u(R_\varepsilon(v^R))$ and $w^\varepsilon := R_\varepsilon(v^R)$ in the following. By Itô's lemma, using $\mathrm{D}\mathcal{H}[v^R]=h'(u^\varepsilon)$, $\mathrm{D}^2\mathcal{H}[v^R]=\mathrm{D}R_\varepsilon(v^R)$, we have

$$\begin{aligned}\mathcal{H}(v^R(t)) &= \mathcal{H}(v(0)) + \int_0^t \big\langle \operatorname{div}\big(B(w^\varepsilon)\nabla h'(u^\varepsilon(s))\big), w^\varepsilon(s)\big\rangle_{D(L)',D(L)}\mathrm{d}s \\ &\quad + \sum_{k=1}^\infty\sum_{i,j=1}^n\int_0^t\int_{\mathcal{O}}\frac{\partial h}{\partial u_i}(u^\varepsilon(s))\sigma_{ij}(u^\varepsilon(s))e_k\mathrm{d}x\mathrm{d}W_j^k(s) \\ &\quad + \frac12\sum_{k=1}^\infty\int_0^t\int_{\mathcal{O}}\mathrm{D}R_\varepsilon[v^R(s)]\big(\sigma(u^\varepsilon(s))e_k\big):\big(\sigma(u^\varepsilon(s))e_k\big)\mathrm{d}x\mathrm{d}s.\end{aligned}\tag{22}$$

Lemma 14 shows that the first term on the right-hand side can be estimated from above by $\int_{\mathcal{O}}h(u^0)\mathrm{d}x$. Using $w^\varepsilon=R_\varepsilon(v^R)=h'(u^\varepsilon)$ and integrating by parts, the second term on the right-hand side can be written as

$$\begin{aligned}&\int_0^t \big\langle \operatorname{div}\big(B(w^\varepsilon)\nabla h'(u^\varepsilon(s))\big), w^\varepsilon(s)\big\rangle_{D(L)',D(L)}\mathrm{d}s \\ &\quad = -\int_0^t\int_{\mathcal{O}}\nabla w^\varepsilon(s):B(w^\varepsilon)\nabla w^\varepsilon(s)\mathrm{d}x\mathrm{d}s \le 0.\end{aligned}$$

The boundary integral vanishes because of the choice of the space $D(L)=H_N^m(\mathcal{O})$. The last inequality follows from Assumption (A3), which implies that $B(w^\varepsilon)=A(u(w^\varepsilon))h''(u(w^\varepsilon))^{-1}$ is positive semidefinite.. We reformulate the last term in (22) by applying Lemma 13 with $a=\sigma(u^\varepsilon)e_k$ and $b=\mathrm{D}R_\varepsilon[v](\sigma(u^\varepsilon)e_k)$:

$$\begin{aligned}&\int_{\mathcal{O}}\mathrm{D}R_\varepsilon[v^R]\big(\sigma(u^\varepsilon)e_k\big):\big(\sigma(u^\varepsilon)e_k\big)\mathrm{d}x \\ &\quad \le \int_{\mathcal{O}}\operatorname{tr}\big[(\sigma(u^\varepsilon)e_k)^T u'(w^\varepsilon)^{-1}\sigma(u^\varepsilon)e_k\big]\mathrm{d}x.\end{aligned}$$

Taking the supremum in (22) over $(0,T_R)$, where $T_R\le T\wedge\tau_R$, and the expectation yields

$$\begin{aligned}&\mathbb{E}\sup_{0<t<T_R}\int_{\mathcal{O}}h(u^\varepsilon(t))\mathrm{d}x + \frac{\varepsilon}{2}\mathbb{E}\sup_{0<t<T_R}\|Lw^\varepsilon\|^2_{L^2(\mathcal{O})} \\ &\quad + \mathbb{E}\sup_{0<t<T_R}\int_0^t\int_{\mathcal{O}}\nabla w^\varepsilon(s):B(w^\varepsilon)\nabla w^\varepsilon(s)\mathrm{d}x\mathrm{d}s - \mathbb{E}\int_{\mathcal{O}}h(u^0)\mathrm{d}x\end{aligned}\tag{23}$$

$$
\begin{aligned}
&\le \mathbb{E}\sup_{0<t<T_R}\sum_{k=1}^\infty\sum_{i,j=1}^n\int_0^t\int_{\mathcal{O}}\frac{\partial h}{\partial u_i}(u^\varepsilon(s))\sigma_{ij}(u^\varepsilon(s))e_k\mathrm{d}x\mathrm{d}W_j^k(s) \\
&\quad+\frac{1}{2}\mathbb{E}\sup_{0<t<T_R}\sum_{k=1}^\infty\int_0^t\int_{\mathcal{O}}\operatorname{tr}\big[(\sigma(u^\varepsilon(s))e_k)^T u'(w^\varepsilon(s))^{-1}\sigma(u^\varepsilon(s))e_k\big]\mathrm{d}x\mathrm{d}s \\
&=: I_1+I_2.
\end{aligned}
$$

We apply the Burkholder–Davis–Gundy inequality [32, Theorem 6.1.2] to I_1 and use Assumption (A5):

$$
\begin{aligned}
I_1 &\le C\mathbb{E}\sup_{0<t<T_R}\bigg\{\int_0^t\sum_{k=1}^\infty\sum_{i,j=1}^n\bigg(\int_{\mathcal{O}}\frac{\partial h}{\partial u_i}(u^\varepsilon(s))\sigma_{ij}(u^\varepsilon(s))e_k\mathrm{d}x\bigg)^2\mathrm{d}s\bigg\}^{1/2} \\
&\le C\mathbb{E}\sup_{0<t<T_R}\bigg(1+\int_0^t\int_{\mathcal{O}}h(u^\varepsilon(s))\mathrm{d}x\mathrm{d}s\bigg).
\end{aligned}
$$

Also the remaining integral I_2 can be bounded from above by Assumption (A5):

$$
I_2 \le C\mathbb{E}\sup_{0<t<T_R}\bigg(1+\int_0^t\int_{\mathcal{O}}h(u^\varepsilon(s))\mathrm{d}x\mathrm{d}s\bigg).
$$

Therefore, (23) becomes

$$
\begin{aligned}
&\mathbb{E}\sup_{0<t<T_R}\int_{\mathcal{O}}h(u^\varepsilon(t))\mathrm{d}x+\frac{\varepsilon}{2}\mathbb{E}\sup_{0<t<T_R}\|Lw^\varepsilon\|_{L^2(\mathcal{O})}^2 \\
&\quad+\mathbb{E}\sup_{0<t<T_R}\int_0^t\int_{\mathcal{O}}\nabla w^\varepsilon(s):B(w^\varepsilon)\nabla w^\varepsilon(s)\mathrm{d}x\mathrm{d}s-\mathbb{E}\int_{\mathcal{O}}h(u^0)\mathrm{d}x \\
&\le C\mathbb{E}\sup_{0<t<T_R}\bigg(1+\int_0^t\int_{\mathcal{O}}h(u^\varepsilon(s))\mathrm{d}x\mathrm{d}s\bigg) \\
&\le C+C\mathbb{E}\int_0^{T_R}\int_{\mathcal{O}}\sup_{0<s<t}h(u^\varepsilon(s))\mathrm{d}x\mathrm{d}t.
\end{aligned}
\tag{24}
$$

We apply Gronwall's lemma to the function $F(t)=\sup_{0<s<t}\int_{\mathcal{O}}h(u^\varepsilon(s))\mathrm{d}x$ to find that

$$
\mathbb{E}\sup_{0<t<T_R}\int_{\mathcal{O}}h(u^\varepsilon(t))\mathrm{d}x\le C(u^0,T).
$$

Using this bound in (24) then finishes the proof. □

The entropy inequality allows us to extend the local solution to a global one.

Proposition 16. *Let (τ_R,v^ε) be a local solution to* (19)–(20)*, constructed in Theorem 12. Then v^ε can be extended to a global solution to* (19)–(20).

Proof. With the notation $u^\varepsilon=u(R_\varepsilon(v^\varepsilon))$ and $w^\varepsilon=R_\varepsilon(v^\varepsilon)$, we observe that $v^\varepsilon=Q_\varepsilon(R_\varepsilon(v^\varepsilon))$ $=u(R_\varepsilon(v^\varepsilon))+\varepsilon L^*LR_\varepsilon(v^\varepsilon)=u^\varepsilon+\varepsilon L^*Lw^\varepsilon$. Thus, we have for $T_R\le T\wedge\tau_R$,

$$
\mathbb{E}\sup_{0<t<T_R}\|v^\varepsilon(t)\|_{D(L)'}\le\mathbb{E}\sup_{0<t<T_R}\|u^\varepsilon\|_{D(L)'}+\varepsilon\mathbb{E}\sup_{0<t<T_R}\|L^*Lw^\varepsilon(t)\|_{D(L)'}
$$

$$\le C\mathbb{E}\sup_{0<t<T_R}\|u^\varepsilon\|_{L^1(\mathcal{O})}+\varepsilon\mathbb{E}\sup_{0<t<T_R}\|L^*Lw^\varepsilon(t)\|_{D(L)'}.$$

We know from Hypothesis (H1) that $|u^\varepsilon|\le C(1+h(u^\varepsilon))$. Therefore, taking into account the entropy inequality and the second inequality in (15),

$$\mathbb{E}\sup_{0<t<T_R}\|v(t)\|_{D(L)'}\le C\mathbb{E}\sup_{0<t<T_R}\|h(u^\varepsilon(t))\|_{L^1(\mathcal{O})}+\varepsilon C\sup_{0<t<T_R}\|Lw^\varepsilon(t)\|_{L^2(\mathcal{O})}\le C(u^0,T).$$

This allows us to perform the limit $R\to\infty$ and to conclude that we have indeed a solution v^ε in $(0,T)$ for any $T>0$. □

5. Proof of Theorem 3

We prove the global existence of martingale solutions to the SKT model with self-diffusion.

5.1. **Uniform estimates.** Let v^ε be a global solution to (19)–(20) and set $u^\varepsilon=u(R_\varepsilon(v^\varepsilon))$. We assume that $A(u)$ is given by (3) and that $a_{ii}>0$ for $i=1,\ldots,n$. We start with some uniform estimates, which are a consequence of the entropy inequality in Proposition 15.

Lemma 17 (Uniform estimates)**.** *There exists a constant $C(u^0,T)>0$ such that for all $\varepsilon>0$ and $i,j=1,\ldots,n$ with $i\neq j$,*

$$\mathbb{E}\|u_i^\varepsilon\|_{L^\infty(0,T;L^1(\mathcal{O}))}\le C(u^0,T), \tag{25}$$

$$a_{i0}^{1/2}\mathbb{E}\|(u_i^\varepsilon)^{1/2}\|_{L^2(0,T;H^1(\mathcal{O}))}+a_{ii}^{1/2}\mathbb{E}\|u_i^\varepsilon\|_{L^2(0,T;H^1(\mathcal{O}))}\le C(u^0,T), \tag{26}$$
$$a_{ij}^{1/2}\mathbb{E}\|\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^2(0,T;L^2(\mathcal{O}))}\le C(u^0,T).$$

Moreover, we have the estimate

$$\varepsilon\mathbb{E}\|LR_\varepsilon(v^\varepsilon)\|^2_{L^\infty(0,T;L^2(\mathcal{O}))}+\mathbb{E}\|v^\varepsilon\|^2_{L^\infty(0,T;D(L)')}\le C(u^0,T). \tag{27}$$

Proof. Let v^ε be a global solution to (19)–(20). We observe that $R_\varepsilon(v^\varepsilon)=h'(u(R_\varepsilon(v^\varepsilon)))=h'(u^\varepsilon)$ implies that $\nabla R_\varepsilon(v^\varepsilon)=h''(u^\varepsilon)\nabla u^\varepsilon$. It is shown in [11, Lemma 4] that for all $z\in\mathbb{R}^n$ and $u\in(0,\infty)^n$,

$$z^Th''(u)A(u)z\ge\sum_{i=1}^n\pi_i\bigg(a_{0i}\frac{z_i^2}{u_i}+2a_{ii}z_i^2\bigg)+\frac12\sum_{i,j=1,\,i\neq j}^n\pi_ia_{ij}\bigg(\sqrt{\frac{u_j}{u_i}}z_i+\sqrt{\frac{u_i}{u_j}}z_j\bigg)^2.$$

Using $B(R_\varepsilon(v^\varepsilon))=A(u^\varepsilon)h''(u^\varepsilon)^{-1}$ and the previous inequality with $z=\nabla u^\varepsilon$, we find that

$$\begin{aligned}\nabla R_\varepsilon(v^\varepsilon):B(R_\varepsilon(v^\varepsilon))\nabla R_\varepsilon(v)&=\nabla u^\varepsilon:h''(u^\varepsilon)\big(A(u^\varepsilon)h''(u^\varepsilon)^{-1}\big)h''(u^\varepsilon)\nabla u^\varepsilon\\&=\nabla u^\varepsilon:h''(u^\varepsilon)A(u^\varepsilon)\nabla u^\varepsilon\\&\ge\sum_{i=1}^n\pi_i\big(4a_{0i}|\nabla(u^\varepsilon)^{1/2}|^2+2a_{ii}|\nabla u^\varepsilon|^2\big)+2\sum_{i\neq j}\pi_ia_{ij}|\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}|^2.\end{aligned} \tag{28}$$

Therefore, the entropy inequality in Proposition 15 becomes

$$\mathbb{E}\sup_{0<t<T}\int_{\mathcal{O}}h(u^\varepsilon(t))\mathrm{d}x+\mathbb{E}\sup_{0<t<T}\frac{\varepsilon}{2}\|LR(v^\varepsilon(t))\|^2_{L^2(\mathcal{O})} \tag{29}$$

$$+ \mathbb{E}\int_0^T \int_{\mathcal{O}} \sum_{i=1}^n \pi_i \big(4a_{0i}|\nabla(u^\varepsilon)^{1/2}|^2 + 2a_{ii}|\nabla u^\varepsilon|^2\big)\mathrm{d}x\mathrm{d}s$$
$$+ 2\mathbb{E}\int_0^T \int_{\mathcal{O}} \sum_{i\neq j} \pi_i a_{ij}|\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}|^2\mathrm{d}x\mathrm{d}s \le C(u^0,T).$$

This is the stochastic analog of the entropy inequality (6). By Hypothesis (H1), we have $|u| \le C(1+h(u))$ and consequently,

$$\mathbb{E}\sup_{0<t<T}\|u^\varepsilon(t)\|_{L^1(\mathcal{O})} \le C\mathbb{E}\sup_{0<t<T}\int_{\mathcal{O}} h(u^\varepsilon(t))\mathrm{d}x + C \le C(u^0,T),$$

which proves (25). Estimate (26) then follows from the Poincaré–Wirtinger inequality.

It remains to show estimate (27). We deduce from the second inequality in (15) that

$$\begin{aligned}\|v^\varepsilon(t)\|_{D(L)'} &= \|Q_\varepsilon(R_\varepsilon(v^\varepsilon(t)))\|_{D(L)'} = \|u(R_\varepsilon(v^\varepsilon(t))) + \varepsilon L^*LR_\varepsilon(v^\varepsilon(t))\|_{D(L)'} \\ &\le C\|u(R_\varepsilon(v^\varepsilon(t)))\|_{L^1(\mathcal{O})} + \varepsilon\|L^*LR_\varepsilon(v^\varepsilon(t))\|_{D(L)'} \\ &\le C\|u^\varepsilon(t)\|_{L^1(\mathcal{O})} + \varepsilon C\|LR_\varepsilon(v^\varepsilon(t))\|_{L^2(\mathcal{O})}.\end{aligned}$$

This shows that

$$\mathbb{E}\sup_{0<t<T}\|v^\varepsilon(t)\|_{D(L)'} \le C\mathbb{E}\sup_{0<t<T}\|u^\varepsilon\|_{L^1(\mathcal{O})} + \varepsilon C\mathbb{E}\sup_{0<t<T}\|LR_\varepsilon(v^\varepsilon(t))\|_{L^2(\mathcal{O})} \le C(u^0,T),$$

ending the proof. ☐

We also need higher-order moment estimates.

Lemma 18 (Higher-order moments I)**.** *Let $p \ge 2$. There exists a constant $C(p,u^0,T)$, which is independent of ε, such that*

$$\mathbb{E}\|u^\varepsilon\|^p_{L^\infty(0,T;L^1(\mathcal{O}))} \le C(p,u^0,T), \tag{30}$$
$$a_{i0}^{p/2}\mathbb{E}\|(u_i^\varepsilon)^{1/2}\|^p_{L^2(0,T;H^1(\mathcal{O}))} + a_{ii}^{p/2}\mathbb{E}\|u_i^\varepsilon\|^p_{L^2(0,T;H^1(\mathcal{O}))} \le C(p,u^0,T), \tag{31}$$
$$a_{ij}^{p/2}\mathbb{E}\|\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|^p_{L^2(0,T;L^2(\mathcal{O}))} \le C(p,u^0,T). \tag{32}$$

Moreover, we have

$$\mathbb{E}\bigg(\varepsilon\sup_{0<t<T}\|LR_\varepsilon(v^\varepsilon(t))\|^2_{L^2(\mathcal{O})}\bigg)^p + \mathbb{E}\bigg(\sup_{0<t<T}\|v^\varepsilon(t)\|_{D(L)'}\bigg)^p \le C(p,u^0,T). \tag{33}$$

Proof. Proceeding as in the proof of Proposition 15 and taking into account identity (22) and inequality (28), we obtain

$$\mathcal{H}(v^\varepsilon(t)) + \int_0^T \int_{\mathcal{O}} \sum_{i=1}^n \pi_i\big(4a_{i0}|\nabla(u^\varepsilon)^{1/2}|^2 + 2a_{ii}|\nabla u^\varepsilon|^2\big)\mathrm{d}x\mathrm{d}s$$
$$+ 2\mathbb{E}\int_0^T \int_{\mathcal{O}} \sum_{i\neq j} \pi_i a_{ij}|\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}|^2\mathrm{d}x\mathrm{d}s$$

$$\leq \mathcal{H}(v^\varepsilon(0)) + \sum_{k=1}^\infty \sum_{i,j=1}^n \int_0^t \int_{\mathcal{O}} \pi_i \log u_i^\varepsilon(s)\sigma_{ij}(u^\varepsilon(s))e_k \mathrm{d}x\mathrm{d}W_j^k(s)$$
$$+ \frac{1}{2}\sum_{k=1}^\infty \int_0^t \int_{\mathcal{O}} \operatorname{tr}\big[(\sigma(u^\varepsilon(s))e_k)^T h''(u^\varepsilon(s))\sigma(u^\varepsilon(s))e_k\big]\mathrm{d}x\mathrm{d}s,$$

recalling Definition 21 of $\mathcal{H}(v^\varepsilon)$. We raise this inequality to the pth power, take the expectation, apply the Burkholder–Davis–Gundy inequality (for the second term on the right-hand side), and use Assumption (A5) to find that

$$\begin{aligned}
(34)\qquad & \mathbb{E}\bigg(\sup_{0<t<T}\int_{\mathcal{O}} h(u^\varepsilon(t))\mathrm{d}x + \varepsilon \sup_{0<t<T}\|LR_\varepsilon(v^\varepsilon(t))\|_{L^2(\mathcal{O})}^2\bigg)^p \\
&\quad + C\mathbb{E}\bigg(\int_0^T\int_{\mathcal{O}}\sum_{i=1}^n \pi_i a_{i0}|\nabla(u_i^\varepsilon(s))^{1/2}|^2\mathrm{d}x\mathrm{d}s\bigg)^p \\
&\quad + C\mathbb{E}\bigg(\int_0^T\int_{\mathcal{O}}\sum_{i=1}^n \pi_i a_{ii}|\nabla u_i^\varepsilon(s)|^2\mathrm{d}x\mathrm{d}s\bigg)^p \\
&\quad + C\mathbb{E}\bigg(\int_0^T\int_{\mathcal{O}}\sum_{i\neq j} \pi_i a_{ij}|\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}|^2\mathrm{d}x\mathrm{d}s\bigg)^p \\
&\leq C(p,u^0) + C\mathbb{E}\bigg(\int_0^T \sum_{k=1}^\infty\sum_{i,j=1}^n\bigg(\int_{\mathcal{O}}\log u_i^\varepsilon(s)\sigma_{ij}(u^\varepsilon(s))e_k\mathrm{d}x\bigg)^2\mathrm{d}s\bigg)^{p/2} \\
&\quad + C\mathbb{E}\bigg(\int_0^T\sum_{k=1}^\infty\int_{\mathcal{O}}\operatorname{tr}\big[(\sigma(u^\varepsilon(s))e_k)^T h''(u^\varepsilon(s))(\sigma(u^\varepsilon(s))e_k)\big]\mathrm{d}x\mathrm{d}s\bigg)^p \\
&\leq C(p,u^0) + C\mathbb{E}\bigg(\int_0^T\int_{\mathcal{O}} h(u^\varepsilon(s))\mathrm{d}x\mathrm{d}s\bigg)^p.
\end{aligned}$$

We neglect the expression $\varepsilon\|LR_\varepsilon(v^\varepsilon(t))\|_{L^2(\mathcal{O})}^2$ and apply Gronwall's lemma. Then, taking into account the fact that the entropy dominates the $L^1(\mathcal{O})$ norm, thanks to Hypothesis (H1), and applying the Poincaré–Wirtinger inequality, we obtain estimates (30)–(32). Going back to (34), we infer that

$$\begin{aligned}
\mathbb{E}\bigg(\varepsilon\sup_{0<t<T}\|LR_\varepsilon(v^\varepsilon(t))\|_{L^2(\mathcal{O})}^2\bigg)^p &\leq C(p,u^0) + C(p,T)\mathbb{E}\int_0^T\bigg(\int_{\mathcal{O}} h(u^\varepsilon(s))\mathrm{d}x\bigg)^p\mathrm{d}s \\
&\leq C(p,u^0,T).
\end{aligned}$$

Combining the previous estimates and arguing as in the proof of Lemma 17, we have

$$\begin{aligned}
\mathbb{E}\bigg(\sup_{0<t<T}\|v^\varepsilon(t)\|_{D(L)'}\bigg)^p &= \mathbb{E}\bigg(\sup_{0<t<T}\|u^\varepsilon(t)+\varepsilon L^*LR_\varepsilon(v^\varepsilon(t))\|_{D(L)'}\bigg)^p \\
&\leq C\mathbb{E}\bigg(\sup_{0<t<T}\|u^\varepsilon(t)\|_{L^1(\mathcal{O})}\bigg)^p + C\mathbb{E}\bigg(\varepsilon^2\sup_{0<t<T}\|LR_\varepsilon(v^\varepsilon(t))\|_{L^2(\mathcal{O})}^2\bigg)^{p/2} \leq C(p,u^0,T).
\end{aligned}$$

This ends the proof. □

Using the Gagliardo–Nirenberg inequality, we can derive further estimates. We recall that $Q_T = \mathcal{O}\times(0,T)$.

Lemma 19 (Higher-order moments II)**.** *Let $p\geq 2$. There exists a constant $C(p,u^0,T) > 0$, which is independent of ε, such that*

$$\mathbb{E}\|u_i^\varepsilon\|^p_{L^{2+2/d}(Q_T)} \leq C(p,u^0,T), \tag{35}$$

$$\mathbb{E}\|u_i^\varepsilon\|^p_{L^{2+4/d}(0,T;L^2(\mathcal{O}))} \leq C(p,u^0,T). \tag{36}$$

Proof. We apply the Gagliardo–Nirenberg inequality:

$$\begin{aligned}\mathbb{E}\bigg(\int_0^T\|u_i^\varepsilon\|^s_{L^r(\mathcal{O})}\mathrm{d}t\bigg)^{p/s} &\leq C\mathbb{E}\bigg(\int_0^T\|u_i^\varepsilon\|^{\theta s}_{H^1(\mathcal{O})}\|u_i^\varepsilon\|^{(1-\theta)s}_{L^1(\mathcal{O})}\mathrm{d}t\bigg)^{p/s} \\ &\leq C\mathbb{E}\bigg(\|u_i^\varepsilon\|^{(1-\theta)s}_{L^\infty(0,T;L^1(\mathcal{O}))}\int_0^T\|u_i^\varepsilon\|^2_{H^1(\mathcal{O})}\mathrm{d}t\bigg)^{p/s} \\ &\leq C\big(\mathbb{E}\|u_i^\varepsilon\|^{2(1-\theta)p}_{L^\infty(0,T;L^1(\mathcal{O}))}\big)^{1/2}\big(\mathbb{E}\|u_i^\varepsilon\|^{4p/s}_{L^2(0,T;H^1(\mathcal{O}))}\big)^{1/2}\leq C,\end{aligned}$$

where $r>1$ and $\theta\in(0,1]$ are related by $1/r = 1-\theta(d+2)/(2d)$ and $s=2/\theta\geq 2$. The right-hand side is bounded in view of estimates (30) and (31). Estimate (35) follows after choosing $r=s$, implying that $r=2+2/d$, and (36) follows from the choice $s=2+4/d$, implying that $r=2$. □

Next, we show some bounds for the fractional time derivative of u^ε. This result is used to establish the tightness of the laws of (u^ε) in a sub-Polish space. Alternatively, the tightness property can be proved by verifying the Aldous condition; see, e.g., [18]. We recall the definition of the Sobolev–Slobodeckij spaces. Let X be a vector space and let $p\geq 1$, $\alpha\in(0,1)$. Then $W^{\alpha,p}(0,T;X)$ is the set of all functions $v\in L^p(0,T;X)$ for which

$$\begin{aligned}\|v\|^p_{W^{\alpha,p}(0,T;X)} &= \|v\|^p_{L^p(0,T;X)} + |v|^p_{W^{\alpha,p}(0,T;X)} \\ &= \int_0^T\|v\|_X^p\mathrm{d}t + \int_0^T\int_0^T\frac{\|v(t)-v(s)\|_X^p}{|t-s|^{1+\alpha p}}\mathrm{d}t\mathrm{d}s<\infty.\end{aligned}$$

With this norm, $W^{\alpha,p}(0,T;X)$ becomes a Banach space. We need the following technical lemma, which is proved in Appendix A.

Lemma 20. *Let $g\in L^1(0,T)$ and $\delta<2$, $\delta\neq 1$. Then*

$$\int_0^T\int_0^T|t-s|^{-\delta}\int_{s\wedge t}^{t\vee s}g(r)\mathrm{d}r\mathrm{d}t\mathrm{d}s<\infty. \tag{37}$$

We obtain the following uniform bounds for u^ε and v^ε in Sobolev–Slobodeckij spaces.

Lemma 21 (Fractional time regularity)**.** *Let $\alpha<1/2$. There exists a constant $C(u^0,T)>0$ such that, for $p:=(2d+4)/d>2$,*

$$\mathbb{E}\|u^\varepsilon\|^p_{W^{\alpha,p}(0,T;D(L)')}\leq C(u^0,T),$$

$$\varepsilon^p \mathbb{E}\|L^*LR_\varepsilon(v^\varepsilon)\|^p_{W^{\alpha,p}(0,T;D(L)')} + \mathbb{E}\|v^\varepsilon\|^p_{W^{\alpha,p}(0,T;D(L)')} \le C(u^0,T). \tag{38}$$

Since $p > 2$, we can choose $\alpha < 1/2$ such that $\alpha p > 1$. Then the continuous embedding $W^{\alpha,p}(0,T) \hookrightarrow C^{0,\beta}([0,T])$ for $\beta = \alpha - 1/p > 0$ implies that

$$\mathbb{E}\|u^\varepsilon\|^p_{C^{0,\beta}([0,T];D(L)')} \le C(u^0,T). \tag{39}$$

Proof. First, we derive the $W^{\alpha,p}$ estimate for v^ε and then we conclude the estimate for u^ε from the definition $v^\varepsilon = u^\varepsilon + \varepsilon L^*LR_\varepsilon(v^\varepsilon)$ and Lemma 19. Equation (17) reads in terms of u^ε as

$$\mathrm{d}v_i^\varepsilon = \operatorname{div}\bigg(\sum_{j=1}^n A_{ij}(u^\varepsilon)\nabla u_j^\varepsilon\bigg)\mathrm{d}t + \sum_{j=1}^n \sigma_{ij}(u^\varepsilon)\mathrm{d}W_j, \quad i=1,\ldots,n.$$

We know from (33) that $\mathbb{E}\|v^\varepsilon\|^p_{L^\infty(0,T;D(L)')}$ is bounded. Thus, to prove the bound for the second term in (38), it remains to estimate the following seminorm:

$$\begin{aligned}
\mathbb{E}|v_i^\varepsilon|^p_{W^{\alpha,p}(0,T;D(L)'} &= \mathbb{E}\int_0^T\int_0^T \frac{\|v_i^\varepsilon(t)-v_i^\varepsilon(s)\|^p_{D(L)'}}{|t-s|^{1+\alpha p}}\mathrm{d}t\mathrm{d}s \\
&\le \mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg\|\int_{s\wedge t}^{t\vee s}\operatorname{div}\sum_{j=1}^n A_{ij}(u^\varepsilon(r))\nabla u_j^\varepsilon(r)\mathrm{d}r\bigg\|^p_{D(L)'}\mathrm{d}t\mathrm{d}s \\
&\quad + \mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg\|\int_{s\wedge t}^{t\vee s}\sum_{j=1}^n \sigma_{ij}(u^\varepsilon(r))\mathrm{d}W_j(r)\bigg\|^p_{D(L)'}\mathrm{d}t\mathrm{d}s \\
&=: J_1 + J_2.
\end{aligned}$$

We need some preparations before we can estimate J_1. We observe that

$$\begin{aligned}
\bigg\|\sum_{j=1}^n A_{ij}(u^\varepsilon)\nabla u_j^\varepsilon\bigg\|_{L^1(\mathcal{O})} &= \bigg\|\bigg(a_{i0} + 2\sum_{j=1}^n a_{ij}u_j^\varepsilon\bigg)\nabla u_i^\varepsilon + \sum_{j\neq i} a_{ij}u_i^\varepsilon\nabla u_j^\varepsilon\bigg\|_{L^1(\mathcal{O})} \\
&\le C\|\nabla u_i^\varepsilon\|_{L^1(\mathcal{O})} + C\|u^\varepsilon\|_{L^2(\mathcal{O})}\|\nabla u^\varepsilon\|_{L^2(\mathcal{O})}.
\end{aligned}$$

It follows from the embedding $L^1(\mathcal{O}) \hookrightarrow D(L)'$ that

$$\begin{aligned}
J_1 &\le \mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg(\int_{s\wedge t}^{t\vee s}\bigg\|\operatorname{div}\sum_{j=1}^n A_{ij}(u^\varepsilon(r))\nabla u_j^\varepsilon(r)\bigg\|_{D(L)'}\mathrm{d}r\bigg)^p\mathrm{d}t\mathrm{d}s \\
&\le C\mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg(\int_{s\wedge t}^{t\vee s}\bigg\|\sum_{j=1}^n A_{ij}(u^\varepsilon(r))\nabla u_j^\varepsilon(r)\bigg\|_{L^1(\mathcal{O})}\mathrm{d}r\bigg)^p\mathrm{d}t\mathrm{d}s \\
&\le C\mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg(\int_{s\wedge t}^{t\vee s}\|\nabla u^\varepsilon(r)\|_{L^2(\mathcal{O})}\mathrm{d}r\bigg)^p\mathrm{d}t\mathrm{d}s \\
&\quad + C\mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg(\int_{s\wedge t}^{t\vee s}\|u^\varepsilon(r)\|_{L^2(\mathcal{O})}\|\nabla u^\varepsilon(r)\|_{L^2(\mathcal{O})}\mathrm{d}r\bigg)^p\mathrm{d}t\mathrm{d}s \\
&=: J_{11} + J_{12}.
\end{aligned}$$

We use Hölder's inequality and fix $p=(2d+4)/d$ to obtain

$$J_{11}\le C\mathbb{E}\int_0^T\int_0^T|t-s|^{-1-\alpha p}|t-s|^{p/2}\bigg(\int_{s\wedge t}^{t\vee s}\|\nabla u^\varepsilon(r)\|_{L^2(\mathcal{O})}^2\mathrm{d}r\bigg)^{p/2}\mathrm{d}t\mathrm{d}s.$$

In view of estimate (31) and (37), the right-hand side is finite if $1+\alpha p-p/2<2$ or, equivalently, $\alpha<(d+1)/(d+2)$, and this holds true since $\alpha<1/2$. Applying Hölder's inequality again, we have

$$\begin{aligned}
J_{12}&\le C\mathbb{E}\int_0^T\int_0^T|t-s|^{-1-\alpha p}\bigg(\int_{s\wedge t}^{t\vee s}\|u^\varepsilon(r)\|_{L^2(\mathcal{O})}^2\mathrm{d}r\bigg)^{p/2}\bigg(\int_{s\wedge t}^{t\vee s}\|\nabla u^\varepsilon(r)\|_{L^2(\mathcal{O})}^2\mathrm{d}r\bigg)^{p/2}\mathrm{d}t\mathrm{d}s\\
&\le C\mathbb{E}\int_0^T\int_0^T|t-s|^{-1-\alpha p}|t-s|^{p/(d+2)}\bigg(\int_{s\wedge t}^{t\vee s}\|u^\varepsilon(r)\|_{L^2(\mathcal{O})}^{(2d+4)/d}\mathrm{d}r\bigg)^{pd/(2d+4)}\\
&\quad\times\bigg(\int_{s\wedge t}^{t\vee s}\|\nabla u^\varepsilon(r)\|_{L^2(\mathcal{O})}^2\mathrm{d}r\bigg)^{p/2}\mathrm{d}t\mathrm{d}s\\
&\le C\bigg\{\mathbb{E}\bigg(\int_0^T\int_0^T|t-s|^{-1-\alpha p+p/(d+2)}\bigg(\int_{s\wedge t}^{t\vee s}\|u^\varepsilon(r)\|_{L^2(\mathcal{O})}^{(2d+4)/d}\mathrm{d}r\bigg)\mathrm{d}t\mathrm{d}s\bigg)^2\bigg\}^{1/2}\\
&\quad\times\bigg\{\mathbb{E}\bigg(\int_0^T\|\nabla u^\varepsilon(r)\|_{L^2(\mathcal{O})}^2\mathrm{d}r\bigg)^p\bigg\}^{1/2}.
\end{aligned}$$

Because of estimates (31), (36), and (37), the right-hand side of is finite if $1+\alpha p-p/(d+2)<2$, which is equivalent to $\alpha<1/2$.

To estimate J_2, we use the embedding $L^2(\mathcal{O})\hookrightarrow D(L)'$, the Burkholder–Davis–Gundy inequality, the linear growth of σ from Assumption (A4), and the Hölder inequality:

$$\begin{aligned}
J_2&\le C\int_0^T\int_0^T|t-s|^{-1-\alpha p}\mathbb{E}\bigg\|\int_{s\wedge t}^{t\vee s}\sum_{j=1}^n\sigma_{ij}(u^\varepsilon(r))\mathrm{d}W_j(r)\bigg\|_{L^2(\mathcal{O})}^p\mathrm{d}t\mathrm{d}s\\
&\le C\int_0^T\int_0^T|t-s|^{-1-\alpha p}\mathbb{E}\bigg(\int_{s\wedge t}^{t\vee s}\sum_{k=1}^\infty\sum_{j=1}^n\|\sigma_{ij}(u^\varepsilon(r))e_k\|_{L^2(\mathcal{O})}^2\mathrm{d}r\bigg)^{p/2}\mathrm{d}t\mathrm{d}s\\
&\le C\int_0^T\int_0^T|t-s|^{-1-\alpha p+(p-2)/2}\int_{s\wedge t}^{t\vee s}\mathbb{E}\sum_{j=1}^n\big(1+\|u_j^\varepsilon(r)\|_{L^2(\mathcal{O})}^p\big)\mathrm{d}r\mathrm{d}t\mathrm{d}s.
\end{aligned}$$

By (36) and (37), the right-hand side is finite if $1+\alpha p-(p-2)/2<2$, which is equivalent to $\alpha<(3d+2)/(2d+4)$, and this is valid due to the condition $\alpha<1/2$. We conclude that (v^ε) is bounded in $L^p(\Omega;W^{\alpha,p}(0,T;D(L)'))$ with $p=(2d+4)/d$.

Next, we derive the uniform bounds for u^ε. By definition of v^ε and the $W^{\alpha,p}$ seminorm,

$$\begin{aligned}
\mathbb{E}|u^\varepsilon|_{W^{\alpha,p}(0,T;D(L)')}^p&=\mathbb{E}|v^\varepsilon-\varepsilon L^*LR_\varepsilon(v^\varepsilon)|_{W^{\alpha,p}(0,T;D(L)')}^p\\
&\le C\mathbb{E}\int_0^T\int_0^T\frac{\|v^\varepsilon(t)-v^\varepsilon(s)\|_{D(L)'}^p}{|t-s|^{1+\alpha p}}\mathrm{d}t\mathrm{d}s
\end{aligned}$$

$$+ C\mathbb{E}\int_0^T\int_0^T \frac{\varepsilon^p\|L^*LR_\varepsilon(v^\varepsilon(t)) - L^*LR_\varepsilon(v^\varepsilon(s))\|_{D(L)'}^p}{|t-s|^{1+\alpha p}}\mathrm{d}t\mathrm{d}s.$$

It follows from (15) and the Lipschitz continuity of R_ε (Lemma 10) that

$$\begin{aligned}\|L^*LR_\varepsilon(v^\varepsilon(t)) - L^*LR_\varepsilon(v^\varepsilon(s))\|_{D(L)'} &\le \|R_\varepsilon(v^\varepsilon(t)) - R_\varepsilon(v^\varepsilon(s))\|_{L^2(\mathcal{O})} \\ &\le \varepsilon^{-1}C\|v^\varepsilon(t) - v^\varepsilon(s)\|_{D(L)'}.\end{aligned}$$

Then we find that

$$\mathbb{E}|u^\varepsilon|^p_{W^{\alpha,p}(0,T;D(L)')} \le C\mathbb{E}\int_0^T\int_0^T \frac{\|v^\varepsilon(t) - v^\varepsilon(s)\|^p_{D(L)'}}{|t-s|^{1+\alpha p}}\mathrm{d}t\mathrm{d}s = C\mathbb{E}|v^\varepsilon|_{W^{\alpha,p}(0,T;D(L)')},$$

which finishes the proof. □

5.2. **Tightness of the laws of** (u^ε)**.** We show that the laws of (u^ε) are tight in a certain sub-Polish space. For this, we introduce the following spaces:

- $C^0([0,T];D(L)')$ is the space of continuous functions $u:[0,T]\to D(L)'$ with the topology $\mathbb{T}_1$ induced by the norm $\|u\|_{C^0([0,T];D(L)')} = \sup_{0<t<T}\|u(t)\|_{D(L)'}$;
- $L^2_w(0,T;H^1(\mathcal{O}))$ is the space $L^2(0,T;H^1(\mathcal{O}))$ with the weak topology $\mathbb{T}_2$.

We define the space

$$\widetilde{Z}_T := C^0([0,T];D(L)') \cap L^2_w(0,T;H^1(\mathcal{O})),$$

endowed with the topology $\widetilde{\mathbb{T}}$ that is the maximum of the topologies $\mathbb{T}_1$ and $\mathbb{T}_2$. The space $\widetilde{Z}_T$ is a sub-Polish space, since $C^0([0,T];D(L)')$ is separable and metrizable and

$$f_m(u) = \int_0^T (u(t), v_m(t))_{H^1(\mathcal{O})}\mathrm{d}t, \quad u \in L^2_w(0,T;H^1(\mathcal{O})),\ m\in\mathbb{N},$$

where $(v_m)_m$ is a dense subset of $L^2(0,T;H^1(\mathcal{O}))$, is a countable family (f_m) of point-separating functionals acting on $L^2(0,T;H^1(\mathcal{O}))$. In the following, we choose a number $s^*\ge 1$ such that

$$s^* < \frac{2d}{d-2} \quad \text{if } d\ge 3, \quad s^*<\infty \quad \text{if } d=2, \quad s^*\le\infty \quad \text{if } d=1. \tag{40}$$

Then the embedding $H^1(\mathcal{O}) \hookrightarrow L^{s^*}(\mathcal{O})$ is compact.

Lemma 22. *The set of laws of* (u^ε) *is tight in*

$$Z_T = \widetilde{Z}_T \cap L^2(0,T;L^{s^*}(\mathcal{O}))$$

with the topology $\mathbb{T}$ *that is the maximum of* $\widetilde{\mathbb{T}}$ *and the topology induced by the* $L^2(0,T;$ $L^{s^*}(\mathcal{O}))$ *norm, where* s^* *is given by* (40).

Proof. We apply Chebyshev's inequality for the first moment and use estimate (39) with $\beta = \alpha - 1/p > 0$, for any $\eta>0$ and $\delta>0$,

$$\sup_{\varepsilon>0}\mathbb{P}\bigg(\sup_{\substack{s,t\in[0,T],\\ |t-s|\le\delta}}\|u^\varepsilon(t)-u^\varepsilon(s)\|_{D(L)'}>\eta\bigg) \le \sup_{\varepsilon>0}\frac{1}{\eta}\mathbb{E}\bigg(\sup_{\substack{s,t\in[0,T],\\ |t-s|\le\delta}}\|u^\varepsilon(t)-u^\varepsilon(s)\|_{D(L)'}\bigg)$$

$$\leq \frac{\delta^\beta}{\eta}\sup_{\varepsilon>0}\mathbb{E}\bigg(\sup_{\substack{s,t\in[0,T],\\|t-s|\leq\delta}}\frac{\|u^\varepsilon(t)-u^\varepsilon(s)\|_{D(L)'}}{|t-s|^\beta}\bigg) \leq \frac{\delta^\beta}{\eta}\sup_{\varepsilon>0}\mathbb{E}\|u^\varepsilon\|_{C^{0,\beta}([0,T];D(L)'))} \leq C\frac{\delta^\beta}{\eta}.$$

This means that for all $\theta>0$ and all $\eta>0$, there exists $\delta>0$ such that

$$\sup_{\varepsilon>0}\mathbb{P}\bigg(\sup_{s,t\in[0,T],\,|t-s|\leq\delta}\|u^\varepsilon(t)-u^\varepsilon(s)\|_{D(L)'}>\eta\bigg)\leq\theta,$$

which is equivalent to the Aldous condition [5, Section 2.2]. Applying [38, Lemma 5, Theorem 3] with the spaces $X=H^1(\mathcal{O})$ and $B=D(L)'$, we conclude that (u^ε) is precompact in $C^0([0,T];D(L)')$. Then, proceeding as in the proof of the basic criterion for tightness [34, Chapter II, Section 2.1], we see that the set of laws of (u^ε) is tight in $C^0([0,T];D(L)')$.

Next, by Chebyshev's inequality again and estimate (26), for all $K>0$,

$$\mathbb{P}\big(\|u^\varepsilon\|_{L^2(0,T;H^1(\mathcal{O}))}>K\big)\leq\frac{1}{K^2}\mathbb{E}\|u^\varepsilon\|^2_{L^2(0,T;H^1(\mathcal{O}))}\leq\frac{C}{K^2}.$$

This implies that for any $\delta>0$, there exists $K>0$ such that $\mathbb{P}(\|u^\varepsilon\|_{L^2(0,T;H^1(\mathcal{O}))}\leq K)\leq 1-\delta$. Since closed balls with respect to the norm of $L^2(0,T;H^1(\mathcal{O}))$ are weakly compact, we infer that the set of laws of (u^ε) is tight in $L^2_w(0,T;H^1(\mathcal{O}))$.

The tightness in $L^2(0,T;L^{s^*}(\mathcal{O}))$ follows from Lemma 36 in Appendix B with $p=q=2$ and $r=2+4/d$. □

Lemma 23. *The set of laws of $(\sqrt{\varepsilon}L^*LR_\varepsilon(v^\varepsilon))$ is tight in*

$$Y_T := L^2_w(0,T;D(L)')\cap L^\infty_{w*}(0,T;D(L)')$$

with the associated topology $\mathbb{T}_Y$.

Proof. We apply the Chebyshev inequality and use the inequality $\|L^*LR_\varepsilon(v^\varepsilon)\|_{D(L)'}\leq C\|LR_\varepsilon(v^\varepsilon)\|_{L^2(\mathcal{O})}$ and estimate (27):

$$\mathbb{P}\big(\sqrt{\varepsilon}\|L^*LR_\varepsilon(v^\varepsilon)\|_{L^2(0,T;D(L)')}>K\big)\leq\frac{\varepsilon}{K^2}\mathbb{E}\|L^*LR_\varepsilon(v^\varepsilon)\|^2_{L^2(0,T;D(L)')}\leq\frac{C}{K^2}$$

for any $K>0$. Since closed balls in $L^2(0,T;D(L)')$ are weakly compact, the set of laws of $(\sqrt{\varepsilon}L^*LR_\varepsilon(v^\varepsilon))$ is tight in $L^2_w(0,T;D(L)')$. The second claim follows from an analogous argument. □

5.3. **Convergence of** (u^ε)**.** Let $\mathrm{P}(X)$ be the space of probability measures on X. We consider the space $Z_T\times Y_T\times C^0([0,T];U_0)$, equipped with the probability measure $\mu^\varepsilon := \mu^\varepsilon_u\times\mu^\varepsilon_w\times\mu^\varepsilon_W$, where

$$\begin{aligned}\mu^\varepsilon_u(\cdot)&=\mathbb{P}(u^\varepsilon\in\cdot)\in\mathrm{P}(Z_T),\\ \mu^\varepsilon_w&=\mathbb{P}(\sqrt{\varepsilon}L^*LR_\varepsilon(v^\varepsilon)\in\cdot)\in\mathrm{P}(Y_T),\\ \mu^\varepsilon_W(\cdot)&=\mathbb{P}(W\in\cdot)\in\mathrm{P}(C^0([0,T];U_0)),\end{aligned}$$

recalling the choice (40) of s^*. The set of measures (μ^ε) is tight, since the set of laws of (u^ε) and $(\sqrt{\varepsilon}L^*LR_\varepsilon(v^\varepsilon))$ are tight in $(Z_T,\mathbb{T})$ and $(Y_T,\mathbb{T}_Y)$, respectively. Moreover, (μ^ε_W) consists of one element only and is consequently weakly compact in $C^0([0,T];U_0)$. By Prohorov's theorem, (μ_W^ε) is tight. Hence, $Z_T\times Y_T\times C^0([0,T];U_0)$ satisfies the assumptions of the Skorokhod–Jakubowski theorem [6, Theorem C.1]. We infer that there exists a subsequence of $(u^\varepsilon,\sqrt{\varepsilon}L^*LR_\varepsilon(v^\varepsilon))$, which is not relabeled, a probability space $(\widetilde\Omega,\widetilde{\mathcal F},\widetilde{\mathbb P})$ and, on this space, $(Z_T\times Y_T\times C^0([0,T];U_0))$-valued random variables $(\widetilde u,\widetilde w,\widetilde W)$ and $(\widetilde u^\varepsilon,\widetilde w^\varepsilon,\widetilde W^\varepsilon)$ such that $(\widetilde u^\varepsilon,\widetilde w^\varepsilon,\widetilde W^\varepsilon)$ has the same law as $(u^\varepsilon,\sqrt{\varepsilon}L^*LR_\varepsilon(v^\varepsilon),W)$ on $\mathcal B(Z_T\times Y_T\times C^0([0,T];U_0))$ and, as $\varepsilon\to 0$,

$$(\widetilde u^\varepsilon,\widetilde w^\varepsilon,\widetilde W^\varepsilon)\to(\widetilde u,\widetilde w,\widetilde W)\quad\text{in } Z_T\times Y_T\times C^0([0,T];U_0)\quad \widetilde{\mathbb P}\text{-a.s.}$$

By the definition of Z_T and Y_T, this convergence means $\widetilde{\mathbb P}$-a.s.,

$$\begin{aligned}
\widetilde u^\varepsilon &\to \widetilde u && \text{strongly in } C^0([0,T];D(L)'),\\
\widetilde u^\varepsilon &\rightharpoonup \widetilde u && \text{weakly in } L^2(0,T;H^1(\mathcal O)),\\
\widetilde u^\varepsilon &\to \widetilde u && \text{strongly in } L^2(0,T;L^{s^*}(\mathcal O)),\\
\widetilde w^\varepsilon &\rightharpoonup \widetilde w && \text{weakly in } L^2(0,T;D(L)'),\\
\widetilde w^\varepsilon &\rightharpoonup \widetilde w && \text{weakly* in } L^\infty(0,T;D(L)'),\\
\widetilde W^\varepsilon &\to \widetilde W && \text{strongly in } C^0([0,T];U_0).
\end{aligned}$$

We derive some regularity properties for the limit $\widetilde u$. We note that $\widetilde u$ is a Z_T-Borel random variable, since $\mathcal B(Z_T\times Y_T\times C^0([0,T];U_0))$ is a subset of $\mathcal B(Z_T)\times\mathcal B(Y_T)\times\mathcal B(C^0([0,T];U_0))$. We deduce from estimates (25) and (26) and the fact that u^ε and $\widetilde u^\varepsilon$ have the same law that

$$\sup_{\varepsilon>0}\widetilde{\mathbb E}\|\widetilde u^\varepsilon\|^p_{L^2(0,T;H^1(\mathcal O))}+\sup_{\varepsilon>0}\widetilde{\mathbb E}\|\widetilde u^\varepsilon\|^p_{L^\infty(0,T;D(L)')}<\infty.$$

We infer the existence of a further subsequence of $(\widetilde u^\varepsilon)$ (not relabeled) that is weakly converging in $L^p(\widetilde\Omega;L^2(0,T;H^1(\mathcal O)))$ and weakly* converging in $L^p(\widetilde\Omega;C^0([0,T];D(L)'))$ as $\varepsilon\to 0$. Because $\widetilde u^\varepsilon\to\widetilde u$ in Z_T $\widetilde{\mathbb P}$-a.s., we conclude that the limit function satisfies

$$\widetilde{\mathbb E}\|\widetilde u\|^p_{L^2(0,T;H^1(\mathcal O))}+\widetilde{\mathbb E}\|\widetilde u\|^p_{L^\infty(0,T;D(L)')}<\infty.$$

Let $\widetilde{\mathbb F}$ and $\widetilde{\mathbb F}^\varepsilon$ be the filtrations generated by $(\widetilde u,\widetilde w,\widetilde W)$ and $(\widetilde u^\varepsilon,\widetilde w^\varepsilon,\widetilde W)$, respectively. By following the arguments of the proof of [7, Proposition B4], we can verify that these new random variables induce actually stochastic processes. The progressive measurability of $\widetilde u^\varepsilon$ is a consequence of [4, Appendix B]. Set $\widetilde W_j^{\varepsilon,k}(t):=(\widetilde W^\varepsilon(t),e_k)_U$. We claim that $\widetilde W_j^{\varepsilon,k}(t)$ for $k\in\mathbb N$ are independent, standard $\widetilde{\mathcal F}_t$-Wiener processes. The adaptedness is a direct consequence of the definition; the independence of $\widetilde W_j^{\varepsilon,k}(t)$ and the independence of the increments $\widetilde W^{\varepsilon,k}(t)-\widetilde W^{\varepsilon,k}(s)$ with respect to $\widetilde{\mathcal F}_s$ are inherited from $(W(t),e_k)_U$. Passing to the limit $\varepsilon\to 0$ in the characteristic function, by using dominated convergence, we find that $\widetilde W(t)$ are $\widetilde{\mathcal F}_t$-martingales with the correct marginal distributions. We deduce from Lévy's characterization theorem that $\widetilde W(t)$ is indeed a cylindrical Wiener process.

By definition, $u_i^\varepsilon=u_i(R_\varepsilon(v^\varepsilon))=\exp(R_\varepsilon(v^\varepsilon))$ is positive in Q_T a.s. We claim that also $\widetilde u_i$ is nonnegative in $\mathcal O$ a.s.

Lemma 24 (Nonnegativity)**.** *It holds that* $\widetilde{u}_i \ge 0$ *a.e. in* Q_T $\widetilde{\mathbb{P}}$*-a.s. for all* $i = 1, \dots, n$.

Proof. Let $i \in \{1, \dots, n\}$. Since $u_i^\varepsilon > 0$ in Q_T a.s., we have $\mathbb{E}\|(u_i^\varepsilon)^-\|_{L^2(0,T;L^2(\mathcal{O}))} = 0$, where $z^- = \min\{0, z\}$. The function u_i^ε is Z_T-Borel measurable and so does its negative part. Therefore, using the equivalence of the laws of u_i^ε and $\widetilde{u}_i^\varepsilon$ in Z_T and writing μ_i^ε and $\widetilde{\mu}_i^\varepsilon$ for the laws of u_i^ε and $\widetilde{u}_i^\varepsilon$, respectively, we obtain

$$\begin{aligned}\widetilde{\mathbb{E}}\|(\widetilde{u}_i^\varepsilon)^-\|_{L^2(Q_T)} &= \int_{L^2(Q_T)} \|y^-\|_{L^2(Q_T)} \mathrm{d}\widetilde{\mu}_i^\varepsilon(y) \\ &= \int_{L^2(Q_T)} \|y^-\|_{L^2 Q_T)} \mathrm{d}\mu_i^\varepsilon(y) = \mathbb{E}\|u_i^\varepsilon\|_{L^2(Q_T)} = 0.\end{aligned}$$

This shows that $\widetilde{u}_i^\varepsilon \ge 0$ a.e. in Q_T $\widetilde{\mathbb{P}}$-a.s. The convergence (up to a subsequence) $\widetilde{u}^\varepsilon \to \widetilde{u}$ a.e. in Q_T $\widetilde{\mathbb{P}}$-a.s. then implies that $\widetilde{u}_i \ge 0$ in Q_T $\widetilde{\mathbb{P}}$-a.s. ☐

The following lemma is needed to verify that $(\widetilde{u}, \widetilde{W})$ is a martingale solution to (1)–(2).

Lemma 25. *It holds for all* $t \in [0, T]$, $i = 1, \dots, n$, *and all* $\phi_1 \in L^2(\mathcal{O})$ *and all* $\phi_2 \in D(L)$ *that*

$$\lim_{\varepsilon \to 0} \widetilde{\mathbb{E}} \int_0^T \big(\widetilde{u}_i^\varepsilon(t) - \widetilde{u}_i(t), \phi_1\big)_{L^2(\mathcal{O})} \mathrm{d}t = 0, \tag{41}$$

$$\lim_{\varepsilon \to 0} \widetilde{\mathbb{E}} \big\langle \widetilde{u}_i^\varepsilon(0) - \widetilde{u}_i(0), \phi_2\big\rangle_{D(L)', D(L)} = 0, \tag{42}$$

$$\lim_{\varepsilon \to 0} \widetilde{\mathbb{E}} \int_0^T \big\langle \sqrt{\varepsilon}\widetilde{w}_i^\varepsilon(t), \phi_2\big\rangle_{D(L)', D(L)} \mathrm{d}t = 0, \tag{43}$$

$$\lim_{\varepsilon \to 0} \widetilde{\mathbb{E}} \langle \sqrt{\varepsilon}\widetilde{w}_i^\varepsilon(0), \phi_2\rangle_{D(L)', D(L)} = 0, \tag{44}$$

$$\lim_{\varepsilon \to 0} \widetilde{\mathbb{E}} \int_0^T \bigg| \sum_{j=1}^n \int_0^t \int_{\mathcal{O}} \big(A_{ij}(\widetilde{u}^\varepsilon(s))\nabla\widetilde{u}_j^\varepsilon(s) - A_{ij}(\widetilde{u}(s))\nabla\widetilde{u}_j(s)\big) \cdot \nabla\phi_2 \mathrm{d}x\mathrm{d}s\bigg| \mathrm{d}t = 0, \tag{45}$$

$$\lim_{\varepsilon \to 0} \widetilde{\mathbb{E}} \int_0^T \bigg| \sum_{j=1}^n \int_0^t \Big(\sigma_{ij}(\widetilde{u}^\varepsilon(s))\mathrm{d}\widetilde{W}_j^\varepsilon(s) - \sigma_{ij}(\widetilde{u}(s))\mathrm{d}\widetilde{W}_j(s), \phi_1\Big)_{L^2(\mathcal{O})}\bigg|^2 \mathrm{d}t = 0. \tag{46}$$

Proof. The proof is a combination of the uniform bounds and Vitali's convergence theorem. Convergences (41) and (42) have been shown in the proof of [18, Lemma 16], and (43) is a direct consequence of (38) and

$$\widetilde{\mathbb{E}}\bigg(\int_0^T \langle\sqrt{\varepsilon}\widetilde{w}_i^\varepsilon(t), \phi_2\rangle_{D(L)', D(L)} \mathrm{d}t\bigg)^p \le \varepsilon^{p/2}\widetilde{\mathbb{E}}\bigg(\int_0^T \|\widetilde{w}_i^\varepsilon(t)\|_{D(L)'}\|\phi_2\|_{D(L)}\mathrm{d}t\bigg)^p \le \varepsilon^{p/2}C.$$

Convergence (44) follows from $\widetilde{w}_i^\varepsilon \rightharpoonup \widetilde{w}_i$ weakly* in $L^\infty(0, T; D(L)')$. We establish (45):

$$\bigg|\int_0^T \bigg| \sum_{j=1}^n \int_0^t \int_{\mathcal{O}} \big(A_{ij}(\widetilde{u}^\varepsilon(s))\nabla\widetilde{u}_j^\varepsilon(s) - A_{ij}(\widetilde{u}(s))\nabla\widetilde{u}_j(s)\big) \cdot \nabla\phi_2 \mathrm{d}x\mathrm{d}s\bigg|$$

$$\leq \int_0^T \|A_{ij}(\widetilde{u}^\varepsilon(s)) - A_{ij}(\widetilde{u}(s))\|_{L^2(\mathcal{O})} \|\nabla \widetilde{u}_j^\varepsilon(s)\|_{L^2(\mathcal{O})} \|\nabla\phi_2\|_{L^\infty(\mathcal{O})} \mathrm{d}s$$
$$+ \left| \int_0^T \int_{\mathcal{O}} A_{ij}(\widetilde{u}(s)) \nabla(\widetilde{u}^\varepsilon(s) - \widetilde{u}(s)) \cdot \nabla\phi_2 \mathrm{d}x\mathrm{d}s \right| =: I_1^\varepsilon + I_2^\varepsilon.$$

By the Lipschitz continuity of A and the uniform bound for $\nabla\widetilde{u}^\varepsilon$, we have $I_1^\varepsilon \to 0$ as $\varepsilon \to 0$ $\widetilde{\mathbb{P}}$-a.s. At this point, we use the embedding $D(L) \hookrightarrow W^{1,\infty}(\mathcal{O})$. Also the second integral I_2^ε converges to zero, since $A_{ij}(\widetilde{u})\nabla\phi_2 \in L^2(0,T;L^2(\mathcal{O}))$ and $\nabla\widetilde{u}_j^\varepsilon \rightharpoonup \nabla\widetilde{u}_j$ weakly in $L^2(0,T;L^2(\mathcal{O}))$. This shows that $\widetilde{\mathbb{P}}$-a.s.,

$$\lim_{\varepsilon\to 0} \int_0^T \int_{\mathcal{O}} A_{ij}(\widetilde{u}^\varepsilon(s)) \nabla\widetilde{u}_j^\varepsilon(s) \cdot \nabla\phi_2 \mathrm{d}x\mathrm{d}s = \int_0^T \int_{\mathcal{O}} A_{ij}(\widetilde{u}(s)) \nabla\widetilde{u}_j(s) \cdot \nabla\phi_2 \mathrm{d}x\mathrm{d}s.$$

A straightforward estimation and bound (31) lead to

$$\widetilde{\mathbb{E}} \left| \int_0^T \int_{\mathcal{O}} A_{ij}(\widetilde{u}^\varepsilon(s)) \nabla\widetilde{u}_j^\varepsilon(s) \cdot \nabla\phi_2 \mathrm{d}x\mathrm{d}s \right|^p$$
$$\leq \|\nabla\phi_2\|_{L^\infty(\mathcal{O})}^p \widetilde{\mathbb{E}} \bigg(\int_0^T \bigg\| \sum_{j=1}^n A_{ij}(\widetilde{u}^\varepsilon(s)) \nabla\widetilde{u}_j^\varepsilon(s) \bigg\|_{L^1(\mathcal{O})} \mathrm{d}s \bigg)^p \leq C,$$

Hence, Vitali's convergence theorem gives (45).

It remains to prove (46). By Assumption (A4), $\widetilde{\mathbb{P}}$-a.s.,

$$\int_0^T \|\sigma_{ij}(\widetilde{u}^\varepsilon(s)) - \sigma_{ij}(\widetilde{u}(s))\|_{\mathcal{L}_2(U;L^2(\mathcal{O}))}^2 \mathrm{d}s \leq C_\sigma \|\widetilde{u}^\varepsilon - \widetilde{u}\|_{L^2(0,T;L^2(\mathcal{O}))} \to 0.$$

This convergence and $\widetilde{W}^\varepsilon \to \widetilde{W}$ in $C^0([0,T];U_0)$ imply that [14, Lemma 2.1]

$$\int_0^T \sigma_{ij}(\widetilde{u}^\varepsilon) \mathrm{d}\widetilde{W}^\varepsilon \to \int_0^T \sigma_{ij}(\widetilde{u}) \mathrm{d}\widetilde{W} \quad \text{in } L^2(0,T;L^2(\mathcal{O}))\ \widetilde{\mathbb{P}}\text{-a.s.}$$

By Assumption (A4) again,

$$\widetilde{\mathbb{E}} \bigg(\int_0^T \|\sigma_{ij}(\widetilde{u}^\varepsilon(s)) - \sigma_{ij}(\widetilde{u}(s))\|_{\mathcal{L}_2(U;L^2(\mathcal{O}))}^2 \mathrm{d}s \bigg)^p$$
$$\leq C + C\widetilde{\mathbb{E}} \bigg(\int_0^T \big(\|\widetilde{u}^\varepsilon(s)\|_{L^2(\mathcal{O})}^2 + \|\widetilde{u}(s)\|_{L^2(\mathcal{O})}^2 \big) \mathrm{d}s \bigg)^p \leq C.$$

We infer from Vitali's convergence theorem that

$$\lim_{\varepsilon\to 0} \widetilde{\mathbb{E}} \int_0^T \|\sigma_{ij}(\widetilde{u}^\varepsilon(s)) - \sigma_{ij}(\widetilde{u}(s))\|_{\mathcal{L}_2(U;L^2(\mathcal{O}))}^2 \mathrm{d}s = 0.$$

The estimate

$$\widetilde{\mathbb{E}} \bigg| \bigg(\int_0^T \sigma_{ij}(\widetilde{u}^\varepsilon(s)) \mathrm{d}\widetilde{W}_j^\varepsilon(s) - \int_0^T \sigma_{ij}(\widetilde{u}(s)) \mathrm{d}\widetilde{W}_j(s), \phi_1 \bigg)_{L^2(\mathcal{O})} \bigg|^2$$

$$
\begin{aligned}
&\le C\|\phi_1\|_{L^2(\mathcal{O})}^2 \widetilde{\mathbb{E}}\int_0^T \big(\|\sigma_{ij}(\widetilde{u}^\varepsilon(s))\|^2_{\mathcal{L}_2(U;L^2(\mathcal{O}))} + \|\sigma_{ij}(\widetilde{u}(s))\|^2_{\mathcal{L}_2(U;L^2(\mathcal{O}))}\big)\mathrm{d}s \\
&\le C\|\phi_1\|_{L^2(\mathcal{O})}^2 \bigg\{1 + \widetilde{\mathbb{E}}\bigg(\int_0^T \big(\|\widetilde{u}^\varepsilon(s)\|^2_{L^2(\mathcal{O})} + \|\widetilde{u}(s)\|^2_{L^2(\mathcal{O})}\big)\mathrm{d}s\bigg)\bigg\} \le C
\end{aligned}
$$

for all $\phi_1 \in L^2(\mathcal{O})$ and the dominated convergence theorem yield (46). ☐

To show that the limit is indeed a solution, we define, for $t \in [0,T]$, $i = 1,\dots,n$, and $\phi \in D(L)$,

$$
\begin{aligned}
\Lambda_i^\varepsilon(\widetilde{u}^\varepsilon, \widetilde{w}^\varepsilon, \widetilde{W}^\varepsilon, \phi)(t) := \langle \widetilde{u}_i(0), \phi\rangle &+ \sqrt{\varepsilon}\langle \widetilde{w}^\varepsilon(0), \phi\rangle \\
&- \sum_{j=1}^n \int_0^t \int_{\mathcal{O}} A_{ij}(\widetilde{u}^\varepsilon(s))\nabla \widetilde{u}_j^\varepsilon(s)\cdot\nabla\phi \mathrm{d}x\mathrm{d}s \\
&+ \sum_{j=1}^n \bigg(\int_0^t \sigma_{ij}(\widetilde{u}^\varepsilon(s))\mathrm{d}\widetilde{W}_j^\varepsilon(s), \phi\bigg)_{L^2(\mathcal{O})}, \\
\Lambda_i(\widetilde{u}, \widetilde{w}, \widetilde{W}, \phi)(t) := \langle \widetilde{u}_i(0), \phi\rangle &- \sum_{j=1}^n \int_0^t \int_{\mathcal{O}} \langle A_{ij}(\widetilde{u}(s))\nabla \widetilde{u}_j(s)\cdot\nabla\phi \mathrm{d}x\mathrm{d}s \\
&+ \sum_{j=1}^n \bigg(\int_0^t \sigma_{ij}(\widetilde{u}(s))\mathrm{d}\widetilde{W}_j(s), \phi\bigg)_{L^2(\mathcal{O})}.
\end{aligned}
$$

The following corollary is a consequence of the previous lemma.

Corollary 26. *It holds for any $\phi_1 \in L^2(\mathcal{O})$ and $\phi_2 \in D(L)$ that*

$$
\begin{aligned}
\lim_{\varepsilon\to 0}\big\|(\widetilde{u}_i^\varepsilon, \phi_1)_{L^2(\mathcal{O})} - (\widetilde{u}_i, \phi_1)_{L^2(\mathcal{O})}\big\|_{L^1(\widetilde{\Omega}\times(0,T))} &= 0, \\
\lim_{\varepsilon\to 0}\|\Lambda_i^\varepsilon(\widetilde{u}^\varepsilon, \sqrt{\varepsilon}\widetilde{w}^\varepsilon, \widetilde{W}^\varepsilon, \phi_2) - \Lambda_i(\widetilde{u}, 0, \widetilde{W}, \phi_2)\|_{L^1(\widetilde{\Omega}\times(0,T))} &= 0.
\end{aligned}
$$

Since v^ε is a strong solution to (17), it satisfies for a.e. $t \in [0,T]$ $\mathbb{P}$-a.s., $i = 1,\dots,n$, and $\phi \in D(L)$,

$$
(v_i^\varepsilon(t), \phi)_{L^2(\mathcal{O})} = \Lambda_i^\varepsilon(u^\varepsilon, \varepsilon L^*LR_\varepsilon(v^\varepsilon), W, \phi)(t)
$$

and in particular,

$$
\int_0^T \mathbb{E}\big|(v_i^\varepsilon(t), \phi)_{L^2(\mathcal{O})} - \Lambda_i^\varepsilon(u^\varepsilon, \varepsilon L^*LR_\varepsilon(v^\varepsilon), W, \phi)(t)\big|\mathrm{d}t = 0.
$$

We deduce from the equivalence of the laws of $(u^\varepsilon, \varepsilon L^*LR_\varepsilon(v^\varepsilon), W)$ and $(\widetilde{u}^\varepsilon, \sqrt{\varepsilon}\widetilde{w}^\varepsilon, \widetilde{W})$ that

$$
\int_0^T \widetilde{\mathbb{E}}\big|\big(\widetilde{u}_i^\varepsilon(t) + \sqrt{\varepsilon}\widetilde{w}_i^\varepsilon, \phi\big)_{L^2(\mathcal{O})} - \Lambda_i^\varepsilon\big(\widetilde{u}^\varepsilon, \sqrt{\varepsilon}\widetilde{w}^\varepsilon, \widetilde{W}^\varepsilon, \phi\big)(t)\big|\mathrm{d}t = 0.
$$

By Corollary 26, we can pass to the limit $\varepsilon \to 0$ to obtain

$$
\int_0^T \widetilde{\mathbb{E}}\big|(\widetilde{u}_i(t), \phi)_{L^2(\mathcal{O})} - \Lambda_i(\widetilde{u}, 0, \widetilde{W}, \phi)(t)\big|\mathrm{d}t = 0.
$$

This identity holds for all $i=1,\ldots,n$ and all $\phi\in D(L)$. This shows that

$$\big|(\widetilde{u}_i(t),\phi)_{L^2(\mathcal{O})}-\Lambda_i(\widetilde{u},0,\widetilde{W},\phi)(t)\big|=0\quad\text{for a.e. } t\in[0,T]\ \widetilde{\mathbb{P}}\text{-a.s.},\ i=1,\ldots,n.$$

We infer from the definition of Λ_i that

$$\begin{aligned}(\widetilde{u}_i(t),\phi)_{L^2(\mathcal{O})}&=(\widetilde{u}_i(0),\phi)_{L^2(\mathcal{O})}-\sum_{j=1}^n\int_0^t\int_{\mathcal{O}}A_{ij}(\widetilde{u}(s))\nabla\widetilde{u}_j(s)\cdot\nabla\phi\mathrm{d}x\mathrm{d}s\\&\quad+\sum_{j=1}^n\bigg(\int_0^t\sigma_{ij}(\widetilde{u}(s))\mathrm{d}\widetilde{W}_j(s),\phi\bigg)_{L^2(\mathcal{O})}\end{aligned}$$

for a.e. $t\in[0,T]$ and all $\phi\in D(L)$. Set $\widetilde{U}=(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}},\widetilde{\mathbb{F}})$. Then $(\widetilde{U},\widetilde{W},\widetilde{u})$ is a martingale solution to (1)–(3).

6. Proof of Theorem 4

We turn to the existence proof of the SKT model without self-diffusion.

6.1. **Uniform estimates.** Let v^ε be a global solution to (19)–(20) and set $u^\varepsilon=u(R_\varepsilon(v^\varepsilon))$. We assume that $A(u)$ is given by (3) and that $a_{i0}>0$, $a_{ii}=0$ for $i=1,\ldots,n$. The uniform estimates of Lemmas 17 and 18 are still valid. Since $a_{ii}=0$, we obtain an $H^1(\mathcal{O})$ bound for $(u_i^\varepsilon)^{1/2}$ instead of u_i^ε, which yields weaker bounds than those in Lemma 19.

Lemma 27. *Let $p\geq 2$ and set $\rho_1:=(d+2)/(d+1)$. Then there exists a constant $C(p,u^0,T)>0$, which is independent of ε, such that*

$$\mathbb{E}\|u_i^\varepsilon\|^p_{L^2(0,T;W^{1,1}(\mathcal{O}))}\leq C(p,u^0,T),\tag{47}$$

$$\mathbb{E}\|u_i^\varepsilon\|^p_{L^{1+2/d}(Q_T)}\leq C(p,u^0,T),\tag{48}$$

$$\mathbb{E}\|u_i^\varepsilon\|^p_{L^{4/d}(0,T;L^2(\mathcal{O}))}\leq C(p,u^0,T),\tag{49}$$

$$\mathbb{E}\|u_i^\varepsilon\|^p_{L^{\rho_1}(0,T;W^{1,\rho_1}(\mathcal{O}))}\leq C(p,u^0,T).\tag{50}$$

Proof. The identity $\nabla u_i^\varepsilon=2(u_i^\varepsilon)^{1/2}\nabla(u_i^\varepsilon)^{1/2}$ and the Hölder inequality show that

$$\begin{aligned}\mathbb{E}\|\nabla u_i^\varepsilon\|^p_{L^2(0,T;L^1(\mathcal{O}))}&\leq C\mathbb{E}\bigg(\int_0^T\|(u_i^\varepsilon)^{1/2}\|^2_{L^2(\mathcal{O})}\|\nabla(u_i^\varepsilon)^{1/2}\|^2_{L^2(\mathcal{O})}\mathrm{d}t\bigg)^{p/2}\\&\leq C\mathbb{E}\bigg(\|u_i^\varepsilon\|_{L^\infty(0,T;L^1(\mathcal{O}))}\int_0^T\|\nabla(u_i^\varepsilon)^{1/2}\|^2_{L^2(\mathcal{O})}\mathrm{d}t\bigg)^{p/2}\\&\leq C\big(\mathbb{E}\|u_i^\varepsilon\|^p_{L^\infty(0,T;L^1(\mathcal{O}))}\big)^{1/2}\big(\mathbb{E}\|\nabla(u_i^\varepsilon)^{1/2}\|^{2p}_{L^2(0,T;L^2(\mathcal{O}))}\big)^{1/2}.\end{aligned}$$

Because of (30) and (31), the right-hand side is bounded. Using (30) again, we infer that (47) holds. Estimate (48) is obtained from the Gagliardo–Nirenberg inequality similarly as in the proof of Lemma 19:

$$\mathbb{E}\bigg(\int_0^T\|(u_i^\varepsilon)^{1/2}\|^s_{L^r(\mathcal{O})}\bigg)^{p/s}\leq C\big(\mathbb{E}\|(u_i^\varepsilon)^{1/2}\|^{2(1-\theta)p}_{L^\infty(0,T;L^2(\mathcal{O}))}\big)^{1/2}$$

$$\times \big(\mathbb{E}\|(u_i^\varepsilon)^{1/2}\|_{L^2(0,T;H^1(\mathcal{O}))}^{4p/s}\big)^{1/2} \le C,$$

where $s = 2/\theta \ge 2$ and $1/r = 1/2 - \theta/d = 1/2 - 2/(ds)$. Choosing $r = (2d+4)/d$ gives $s = r$, and $r = 4$ leads to $s = 8/d$; this proves estimates (48) and (49). Finally, (50) follows from Hölder's inequality:

$$\begin{aligned}\|u_i^\varepsilon\|_{L^{\rho_1}(Q_T)} &= 2\|(u_i^\varepsilon)^{1/2}\nabla(u_i^\varepsilon)^{1/2}\|_{L^{\rho_1}(Q_T)} \le 2\|(u_i^\varepsilon)^{1/2}\|_{L^{(2d+4)/d}(Q_T)}\|\nabla(u_i^\varepsilon)^{1/2}\|_{L^2(Q_T)} \\ &\le 2\|u_i^\varepsilon\|_{L^{1+2/d}(Q_T)}^{1/2}\|(u_i^\varepsilon)^{1/2}\|_{L^2(0,T;H^1(\mathcal{O}))}\end{aligned}$$

and taking the expectation and using (48) and (31) ends the proof. □

The following lemma is needed to derive the fractional time estimate.

Lemma 28. *Let $p \ge 2$ and set $\rho_2 := (2d+2)/(2d+1)$. Then it holds for any $i, j = 1, \ldots, n$ with $i \ne j$:*

$$\mathbb{E}\|u_i^\varepsilon u_j^\varepsilon\|_{L^{\rho_2}(0,T;W^{1,\rho_2}(\mathcal{O}))}^p \le C(p, u^0, T). \tag{51}$$

Proof. The Hölder inequality and (30) immediately yield

$$\mathbb{E}\|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^\infty(0,T;L^1(\mathcal{O}))}^p \le C,$$

and we conclude from the Poincaré–Wirtinger inequality, estimate (32), and the previous estimate that

$$\mathbb{E}\|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^2(0,T;H^1(\mathcal{O}))}^p \le C. \tag{52}$$

By the Gagliardo–Nirenberg inequality, with $\theta = d/(d+1)$,

$$\begin{aligned}\int_0^T \|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^{2(d+1)/d}(\mathcal{O})}^{2(d+1)/d}\mathrm{d}t &\le C\int_0^T \|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{H^1(\mathcal{O})}^{2\theta(d+1)/d}\|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^1(\mathcal{O})}^{2(1-\theta)(d+1)/d}\mathrm{d}t \\ &\le C\|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^\infty(0,T;L^1(\mathcal{O}))}^{2(1-\theta)(d+1)/d}\int_0^T \|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{H^1(\mathcal{O})}^2\mathrm{d}t \\ &= C\|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^\infty(0,T;L^1(\mathcal{O}))}^{2/d}\|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^2(0,T;H^1(\mathcal{O}))}^2.\end{aligned}$$

Taking the expectation and applying the Hölder inequality, we infer that

$$\mathbb{E}\|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^{2(d+1)/d}(Q_T)}^p \le C. \tag{53}$$

Finally, the identity $\nabla(u_i^\varepsilon u_j^\varepsilon) = 2(u_i^\varepsilon u_j^\varepsilon)^{1/2}\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}$ and Hölder's inequality lead to

$$\begin{aligned}\int_0^T \|\nabla(u_i^\varepsilon u_j^\varepsilon)\|_{L^{\rho_2}(\mathcal{O})}^{\rho_2}\mathrm{d}t &\le C\int_0^T \|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^{2(d+1)/d}(\mathcal{O})}^2\|\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^2(\mathcal{O})}^2\mathrm{d}t \\ &\le C\bigg(\int_0^T \|(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^{2(d+1)/d}(\mathcal{O})}^{2(d+1)/d}\mathrm{d}t\bigg)^{1-\rho_2/2}\bigg(\int_0^T \|\nabla(u_i^\varepsilon u_j^\varepsilon)^{1/2}\|_{L^2(\mathcal{O})}^2\mathrm{d}t\bigg)^{\rho_2/2}.\end{aligned}$$

The bounds (52)–(53) yield, after taking the expectation and applying Hölder's inequality again, the conclusion (51). □

We show now that the fractional time derivative of u^ε is uniformly bounded.

Lemma 29 (Fractional time regularity)**.** *Let $d \le 2$. Then there exist $0 < \alpha < 1$, $p > 1$, and $\beta > 0$ such that $\alpha p > 1$ and*

$$\mathbb{E}\|u^\varepsilon\|^p_{W^{\alpha,p}(0,T;D(L)')} + \mathbb{E}\|u^\varepsilon\|^p_{C^{0,\beta}([0,T];D(L)')} \le C. \tag{54}$$

Proof. We proceed similarly as in the proof of Lemma 21. First, we estimate the diffusion part, setting

$$g(t) = \int_0^t \bigg\| a_{i0}\nabla u_i^\varepsilon + \sum_{j\neq i} a_{ij}\nabla(u_i^\varepsilon u_j^\varepsilon)\bigg\|_{L^1(\mathcal{O})} \mathrm{d}r.$$

Then, using $D(L) \subset W^{1,\infty}(\mathcal{O})$ (which holds due to the assumption $m > d/2+1$),

$$\begin{aligned}
\mathbb{E}&\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg\|\int_{s\wedge t}^{t\vee s} \operatorname{div}\sum_{j=1}^n A_{ij}(u^\varepsilon(r))\nabla u_j^\varepsilon(r)\mathrm{d}r\bigg\|^p_{D(L)'}\mathrm{d}t\mathrm{d}s \\
&\le C\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg(\int_{s\wedge t}^{t\vee s}\bigg\| a_{i0}\nabla u_i^\varepsilon + \sum_{j\neq i} a_{ij}\nabla(u_i^\varepsilon u_j^\varepsilon)\bigg\|_{L^1(\mathcal{O})}\mathrm{d}r\bigg)^p\mathrm{d}t\mathrm{d}s \\
&\le C\mathbb{E}\int_0^T\int_0^T \frac{|g(t)-g(s)|^p}{|t-s|^{1+\alpha p}}\mathrm{d}t\mathrm{d}s \le C\mathbb{E}\|g\|^p_{W^{\alpha,p}(0,T;\mathbb{R})}.
\end{aligned}$$

The embedding $W^{1,p}(0,T;\mathbb{R}) \hookrightarrow W^{\alpha,p}(0,T;\mathbb{R})$ and estimates (51) and (50) show that for $1 \le p \le \rho_1 = (d+2)/(d+1)$,

$$\begin{aligned}
\mathbb{E}\|g\|^p_{W^{\alpha,p}(0,T;\mathbb{R})} &\le C\mathbb{E}\|g\|^p_{W^{1,p}(0,T;\mathbb{R})} = C\mathbb{E}\|\partial_t g\|^p_{L^p(0,T;\mathbb{R})} + C\mathbb{E}\|g\|^p_{L^p(0,T;\mathbb{R})} \\
&\le C\mathbb{E}\int_0^T \bigg\| a_{i0}\nabla u_i^\varepsilon(t) + \sum_{j\neq i} a_{ij}\nabla(u_i^\varepsilon u_j^\varepsilon)(t)\bigg\|^p_{L^1(\mathcal{O})}\mathrm{d}t \\
&\quad + C\mathbb{E}\int_0^T\int_0^t \bigg\| a_{i0}\nabla u_i^\varepsilon(r) + \sum_{j\neq i} a_{ij}\nabla(u_i^\varepsilon u_j^\varepsilon)(r)\bigg\|^p_{L^1(\mathcal{O})}\mathrm{d}r\mathrm{d}t \le C.
\end{aligned}$$

Next, we consider the stochastic part, using the Burkholder–Davis–Gundy inequality, Hölder's inequality, and the sublinear growth condition in the statement of the theorem:

$$\begin{aligned}
\mathbb{E}&\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg\|\int_{s\wedge t}^{t\vee s}\sum_{j=1}^n \sigma_{ij}(u^\varepsilon(r))\mathrm{d}W_j(r)\bigg\|^p_{L^2(\mathcal{O})}\mathrm{d}t\mathrm{d}s \\
&\le C\mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\bigg(\int_{s\wedge t}^{t\vee s}\sum_{j=1}^n \|\sigma_{ij}(u^\varepsilon(r))\|^2_{\mathcal{L}_2(U;L^2(\mathcal{O}))}\mathrm{d}r\bigg)^{p/2}\mathrm{d}t\mathrm{d}s \\
&\le C\int_0^T\int_0^T |t-s|^{-1-\alpha p+p/2-1}\mathbb{E}\int_{s\wedge t}^{t\vee s}\sum_{j=1}^n \|\sigma_{ij}(u^\varepsilon(r))\|^p_{\mathcal{L}_2(U;L^2(\mathcal{O}))}\mathrm{d}r\mathrm{d}t\mathrm{d}s \\
&\le C\int_0^T\int_0^T |t-s|^{-1-\alpha p+p/2-1}\mathbb{E}\int_{s\wedge t}^{t\vee s}\sum_{j=1}^n (1+\|u^\varepsilon(r)\|^{\gamma p}_{L^2(\mathcal{O})})\mathrm{d}r\mathrm{d}t\mathrm{d}s \le C.
\end{aligned}$$

The last step follows from estimate (49) (assuming that $1 \leq \gamma p \leq 4/d$) and Lemma 20, since $1 + \alpha p - p/2 + 1 < 2$ if and only if $\alpha < 1/2$. We conclude that the second term of the right-hand side of

$$v^\varepsilon(t) = v^\varepsilon(0) + \int_0^t \operatorname{div}(A(u^\varepsilon(s))\nabla u^\varepsilon(s))\mathrm{d}s + \int_0^t \sigma(u^\varepsilon(s))\mathrm{d}W(s)$$

is uniformly bounded in $\mathbb{E}|\cdot|_{W^{\alpha,p}(0,T;D(L)')}$ for $\alpha < 1$ and $p \leq (d+2)/(d+1)$, while the third term is uniformly bounded in that norm for $\alpha < 1/2$ and $p \leq 4/(\gamma d)$. In both cases, we can choose α such that $\alpha p > 1$. At this point, we need the condition $\gamma < 1$ if $d = 2$. (The result holds for any space dimension if $\gamma < 2/d$.) Taking into account (33), (v^ε) is bounded in $W^{\alpha,p}(0,T;D(L)')$. The embedding $W^{\alpha,p}(0,T;D(L)') \hookrightarrow C^{0,\beta}([0,T];D(L)')$ for $\beta = \alpha - 1/p > 0$ implies that (v^ε) is bounded in the latter space.

We turn to the estimate of u^ε in the $W^{\alpha,p}(0,T;D(L)')$ norm:

$$\mathbb{E}\|u^\varepsilon\|^p_{W^{\alpha,p}(0,T;D(L)')} \leq C\big(\mathbb{E}\|v^\varepsilon\|^p_{W^{\alpha,p}(0,T;D(L)')} + \varepsilon\mathbb{E}\|L^*LR_\varepsilon(v^\varepsilon)\|^p_{W^{\alpha,p}(0,T;D(L)')}\big).$$

It remains to consider the last term. In view of estimate (15) and the Lipschitz continuity of R_ε with Lipschitz constant C/ε, we obtain

$$\begin{aligned}
&\mathbb{E}|\varepsilon L^*LR_\varepsilon(v^\varepsilon)|^p_{W^{\alpha,p}(0,T;D(L)')} \\
&\quad = \varepsilon^p\mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\|L^*LR_\varepsilon(v^\varepsilon(t)) - L^*LR_\varepsilon(v^\varepsilon(s))\|^p_{D(L)'}\mathrm{d}t\mathrm{d}s \\
&\quad \leq \varepsilon^p C\mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\|R_\varepsilon(v^\varepsilon(t)) - R_\varepsilon(v^\varepsilon(s))\|^p_{D(L)}\mathrm{d}t\mathrm{d}s \\
&\quad \leq \varepsilon^p C\mathbb{E}\int_0^T\int_0^T |t-s|^{-1-\alpha p}\frac{C}{\varepsilon^p}\|v^\varepsilon(t) - v^\varepsilon(s)\|^p_{D(L)'}\mathrm{d}t\mathrm{d}s \\
&\quad = C\mathbb{E}\|v^\varepsilon\|^p_{W^{\alpha,p}(0,T;D(L)')} \leq C.
\end{aligned}$$

Moreover, by (15) and the Lipschitz continuity of R_ε again,

$$\|\varepsilon L^*LR_\varepsilon(v^\varepsilon)\|^p_{L^p(0,T;D(L)')} \leq \varepsilon^p C\|R_\varepsilon(v^\varepsilon)\|^p_{L^p(0,T;D(L))} \leq \varepsilon^p C\|v^\varepsilon\|^p_{L^p(0,T;D(L)')} \leq C,$$

where we used estimate (33). This finishes the proof. □

6.2. **Tightness of the laws of** (u^ε)**.** The tightness is shown in a different sub-Polish space than in Section 5.2:

$$\widetilde{Z}_T := C^0([0,T];D(L)') \cap L^{\rho_1}_w(0,T;W^{1,\rho_1}(\mathcal{O})),$$

endowed with the topology $\widetilde{\mathbb{T}}$ that is the maximum of the topology of $C^0([0,T];D(L)')$ and the weak topology of $L^{\rho_1}_w(0,T;W^{1,\rho_1}(\mathcal{O}))$, recalling that $\rho_1 = (d+2)/(d+1) > 1$.

Lemma 30. *The family of laws of* (u^ε) *is tight in*

$$Z_T := \widetilde{Z}_T \cap L^2(0,T;L^2(\mathcal{O}))$$

with the topology that is the maximum of $\widetilde{\mathbb{T}}$ *and the topology induced by the* $L^2(0,T;L^2(\mathcal{O}))$ *norm.*

Proof. The tightness in $L^2(0,T;L^q(\mathcal{O}))$ for $q < d/(d-1) = 2$ is a consequence of the compact embedding $W^{1,1}(\mathcal{O}) \hookrightarrow L^q(\mathcal{O})$ as well as estimates (47) and (54). In fact, we can extend this result up to $q = 2$ because of the uniform bound of $u_i^\varepsilon \log u_i^\varepsilon$ in $L^\infty(0,T;L^1(\mathcal{O}))$, which originates from the entropy estimate. Indeed, we just apply [3, Prop. 1], using additionally (26) with $a_{i0} > 0$. Then the tightness in $L^2(0,T;L^2(\mathcal{O}))$ follows from Lemma 36. Finally, the tightness in $\widetilde{Z}_T$ is shown as in the proof of Lemma 22 in Appendix B. □

In three space dimensions, we do not obtain tightness in $L^2(0,T;L^2(\mathcal{O}))$ but in the larger space $L^{4/3}(0,T;L^2(\mathcal{O}))$. This follows similarly as in the proof of Lemma 22 taking into account the compact embedding $W^{1,\rho_1}(\mathcal{O}) \hookrightarrow L^2(\mathcal{O})$, which holds as long as $d \le 3$, as well as estimates (50) and (54). Unfortunately, this result seems to be not sufficient to identify the limit of the product $\widetilde{u}_i^\varepsilon \widetilde{u}_j^\varepsilon$. Therefore, we restrict ourselves to the two-dimensional case.

The following result is shown exactly as in Lemma 23.

Lemma 31. *The family of laws of* $(\sqrt{\varepsilon} L^* L R_\varepsilon(v^\varepsilon))$ *is tight in* $Y_T = L^2_w(0,T;D(L)') \cap L^\infty_{w*}(0,T;D(L)')$.

Arguing as in Section 5.3, the Skorokhod–Jakubowski theorem implies the existence of a subsequence, a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$, and, on this space, $(Z_T \times Y_T \times C^0([0,T];U_0))$-valued random variables $(\widetilde{u}^\varepsilon, \widetilde{w}^\varepsilon, \widetilde{W}^\varepsilon)$ and $(\widetilde{u}, \widetilde{w}, \widetilde{W})$ such that $(\widetilde{u}^\varepsilon, \widetilde{w}^\varepsilon, \widetilde{W}^\varepsilon)$ has the same law as $(u^\varepsilon, \sqrt{\varepsilon} L^* L R_\varepsilon(v^\varepsilon), W)$ on $\mathcal{B}(Z_T \times Y_T \times C^0([0,T];U_0))$ and, as $\varepsilon \to 0$ and $\widetilde{\mathbb{P}}$-a.s.,

$$(\widetilde{u}^\varepsilon, \widetilde{w}^\varepsilon, \widetilde{W}^\varepsilon) \to (\widetilde{u}, \widetilde{w}, \widetilde{W}) \quad \text{in } Z_T \times Y_T \times C^0([0,T];U_0).$$

This convergence means that $\widetilde{\mathbb{P}}$-a.s.,

$$\begin{aligned}
\widetilde{u}^\varepsilon &\to \widetilde{u} && \text{strongly in } C^0([0,T];D(L)'), \\
\nabla \widetilde{u}^\varepsilon &\rightharpoonup \nabla \widetilde{u} && \text{weakly in } L^{\rho_1}(Q_T), \\
\widetilde{u}^\varepsilon &\to \widetilde{u} && \text{strongly in } L^2(Q_T), \\
\widetilde{w}^\varepsilon &\rightharpoonup \widetilde{w} && \text{weakly in } L^2(0,T;D(L)'), \\
\widetilde{w}^\varepsilon &\rightharpoonup \widetilde{w} && \text{weakly* in } L^\infty(0,T;D(L)'), \\
\widetilde{W}^\varepsilon &\to \widetilde{W} && \text{strongly in } C^0([0,T];U_0).
\end{aligned}$$

The remainder of the proof is very similar to that one of Section 5.3, using slightly weaker convergence results. The most difficult part is the convergence of the nonlinear term $\nabla(\widetilde{u}_i^\varepsilon \widetilde{u}_j^\varepsilon)$, since the previous convergences do not allow us to perform the limit $\widetilde{u}_i^\varepsilon \nabla \widetilde{u}_j^\varepsilon$ because of $\rho_1 < 2$. The idea is to consider the "very weak" formulation by performing the limit in $\widetilde{u}_i^\varepsilon \widetilde{u}_j^\varepsilon \Delta\phi$ instead of $\nabla(\widetilde{u}_i^\varepsilon \widetilde{u}_j^\varepsilon) \cdot \nabla\phi$ for suitable test functions ϕ. Indeed, let $\phi \in L^\infty(0,T;C_0^\infty(\mathcal{O}))$. Since $\widetilde{u}_i^\varepsilon \to \widetilde{u}$ strongly in $L^2(0,T;L^2(\mathcal{O}))$ $\widetilde{\mathbb{P}}$-a.s., we have

$$\int_0^T \int_{\mathcal{O}} \nabla(\widetilde{u}_i^j \widetilde{u}_j^\varepsilon) \cdot \nabla\phi \mathrm{d}x\mathrm{d}t = -\int_0^T \int_{\mathcal{O}} \widetilde{u}_i^\varepsilon \widetilde{u}_j^\varepsilon \Delta\phi \mathrm{d}x\mathrm{d}t \to -\int_0^T \int_{\mathcal{O}} \widetilde{u}_i \widetilde{u}_j \Delta\phi \mathrm{d}x\mathrm{d}t.$$

It follows from the equivalence of the laws that

$$\widetilde{\mathbb{E}}\bigg(\int_0^T\int_{\mathcal{O}}\widetilde{u}_i^\varepsilon\widetilde{u}_j^\varepsilon\Delta\phi\mathrm{d}x\mathrm{d}t\bigg)^2\le C,$$

and we conclude from Vitali's theorem that

$$\widetilde{E}\bigg|\int_0^T\int_{\mathcal{O}}\big(\widetilde{u}_i^\varepsilon\widetilde{u}_j^\varepsilon-\widetilde{u}_i\widetilde{u}_j\big)(t)\Delta\phi\mathrm{d}x\mathrm{d}t\bigg|\to 0\quad\text{as }\varepsilon\to 0.$$

By density, this convergence holds for all test functions $\phi\in L^\infty(0,T;W^{2,\infty}(\mathcal{O}))$ such that $\nabla\phi\cdot\nu=0$ on $\partial\mathcal{O}$. This ends the proof of Theorem 4.

Remark 32 (Three space dimensions)**.** The three-dimensional case is delicate since u_i^ε lies in a space larger than $L^2(Q_T)$. We may exploit the regularity (51) for $\nabla(u_i^\varepsilon u_j^\varepsilon)$, but this leads only to the existence of random variables $\widetilde{\eta}_{ij}^\varepsilon$ and $\widetilde{\eta}_{ij}$ with $i,j=1,\ldots,n$ and $i\neq j$ on the space $X_T=L_w^{\rho_2}(0,T;L^{\rho_2}(\mathcal{O}))$ such that $\widetilde{\eta}_{ij}^\varepsilon$ and $u_i^\varepsilon u_j^\varepsilon$ have the same law on $\mathcal{B}(X_T)$ and, as $\varepsilon\to 0$,

$$\widetilde{\eta}_{ij}^\varepsilon\rightharpoonup\widetilde{\eta}_{ij}\quad\text{weakly in }X_T.$$

Similar arguments as before lead to the limit

$$\widetilde{\mathbb{E}}\bigg|\int_0^T\int_{\mathcal{O}}\nabla(\widetilde{\eta}_{ij}^\varepsilon-\widetilde{\eta}_{ij})(t)\cdot\nabla\phi(t)\mathrm{d}x\mathrm{d}t\bigg|\to 0,$$

but we cannot easily identify $\widetilde{\eta}_{ij}$ with $\widetilde{u}_i\widetilde{u}_j$. □

7. Discussion of the noise terms

We present some examples of admissible terms $\sigma(u)$. Recall that $(e_k)_{k\in\mathbb{N}}$ is an orthonormal basis of U.

Lemma 33. *The stochastic diffusion*

$$\sigma_{ij}(u)=\delta_{ij}s(u_i)\sum_{\ell=1}^\infty a_\ell(e_\ell,\cdot)_U,\quad s(u_i)=\frac{u_i}{1+u_i^{1/2+\eta}}$$

satisfies Assumption (A5) for $\eta>0$ and $(a_\ell)\in\ell^2(\mathbb{R})$.

Proof. With the entropy density h given by (5), we compute $(\partial h/\partial u_i)(u)=\pi_i\log u_i$ and $(\partial^2h/\partial u_i\partial u_j)(u)=(\pi_i/u_i)\delta_{ij}$. Therefore, by Jensen's inequality and the elementary inequalities $|u_i\log u_i|\le C(1+u_i^{1+\eta})$ for any $\eta>0$ and $|u|\le C(1+h(u))$,

$$\begin{aligned}J_1&:=\bigg\{\int_0^T\sum_{k=1}^\infty\sum_{i,j=1}^n\bigg(\int_{\mathcal{O}}\frac{\partial h}{\partial u_i}(u)\sigma_{ij}(u)e_k\mathrm{d}x\bigg)^2\mathrm{d}s\bigg\}^{1/2}\\&=\bigg\{\sum_{k=1}^\infty a_k^2\int_0^T\sum_{i=1}^n\bigg(\int_{\mathcal{O}}\pi_i\frac{u_i\log u_i}{1+u_i^{1/2+\eta}}\mathrm{d}x\bigg)^2\mathrm{d}s\bigg\}^{1/2}\\&\le C\bigg\{\sum_{i=1}^n\int_0^T\bigg(\int_{\mathcal{O}}\frac{1+u_i^{1+\eta}}{1+u_i^{1/2+\eta}}\mathrm{d}x\bigg)^2\mathrm{d}s\bigg\}^{1/2}\end{aligned}$$

$$\le C\bigg\{\sum_{i=1}^n\int_0^T\int_{\mathcal{O}}\bigg(\frac{1+u_i^{1+\eta}}{1+u_i^{1/2+\eta}}\bigg)^2 \mathrm{d}x\mathrm{d}s\bigg\}^{1/2}$$
$$\le C\bigg\{\sum_{i=1}^n\int_0^T\int_{\mathcal{O}}(1+u_i)\mathrm{d}x\mathrm{d}s\bigg\}^{1/2} \le C\bigg(1+\int_0^T\int_{\mathcal{O}}h(u)\mathrm{d}x\mathrm{d}s\bigg).$$

The second condition in Assumption (A5) becomes

$$J_2 := \int_0^T\sum_{k=1}^\infty\int_{\mathcal{O}}\operatorname{tr}\big[(\sigma(u)e_k)^T h''(u)\sigma(u)e_k\big]\mathrm{d}x\mathrm{d}s$$
$$= \sum_{k=1}^\infty a_k^2\sum_{i=1}^n\int_0^T\int_{\mathcal{O}}\frac{\pi_i u_i}{(1+u_i^{1/2+\eta})^2}\mathrm{d}x\mathrm{d}s \le C(\mathcal{O},T).$$

Thus, Assumption (A5) is satisfied. □

The proof shows that J_1 can be estimated if $s(u_i)^2\log(u_i)^2$ is bounded from above by $C(1+h(u))$. This is the case if $s(u_i)$ behaves like u_i^α with $\alpha<1/2$. Furthermore, J_2 can be estimated if $s(u_i)^2/u_i$ is bounded, which is possible if $s(u_i)=u_i^\alpha$ with $\alpha\ge 1/2$. Thus, to both satisfy the growth restriction and avoid the singularity at $u_i=0$, we have chosen σ_{ij} as in Lemma 33. This example is rather artificial. To include more general choices, we generalize our approach. In fact, it is sufficient to estimate the integrals in inequality (23) in such a way that the entropy inequality of Proposition 15 holds. The idea is to exploit the gradient bound for u_i for the estimatation of J_1 and J_2.

Consider a trace-class, positive, and symmetric operator Q on $L^2(\mathcal{O})$ and the space $U = Q^{1/2}(L^2(\mathcal{O}))$, equipped with the norm $\|Q^{1/2}(\cdot)\|_{L^2(\mathcal{O})}$. We will work in the following with an U-cylindrical Wiener process W^Q. This setting is equivalent to a spatially colored noise on $L^2(\mathcal{O})$ in the form of a Q-Wiener process (with $Q\neq\mathrm{Id}$). The latter viewpoint provides, in our opinion, a more intuitive insight. In particular, the operator Q is constructed from the eigenfunctions and eigenvalues described below.

Let $(\eta_k)_{k\in\mathbb{N}}$ be a basis of $L^2(\mathcal{O})$, consisting of the normalized eigenfunctions of the Laplacian subject to Neumann boundary conditions with eigenvalues $\lambda_k\ge 0$, and set $a_k=(1+\lambda_k)^{-\rho}$ for some $\rho>0$ such that $\sum_{k=1}^\infty a_k^2\|\eta_k\|_{L^\infty(\mathcal{O})}^2<\infty$. Since $\lambda_k\le Ck^{2/d}$ [28, Corollary 2] and $\|\eta_k\|_{L^\infty(\mathcal{O})}\le Ck^{(d-1)/2}$ [23, Theorem 1], we may choose $\rho>(d/2)^2$. Considering a sequence of independent Brownian motions $(W_1^k,\ldots,W_n^k)_{k\in\mathbb{N}}$, we assume the noise to be of the form $W^Q=(W_1^Q,\ldots,W_n^Q)$, where

$$W_j^Q(t)=\sum_{k=1}^\infty a_k e_k W_j^k(t),\quad j=1,\ldots,n,\ t>0,$$

and $(e_k)_{k\in\mathbb{N}}=(a_k\eta_k)_{k\in\mathbb{N}}$ is a basis of $U=Q^{1/2}(L^2(\mathcal{O}))$.

Lemma 34. *For the SKT model with self-diffusion, let $\sigma_{ij}(u)=\delta_{ij}u_i^\alpha$ for $1/2\le\alpha\le 1$, $i,j=1,\ldots,n$, interpreted as a map from $L^2(\mathcal{O})$ to $\mathcal{L}_2(H^\beta(\mathcal{O});L^2(\mathcal{O}))$, where $\beta>\rho$. Then the entropy inequality* (29) *holds, i.e., σ_{ij} is admissible for Theorem 3.*

Proof. We can write inequality (23) for $0 < T < T_R$ as

$$
\begin{aligned}
(55)\qquad & \mathbb{E}\sup_{0<t<T}\int_{\mathcal{O}} h(u^\varepsilon(t))\mathrm{d}x + \frac{\varepsilon}{2}\mathbb{E}\sup_{0<t<T}\|Lw^\varepsilon(t)\|^2_{L^2(\mathcal{O})} \\
&\quad + \mathbb{E}\sup_{0<t<T}\int_0^t\int_{\mathcal{O}}\nabla w^\varepsilon(s):B(w^\varepsilon)\nabla w^\varepsilon(s)\mathrm{d}x\mathrm{d}s - \mathbb{E}\int_{\mathcal{O}} h(u^0)\mathrm{d}x \\
&\le \mathbb{E}\sup_{0<t<T}\bigg\{\int_0^t\sum_{k=1}^\infty\sum_{i,j=1}^n\bigg(\int_{\mathcal{O}}\pi_i\log u_i^\varepsilon(s)\sigma_{ij}(u^\varepsilon(s))e_k\mathrm{d}x\bigg)^2\mathrm{d}s\bigg\}^{1/2} \\
&\quad + \frac{1}{2}\mathbb{E}\sup_{0<t<T}\sum_{k=1}^\infty\sum_{i=1}^n\int_0^t\int_{\mathcal{O}}(\sigma_{ii}(u^\varepsilon)e_k\frac{\pi_i}{u_i^\varepsilon}\sigma_{ii}(u^\varepsilon)e_k\mathrm{d}x\mathrm{d}s \\
&=: J_3 + J_4,
\end{aligned}
$$

recalling that $w^\varepsilon = R_\varepsilon(v^\varepsilon)$ and $u^\varepsilon = u(w^\varepsilon)$. We simplify J_3 and J_4, using the definition $e_k = a_k\eta_k$:

$$
\begin{aligned}
J_3 &= \mathbb{E}\sup_{0<t<T}\bigg\{\sum_{k=1}^\infty a_k^2\int_0^t\sum_{i=1}^n\pi_i^2\bigg(\int_{\mathcal{O}} u_i^\varepsilon(s)^\alpha\log u_i^\varepsilon(s)\eta_k\mathrm{d}x\bigg)^2\mathrm{d}s\bigg\}^{1/2} \\
&\le C\mathbb{E}\sup_{0<t<T}\bigg\{\sum_{k=1}^\infty a_k^2\int_{\mathcal{O}}\eta_k^2\mathrm{d}x\int_0^t\sum_{i=1}^n\int_{\mathcal{O}}(u_i^\varepsilon(s)^\alpha\log u_i^\varepsilon(s))^2\mathrm{d}x\mathrm{d}s\bigg\}^{1/2} \\
&\le C\sum_{i=1}^n\mathbb{E}\|(u_i^\varepsilon)^\alpha\log u_i^\varepsilon\|_{L^2(0,T;L^2(\mathcal{O}))}, \\
J_4 &= \sum_{k=1}^\infty a_k^2\mathbb{E}\sup_{0<t<T}\sum_{i=1}^n\pi_i\int_0^t\int_{\mathcal{O}}(u_i^\varepsilon)^{2\alpha}(u_i^\varepsilon)^{-1}\eta_k^2\mathrm{d}x\mathrm{d}s \\
&\le C\sum_{k=1}^\infty a_k^2\|\eta_k\|^2_{L^\infty(\mathcal{O})}\sum_{i=1}^n\mathbb{E}\|(u_i^\varepsilon)^{2\alpha-1}\|_{L^1(0,T;L^1(\mathcal{O}))} \\
&\le C\sum_{i=1}^n\mathbb{E}\|(u_i^\varepsilon)^{2\alpha-1}\|_{L^1(0,T;L^1(\mathcal{O}))}.
\end{aligned}
$$

The last inequality follows from our assumption on (a_k). By (28), we can estimate the integrand of the third integral on the left-hand side of (55) according to

$$
\nabla w^\varepsilon : B(w^\varepsilon)\nabla w^\varepsilon \ge 2\sum_{i=1}^n\pi_i a_{ii}|\nabla u^\varepsilon|^2.
$$

Hence, because of $|u| \le C(1 + h(u))$, we can formulate (55) as

$$
\mathbb{E}\sup_{0<t<T}\|h(u^\varepsilon(t))\|_{L^1(\mathcal{O})} + \frac{\varepsilon}{2}\mathbb{E}\sup_{0<t<T}\|Lw^\varepsilon(t)\|^2_{L^2(\mathcal{O})} + C\mathbb{E}\|\nabla u^\varepsilon(s)\|^2_{L^2(0,T;L^2(\mathcal{O}))}
$$

$$\le C + C\sum_{i=1}^n \mathbb{E}\|(u_i^\varepsilon)^\alpha \log u_i^\varepsilon\|_{L^2(0,T;L^2(\mathcal{O}))} + C\sum_{i=1}^n \mathbb{E}\|(u_i^\varepsilon)^{2\alpha-1}\|_{L^1(0,T;L^1(\mathcal{O}))}.$$

It is sufficient to continue with the case $\alpha = 1$, since the proof for $\alpha < 1$ follows from the case $\alpha = 1$. Then, using $|u^\varepsilon| \le C(1 + h(u^\varepsilon))$,

$$\begin{aligned} &\mathbb{E}\|h(u^\varepsilon)\|_{L^\infty(0,T;L^1(\mathcal{O}))} + \mathbb{E}\|u^\varepsilon\|_{L^\infty(0,T;L^1(\mathcal{O}))} \\ &\quad + \varepsilon\mathbb{E}\|Lw^\varepsilon\|^2_{L^\infty(0,T;L^2(\mathcal{O}))} + \mathbb{E}\|\nabla u^\varepsilon\|^2_{L^2(0,T;L^2(\mathcal{O}))} \\ &\le C + C\sum_{i=1}^n \mathbb{E}\|u_i^\varepsilon \log u_i^\varepsilon\|_{L^2(0,T;L^2(\mathcal{O}))} + C\mathbb{E}\|u^\varepsilon\|_{L^1(0,T;L^1(\mathcal{O}))}. \end{aligned} \tag{56}$$

Now, we use the following lemma which is proved in Appendix A.

Lemma 35. *Let $d \ge 2$ and let $v \in L^2(0,T;H^1(\mathcal{O}))$ satisfy $v\log v \in L^\infty(0,T;L^1(\mathcal{O}))$. Then for any $\delta > 0$, there exists $C(\delta) > 0$ such that*

$$\begin{aligned} \|v\log v\|_{L^2(0,T;L^2(\mathcal{O}))} &\le \delta\big(\|v\log v\|_{L^1(0,T;L^1(\mathcal{O}))} + \|v\|_{L^\infty(0,T;L^1(\mathcal{O}))} + \|\nabla v\|^2_{L^2(0,T;L^2(\mathcal{O}))}\big) \\ &\quad + C(\delta)\|v\|_{L^1(0,T;L^1(\mathcal{O}))}. \end{aligned}$$

It follows from (56) that, for any $\delta > 0$,

$$\begin{aligned} &\mathbb{E}\|h(u^\varepsilon)\|_{L^\infty(0,T;L^1(\mathcal{O}))} + \mathbb{E}\|u^\varepsilon\|_{L^\infty(0,T;L^1(\mathcal{O}))} \\ &\quad + \varepsilon\mathbb{E}\|Lw^\varepsilon\|^2_{L^\infty(0,T;L^2(\mathcal{O}))} + \mathbb{E}\|\nabla u^\varepsilon\|^2_{L^2(0,T;L^2(\mathcal{O}))} \\ &\le C + C(\delta)\mathbb{E}\|u^\varepsilon\|_{L^1(0,T;L^1(\mathcal{O}))} + \delta C\sum_{i=1}^n \mathbb{E}\|u_i^\varepsilon \log u_i^\varepsilon\|_{L^1(0,T;L^1(\mathcal{O}))} \\ &\quad + \delta C\big(\mathbb{E}\|u^\varepsilon\|_{L^\infty(0,T;L^1(\mathcal{O}))} + \mathbb{E}\|\nabla u^\varepsilon\|^2_{L^2(0,T;L^2(\mathcal{O}))}\big). \end{aligned}$$

For sufficiently small $\delta > 0$, the last terms on the right-hand side can be absorbed by the corresponding terms on the left-hand side, leading to

$$\begin{aligned} &\mathbb{E}\|h(u^\varepsilon)\|_{L^\infty(0,T;L^1(\mathcal{O}))} + \mathbb{E}\|u^\varepsilon\|_{L^\infty(0,T;L^1(\mathcal{O}))} \\ &\quad + \varepsilon\mathbb{E}\|Lw^\varepsilon\|^2_{L^\infty(0,T;L^2(\mathcal{O}))} + \mathbb{E}\|\nabla u^\varepsilon\|^2_{L^2(0,T;L^2(\mathcal{O}))} \\ &\le C + C\int_0^T \|u^\varepsilon\|_{L^\infty(0,t;L^1(\mathcal{O}))}\mathrm{d}t \quad \text{for all } T > 0. \end{aligned}$$

Gronwall's lemma ends the proof. ☐

In the case without self-diffusion, we have an $H^1(\mathcal{O})$ estimate for $(u_i^\varepsilon)^{1/2}$ only, and it can be seen that stochastic diffusion terms of the type $\delta_{ij}u_i^\alpha$ for $\alpha > 1/2$ are not admissible. However, we may choose $\sigma_{ij}(u)e_k = \delta_{ij}u_i^\alpha(1 + (u_i^\varepsilon)^\beta)^{-1}a_k\eta_k$ for $1/2 \le \alpha < 1$ and $\beta \ge \alpha/2$.

Appendix A. Proofs of some lemmas

A.1. **Proof of Lemma 13.** The operator equation $\mathrm{D}R_\varepsilon[w](a) = b$ can be written as $a = \mathrm{D}Q_\varepsilon[w](b) = u'(w)b + \varepsilon L^*Lb$. Hence,

$$\int_{\mathcal{O}} a : b\mathrm{d}x = \int_{\mathcal{O}} u'(w)b : b\mathrm{d}x + \varepsilon \int_{\mathcal{O}} Lb : Lb\mathrm{d}x. \tag{57}$$

The matrix $u'(w) = (h'')^{-1}(u(w))$ is symmetric and positive semidefinite (since h is convex). Thus, the square root operator $\sqrt{u'(w)}$ exists and is symmetric. This shows that

$$\begin{aligned} u'(w)b : b &= \sqrt{u'(w)}\sqrt{u'(w)}b : b = \operatorname{tr}\big[\big(\sqrt{u'(w)}\sqrt{u'(w)}b\big)^T b\big] \\ &= \operatorname{tr}\big[\big(\sqrt{u'(w)}b\big)^T\big(\sqrt{u'(w)}b\big)\big] = \big\|\sqrt{u'(w)}b\big\|_F^2. \end{aligned}$$

Furthermore, by the Cauchy–Schwarz inequality $\operatorname{tr}[A^TB] \le \operatorname{tr}[A^TA]\operatorname{tr}[B^TB]$ and the property $\operatorname{tr}[AB] = \operatorname{tr}[BA]$ for matrices A and B,

$$\begin{aligned} a : b &= \operatorname{tr}\big[a^T\sqrt{u'(w)}^{\,-1}\sqrt{u'(w)}b\big] \\ &\le \operatorname{tr}\big[\big(\sqrt{u'(w)}b\big)^T\sqrt{u'(w)}b\big]^{1/2}\operatorname{tr}\big[\big(a^T\sqrt{u'(w)}^{\,-1}\big)^T a^T\sqrt{u'(w)}^{\,-1}\big]^{1/2} \\ &\le \frac{1}{2}\operatorname{tr}\big[\big(\sqrt{u'(w)}b\big)^T\sqrt{u'(w)}b\big] + \frac{1}{2}\operatorname{tr}\big[\big(\sqrt{u'(w)}^{\,-1}a\big)\big(a^T\sqrt{u'(w)}^{\,-1}\big)\big] \\ &= \frac{1}{2}\big\|\sqrt{u'(w)}b\big\|_F^2 + \frac{1}{2}\operatorname{tr}\big[\big(a^T\sqrt{u'(w)}^{\,-1}\big)\big(\sqrt{u'(w)}^{\,-1}a\big)\big] \\ &= \frac{1}{2}\big\|\sqrt{u'(w)}b\big\|_F^2 + \frac{1}{2}\operatorname{tr}[a^Tu'(w)^{-1}a]. \end{aligned}$$

Inserting these relations into (57) leads to

$$\begin{aligned} \int_{\mathcal{O}}\|\sqrt{u'(w)}b\|_F^2\mathrm{d}x + \varepsilon\int_{\mathcal{O}} Lb : Lb\mathrm{d}x &= \int_{\mathcal{O}} a : b\mathrm{d}x \\ &\le \frac{1}{2}\int_{\mathcal{O}}\|\sqrt{u'(w)}b\|_F^2\mathrm{d}x + \frac{1}{2}\int_{\mathcal{O}}\operatorname{tr}[a^Tu'(w)^{-1}a]\mathrm{d}x \end{aligned} \tag{58}$$

and consequently,

$$\int_{\mathcal{O}}\|\sqrt{u'(w)}b\|_F^2\mathrm{d}x \le \int_{\mathcal{O}}\operatorname{tr}[a^Tu'(w)^{-1}a]\mathrm{d}x.$$

Together with (58) we obtain the statement.

A.2. **Proof of Lemma 14.** It follows from the convexity of h that

$$h(v^0) \ge h(u(R_\varepsilon(v^0))) + \big(v^0 - u(R_\varepsilon(v^0))\big)\cdot h'(u(R_\varepsilon(v^0))).$$

Since R_ε and Q_ε are inverse to each other, we can replace v^0 by $Q_\varepsilon(R_\varepsilon(v^0)) = u(R_\varepsilon(v^0)) + \varepsilon L^*LR_\varepsilon(v^0)$:

$$\begin{aligned} h(v^0) &\ge h(u(R_\varepsilon(v^0))) + \big\langle u(R_\varepsilon(v^0)) + \varepsilon L^*LR_\varepsilon(v^0) - u(R_\varepsilon(v^0)), h'(u(R_\varepsilon(v^0)))\big\rangle_{D(L)',D(L)} \\ &= h(u(R_\varepsilon(v^0))) + \varepsilon\langle L^*LR_\varepsilon(v^0), R_\varepsilon(v^0)\rangle_{D(L)',D(L)}. \end{aligned}$$

We find after an integration that

$$\int_{\mathcal{O}} h(v^0)\mathrm{d}x \geq \int_{\mathcal{O}} h(u(R_\varepsilon(v^0)))\mathrm{d}x + \varepsilon \int_{\mathcal{O}} LR_\varepsilon(v^0)\cdot LR_\varepsilon(v^0)\mathrm{d}x,$$

which yields the statement.

A.3. **Proof of Lemma 20.** We show that

$$I := \int_0^T \int_0^T |t-s|^{-\delta} \int_{s\wedge t}^{t\vee s} g(r)\mathrm{d}r\mathrm{d}t\mathrm{d}s < \infty.$$

A change of the integration domain and an integration by parts lead to

$$\begin{aligned} I &= 2\int_0^T \int_s^T (t-s)^{-\delta}\bigg(\int_s^t g(r)dr\bigg)\mathrm{d}t\mathrm{d}s \\ &= -\frac{2}{1-\delta}\int_0^T \int_s^T (t-s)^{1-\delta} g(t)\mathrm{d}t\mathrm{d}s + \frac{2}{1-\delta}\int_0^T (T-s)^{1-\delta}\int_s^t g(r)\mathrm{d}r\mathrm{d}s, \end{aligned} \tag{59}$$

observing that $\lim_{t\to s}(t-s)^{1-\delta}\int_s^t g(r)\mathrm{d}r = 0$ for $1-\delta > -1$, since the integrability of g implies that $\lim_{t\to s}(t-s)^{-1}\int_s^t g(r)\mathrm{d}r = g(s)$ for a.e. s. The result follows as the integrals on the right-hand side of (59) are finite.

A.4. **Proof of Lemma 35.** We use the interpolation inequality with $1/2 = \theta_1 + (1-\theta_1)/(2p)$ and some $1 < p < d/(d-2)$ (and $p > 1$ if $d = 2$) as well as the Young inequality with $\delta > 0$:

$$\begin{aligned} \|v\log v\|_{L^2(0,T;L^2(\mathcal{O}))} &\leq \bigg(\int_0^T \|v\log v\|_{L^1(\mathcal{O})}^{2\theta_1}\|v\log v\|_{L^{2p}(\mathcal{O})}^{2(1-\theta_1)}\mathrm{d}t\bigg)^{1/2} \\ &\leq \bigg(C(\delta)^2\int_0^T \|v\log v\|_{L^1(\mathcal{O})}^2\mathrm{d}t + \delta^2\int_0^T \|v\log v\|_{L^{2p}(\mathcal{O})}^2\mathrm{d}t\bigg)^{1/2} \\ &\leq C(\delta)\|v\log v\|_{L^2(0,T;L^1(\mathcal{O}))} + \delta\|v\log v\|_{L^2(0,T;L^{2p}(\mathcal{O}))}. \end{aligned} \tag{60}$$

The first term on the right-hand side is estimated in a similar way as before, where $\eta > 0$:

$$\begin{aligned} \|v\log v\|_{L^2(0,T;L^1(\mathcal{O}))} &\leq \|v\log v\|_{L^\infty(0,T;L^1(\mathcal{O}))}^{1/2}\|v\log v\|_{L^1(0,T;L^1(\mathcal{O}))}^{1/2} \\ &\leq \eta\|v\log v\|_{L^\infty(0,T;L^1(\mathcal{O}))} + C(\eta)\|v\log v\|_{L^1(0,T;L^1(\mathcal{O}))}. \end{aligned} \tag{61}$$

For the second term on the right-hand side of (60), we introduce the function $g(v) = \max\{2, v\log v\}$ for $v \geq 0$. Then $g \in C^1([0,\infty))$. (The function $v \mapsto v\log v$ is not C^1 at $v = 0$, therefore we need to truncate.) We use the Sobolev inequality:

$$\begin{aligned} \|v\log v\|_{L^2(0,T;L^{2p}(\mathcal{O}))} &\leq \|g(v)\|_{L^2(0,T;L^{2p}(\mathcal{O}))} \leq C\|g(v)\|_{L^2(0,T;W^{1,q}(\mathcal{O}))} \\ &\leq C\big(\|g(v)\|_{L^2(0,T;L^q(\mathcal{O}))} + \|\nabla g(v)\|_{L^2(0,T;L^q(\mathcal{O}))}\big), \end{aligned}$$

where $q = 2dp/(d+2p)$. The condition $p < d/(d-2)$ guarantees that $q < 2$, while $d \geq 2$ yields $q > 1$; thus $q \in (1,2)$. Applying the Gagliardo–Nirenberg inequality, combined with the Poincaré–Wirtinger inequality, with $\theta_2 = d(q-1)/(d(q-1)+q) \le 1$, and then the Young inequality, we find that

$$\begin{aligned}\|g(v)\|_{L^q(\mathcal{O})} &\le C\|\nabla g(v)\|_{L^q(\mathcal{O})}^{\theta_2}\|g(v)\|_{L^1(\mathcal{O})}^{1-\theta_2} + \|g(v)\|_{L^1(\mathcal{O})} \\ &\le \|\nabla g(v)\|_{L^q(\mathcal{O})} + C\|g(v)\|_{L^1(\mathcal{O})} \le \|\nabla g(v)\|_{L^q(\mathcal{O})} + C\big(1 + \|v\log v\|_{L^1(\mathcal{O})}\big).\end{aligned}$$

This yields

$$\|v\log v\|_{L^2(0,T;L^{2p}(\mathcal{O}))} \le C + C\|\nabla g(v)\|_{L^2(0,T;L^q(\mathcal{O}))} + C\|v\log v\|_{L^2(0,T;L^1(\mathcal{O}))}. \tag{62}$$

The last term is estimated as in (61). We consider the norm of $\nabla g(v) = 1_{\{v\log v>2\}}(1+\log v)\nabla v$. For this, we observe that $1_{\{v\log v>2\}}\log v \le C(1+v^\gamma)$ for some $0<\gamma<(2-q)/(2q)$ and use the Hölder inequality:

$$\begin{aligned}\|\nabla g(v)\|_{L^q(\mathcal{O})} &\le \|(1+v^\gamma)\nabla v\|_{L^q(\mathcal{O})} \le \big(1 + \|v^\gamma\|_{L^{2q/(2-q)}(\mathcal{O})}\big)\|\nabla v\|_{L^2(\mathcal{O})} \\ &\le C\big(1 + \|v\|_{L^1(\mathcal{O})}^{(2-q)/(2q)}\big)\|\nabla v\|_{L^2(\mathcal{O})},\end{aligned}$$

since the property $2\gamma q/(2-q) < 1$ gives $v^{2\gamma q/(2-q)} \le C(1+v)$ for $v \ge 0$. Consequently, by Young's inequality,

$$\|\nabla g(v)\|_{L^q(\mathcal{O})} \le C\big(1 + \|v\|_{L^1(\mathcal{O})}^{(2-q)/q} + \|\nabla v\|_{L^2(\mathcal{O})}^2\big),$$

and an integration over time gives

$$\begin{aligned}\|\nabla g(v)\|_{L^2(0,T;L^q(\mathcal{O}))} &\le C\big(1 + \|v\|_{L^\infty(0,T;L^1(\mathcal{O}))}^{(2-q)/q} + \|\nabla v\|_{L^2(0,T;L^2(\mathcal{O}))}^2\big) \\ &\le C\big(1 + \|v\|_{L^\infty(0,T;L^1(\mathcal{O}))} + \|\nabla v\|_{L^2(0,T;L^2(\mathcal{O}))}^2\big),\end{aligned}$$

where we used $(2-q)/q < 1$. Thus, (62) becomes

$$\begin{aligned}\|v\log v\|_{L^2(0,T;L^{2p}(\mathcal{O}))} &\le C + C\|\nabla v\|_{L^2(0,T;L^2(\mathcal{O}))}^2 + C\|v\|_{L^\infty(0,T;L^1(\mathcal{O}))} \\ &\quad + C\|v\log v\|_{L^2(0,T;L^1(\mathcal{O}))}.\end{aligned}$$

It remains to insert (61) and the previous estimate into (60) to conclude that

$$\begin{aligned}\|v\log v\|_{L^2(0,T;L^2(\mathcal{O}))} &\le \eta C(\delta)\|v\log v\|_{L^\infty(0,T;L^1(\mathcal{O}))} + C(\delta,\eta)\|v\log v\|_{L^1(0,T;L^1(\mathcal{O}))} \\ &\quad + \delta C\big(\|\nabla v\|_{L^2(0,T;L^2(\mathcal{O}))}^2 + \|v\|_{L^\infty(0,T;L^1(\mathcal{O}))}\big).\end{aligned}$$

Choosing first $\delta > 0$ and then $\eta > 0$ sufficiently small finishes the proof.

Appendix B. Tightness criterion

Lemma 36 (Tightness criterion). *Let $\mathcal{O} \subset \mathbb{R}^d$ $(d \ge 1)$ be a bounded domain with Lipschitz boundary and let $T > 0$, $p,q,r \ge 1$, $\alpha \in (0,1)$ if $r \ge p$ and $\alpha \in (1/r - 1/p, 1)$ if $r < p$. Let $s \ge 1$ be such that the embedding $W^{1,q}(\mathcal{O}) \hookrightarrow L^s(\mathcal{O})$ is compact, and let Y be a Banach space such that the embedding $L^s(\mathcal{O}) \hookrightarrow Y$ is continuous. Furthermore, let $(u_n)_{n\in\mathbb{N}}$ be a sequence of functions such that there exists $C > 0$ such that for all $n \in \mathbb{N}$,*

$$\mathbb{E}\|u_n\|_{L^p(0,T;W^{1,q}(\mathcal{O}))} + \mathbb{E}\|u_n\|_{W^{\alpha,r}(0,T;Y))} \le C.$$

Then the laws of (u_n) are tight in $L^p(0,T;L^s(\mathcal{O}))$ if $q \le d$ and in $L^p(0,T;C^0(\overline{\mathcal{O}}))$ if $q > d$. If $p = \infty$, the space $L^p(0,T;\cdot)$ is replaced by $C^0([0,T];\cdot)$.

Proof. By Theorem 3 and Lemma 5 of [38], the set

$$\begin{aligned} B_R = \big\{ u_n \in L^p(0,T;W^{1,q}(\mathcal{O})) \cap W^{\alpha,r}(0,T;Y) : \\ \|u_n\|_{L^p(0,T;W^{1,q}(\mathcal{O}))} \le R \text{ and } \|u_n\|_{W^{\alpha,r}(0,T;Y)} \le R \big\} \end{aligned}$$

is relatively compact in $L^p(0,T;L^s(\mathcal{O}))$. We deduce from Chebyshev's inequality that

$$\begin{aligned} \mathbb{P}(B_R^c) &\le \mathbb{P}(\|u_n\|_{L^p(0,T;W^{1,q}(\mathcal{O}))} > R) + \mathbb{P}(\|u_n\|_{W^{\alpha,r}(0,T;Y)} > R) \\ &\le \frac{1}{R}\big(\mathbb{E}\|u_n\|_{L^p(0,T;W^{1,q}(\mathcal{O}))} + \mathbb{E}\|u_n\|_{W^{\alpha,r}(0,T;Y)}\big) \le \frac{C}{R}. \end{aligned}$$

The definition of tightness finishes the proof. □

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] H. Amann. Dynamic theory of quasilinear parabolic systems. III. Global existence. *Math. Z.* 202 (1989), 219–250.

[2] V. Bansaye, A. Moussa, and F. Muñoz-Hernández. Stability of a cross-diffusion system and approximation by repulsive random walks: a duality approach. Submitted for publication, 2021. arXiv:2109.07146.

[3] M. Braukhoff, X. Chen, and A. Jüngel. Corrigendum: Cross diffusion preventing blow up in the two-dimensional Keller-Segel model. *SIAM J. Math. Anal.* 52 (2020), 2198–2200.

[4] Z. Brzeźniak, E. Hausenblas, and E. Motyl. Uniqueness in law of the stochastic convolution process driven by Lévy noise. *Electron. J. Probab.* 18 (2013), 1–15.

[5] Z. Brzeźniak and E. Motyl. The existence of martingale solutions to the stochastic Boussinesq equations. *Global Stoch. Anal.* 1 (2014), 175–216.

[6] Z. Brzeźniak and M. Ondreját. Stochastic wave equations with values in Riemanninan manifolds. *Stochastic Partial Differential Equations and Applications, Quad. Mat.* 25 (2010), 65–97.

[7] Z. Brzeźniak and M. Ondreját. Stochastic geometric wave equations with values in compact Riemannian homogeneous spaces. *Ann. Prob.* 41 (2013), 1938–1977.

[8] L. Chen, E. Daus, A. Holzinger, and A. Jüngel. Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems. *J. Nonlin. Sci.* 31 (2021), no. 94, 38 pages.

[9] L. Chen and A. Jüngel. Analysis of a multi-dimensional parabolic population model with strong cross-diffusion. *SIAM J. Math. Anal.* 36 (2004), 301–322.

[10] L. Chen and A. Jüngel. Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Diff. Eqs.* 224 (2006), 39–59.

[11] X. Chen, E. Daus, and A. Jüngel. Global existence analysis of cross-diffusion population systems for multiple species. *Arch. Ration. Mech. Anal.* 227 (2018), 715–747.

[12] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Second edition. Cambridge University Press, Cambridge, 2014.

[13] E. Daus, L. Desvillettes, and H. Dietert. About the entropic structure of detailed balanced multi-species cross-diffusion equations. *J. Diff. Eqs.* 266 (2019), 3861–3882.

[14] A. Debussche, N. Glatt-Holtz, and R. Temam. Local martingale and pathwise solutions for an abstract fluids model. *Physica D* 240 (2011), 1123–1144.

[15] L. Desvillettes, T. Lepoutre, and A. Moussa. Entropy, duality, and cross diffusion. *SIAM J. Math. Anal.* 46 (2014), 820–853.

[16] L. Desvillettes, T. Lepoutre, A. Moussa, and A. Trescases. On the entropic structure of reaction-cross diffusion systems. *Commun. Partial Diff. Eqs.* 40 (2015), 1705–1747.

[17] P. Deuring. An initial-boundary value problem for a certain density-dependent diffusion system. *Math. Z.* 194 (1987), 375–396.

[18] G. Dhariwal, A. Jüngel, and N. Zamponi. Global martingale solutions for a stochastic population cross-diffusion system. *Stoch. Process. Appl.* 129 (2019), 3792–3820.

[19] H. Dietert and A. Moussa. Persisting entropy structure for nonlocal cross-diffusion systems. Submitted for publication, 2021. arXiv:2101.02893.

[20] P. Drábek and J. Milota. *Methods of Nonlinear Analysis: Applications to Differential Equations.* 2dn edition. Springer, Basel, 2013.

[21] M. Dreher. Analysis of a population model with strong cross-diffusion in unbounded domains. *Proc. Roy. Soc. Edinb. Sec. A* 138 (2008), 769–786.

[22] G. Galiano, M. Garźon, and A. Jüngel. Semi-discretization in time and numerical convergence of solutions of a nonlinear cross-diffusion population model. *Numer. Math.* 93 (2003), 655–673.

[23] D. Grieser. Uniform bounds for eigenfunctions of the Laplacian on manifolds with boundary. *Commun. Partial Diff. Eqs.* 27 (2002), 1283–1299.

[24] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity* 28 (2015), 1963–2001.

[25] A. Jüngel. *Entropy Methods for Diffusive Partial Differential Equations.* Springer Briefs Math., Springer, 2016.

[26] J. Kim. Smooth solutions to a quasi-linear system of diffusion equations for a certain population model. *Nonlin. Anal.* 8 (1984), 1121–1144.

[27] S. Krein, J. I. Petunin, and E. Semenov. *Interpolation of Linear Operators.* Amer. Math. Soc., Providence, 1982.

[28] P. Kröger. Upper bounds for the Neumann eigenvalues on a bounded domain in Euclidean spaces. *J. Funct. Anal.* 106 (1992), 353–357.

[29] N. Krylov. A relatively short proof of Itô's formula for SPDEs and its applications. *Stoch. Partial Diff. Eqs.: Anal. Comput.* 1 (2013), 152–174.

[30] C. Kuehn and A. Neamţu. Pathwise mild solutions for quasilinear stochastic partial differential equations. *J. Diff. Eqs.* 269 (2020), 2185–2227.

[31] T. Lepoutre and A. Moussa. Entropic structure and duality for multiple species cross-diffusion systems. *Nonlin. Anal.* 159 (2017), 298–315.

[32] W. Liu and M. Röckner. *Stochastic Partial Differential Equations: an Introduction.* Springer, Cham, 2015.

[33] Y. Lou, W.-M. Ni, and Y. Wu. On the global existence of a cross-diffusion system. *Discrete Contin. Dyn. Syst.* 4 (1998), 193–203.

[34] M. Métivier. *Stochastic Partial Differential Equations in Infinite Dimensional Spaces.* Scuola Normale Superiore, Pisa, 1988.

[35] A. Moussa. From nonlocal to classical Shigesada–Kawasaki–Teramoto systems: triangular case with bounded coefficients. *SIAM J. Math. Anal.* 52 (2020), 42–64.

[36] C. Prévôt and M. Röckner. *A Concise Course on Stochastic Partial Differential Equations.* Lecture Notes Math. 1905. Springer, Berlin, 2007.

[37] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segregation of interacting species. *J. Theor. Biol.* 79 (1979), 83–99.

[38] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Math. Pura. Appl.* 146 (1987), 65–96.

Mathematisches Institut, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany
E-mail address: marcel.braukhoff@hhu.de

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria; Department of Statistics and Operations Research, University of Vienna, Kolingasse 14–16, 1090 Vienna, Austria
E-mail address: fhuber@univie.ac.at

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
E-mail address: juengel@tuwien.ac.at